

**Studien zur Messung des  
 $t\bar{t}Z$ -Produktionswirkungsquerschnitts im 1-Lepton-Kanal bei  
13 TeV mit dem ATLAS-Detektor**

**Studies on the measurement of the  $t\bar{t}Z$  production cross  
section in the 1-lepton channel at 13 TeV with the ATLAS  
detector**



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

vorgelegt von

**Florian Fischer**

geboren am 29. Oktober 1993 in Annweiler am Trifels

München, den 15. September 2017



Gutachter: Prof. Dr. Otmar Biebel



*“An experiment is a question which science poses to Nature  
and a measurement is the recording of Nature’s answer. ”*

Max Planck



## Abstract

In the Standard Model of particle physics the coupling of the top quark to the  $Z$  boson can be calculated with high precision. However, it has not yet been well constrained by direct measurements, thus allowing for various models beyond the Standard Model which would alter the coupling strength. In this analysis, first studies on the measurement of the production cross section of top quark pairs in association with a  $Z$  boson in the so-called 1-lepton channel are performed, exploiting the high branching fractions of the hadronic decay of the  $Z$  boson and of the lepton+jets channel of top quark pair decay, respectively. For this purpose, Monte Carlo simulations are used and have been scaled to the integrated luminosity of  $36.1 \text{ fb}^{-1}$ , which corresponds to the total amount of data suited for physics recorded by the ATLAS experiment at the LHC during the years 2015 and 2016 with a centre-of-mass energy of  $13 \text{ TeV}$ . The  $t\bar{t}Z$  system is kinematically reconstructed to find dedicated variables which discriminate between signal and background. In order to separate the signal from the vast background originating from the associated top quark pair production with jets, a boosted decision tree is made use of in a subsequent step. In addition, the impact of systematic uncertainties is qualitatively estimated. Finally, the sensitivity of this analysis for the whole LHC Run 2 dataset, corresponding to  $100 \text{ fb}^{-1}$  of collected data, is probed.



# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Theory</b>	<b>3</b>
2.1. The Standard model of particle physics . . . . .	3
2.1.1. Gauge invariance and renormalisation . . . . .	5
2.1.2. Quantum chromodynamics . . . . .	8
2.1.3. The electroweak theory . . . . .	10
2.1.4. The weak decay . . . . .	13
2.2. Parton distribution functions . . . . .	13
2.3. The top quark . . . . .	15
2.3.1. Top quark production . . . . .	16
2.3.2. Top quark decay . . . . .	18
2.4. Top quarks in association with a Z boson . . . . .	21
2.4.1. 1-lepton channel . . . . .	22
<b>3. Experimental setup</b>	<b>25</b>
3.1. Experiments at CERN . . . . .	25
3.2. The Large Hadron Collider . . . . .	25
3.3. The ATLAS experiment . . . . .	28
3.3.1. The ATLAS coordinate system . . . . .	29
3.3.2. The ATLAS magnet system . . . . .	30
3.3.3. The ATLAS inner detector . . . . .	31
3.3.4. The ATLAS calorimeter system . . . . .	34
3.3.5. The ATLAS muon spectrometer . . . . .	36
3.3.6. The ATLAS forward detectors . . . . .	38
3.3.7. The ATLAS trigger system and data acquisition . . . . .	38
3.4. Pile-up and underlying event . . . . .	40
<b>4. Data and Monte Carlo Samples</b>	<b>43</b>
4.1. Monte Carlo event generation . . . . .	43
4.2. Monte Carlo event generators . . . . .	46
4.3. Monte Carlo simulation samples . . . . .	47
4.4. Preprocessing and normalisation of simulated events . . . . .	49
4.5. Data recorded by ATLAS . . . . .	49
<b>5. Physics objects</b>	<b>53</b>
5.1. Primary vertex . . . . .	53
5.2. Electrons . . . . .	54
5.3. Muons . . . . .	55
5.4. Jets . . . . .	57
5.5. Flavour tagging . . . . .	59
5.6. Missing transverse energy . . . . .	59
5.7. Overlap removal . . . . .	60

<b>6. Event selection and reconstruction</b>	<b>61</b>
6.1. Event selection . . . . .	61
6.2. Monte Carlo distributions . . . . .	65
6.3. Event reconstruction . . . . .	69
6.3.1. KLFitter . . . . .	71
6.4. Signal region definition . . . . .	73
6.5. Reconstruction results . . . . .	76
6.6. Reconstruction performance . . . . .	81
6.6.1. Investigation of the Z boson mass . . . . .	81
6.6.2. Investigation of the KLFitter likelihood . . . . .	81
6.6.3. Investigation of the KLFitter jet indices . . . . .	82
6.6.4. Summary . . . . .	85
<b>7. Multivariate Analysis</b>	<b>91</b>
7.1. Boosted decision tree . . . . .	91
7.1.1. General concepts . . . . .	91
7.1.2. Boosting . . . . .	93
7.1.3. Data preprocessing . . . . .	93
7.1.4. Performance . . . . .	95
7.2. BDT analysis . . . . .	95
7.2.1. BDT configuration . . . . .	95
7.2.2. Discriminating variables . . . . .	96
7.2.3. BDT training results . . . . .	98
7.2.4. BDT evaluation . . . . .	99
7.2.5. Variable decorrelation . . . . .	99
<b>8. Sensitivity studies towards <math>100 \text{ fb}^{-1}</math></b>	<b>107</b>
8.1. Cut-based analysis . . . . .	107
8.2. BDT analysis . . . . .	107
<b>9. Systematic uncertainties</b>	<b>109</b>
9.1. Experimental uncertainties . . . . .	109
9.2. Theoretical uncertainties . . . . .	109
9.3. Impact of systematic uncertainties on analysis . . . . .	110
<b>10. Conclusion</b>	<b>117</b>
<b>A. Top quark pair production at the LHC</b>	<b>121</b>
<b>B. The KLFitter transfer functions</b>	<b>123</b>
<b>C. Reconstruction of the neutrino momentum</b>	<b>125</b>
<b>D. Boosting algorithms in TMVA</b>	<b>127</b>
D.1. AdaBoost . . . . .	127
D.2. Gradient Boost . . . . .	128
D.3. Bagging . . . . .	129
<b>E. Additional plots</b>	<b>131</b>
E.1. Signal region definition . . . . .	131

E.2. Reconstruction performance . . . . .	132
E.3. Discriminating variables for the BDT . . . . .	138
E.4. BDT training results . . . . .	147
E.5. BDT evaluation . . . . .	149
E.6. Variable correlations . . . . .	152
E.7. Systematic uncertainties . . . . .	155
<b>F. Additional tables</b>	<b>161</b>
F.1. BDT rankings . . . . .	161
<b>References</b>	<b>165</b>
<b>List of Figures</b>	<b>181</b>
<b>List of Tables</b>	<b>185</b>



# 1. Introduction

At the beginning of the 20th century, Ernest Rutherford laid the foundations for the modern atomic model with the discovery of atomic nuclei by the scattering of  $\alpha$ -particles on gold nuclei [1]. Two cornerstones of the following decades are the discovery of the proton in 1919 by Ernest Rutherford [2] and the discovery of the neutron in 1932 by James Chadwick [3, 4], which significantly contributed to the development of nuclear physics. With the help of various experiments the atom has been found to consist of a large negatively charged shell of electrons and of a tiny massive nucleus consisting of neutrons and positively charged protons, which have been assumed to be fundamental. However, in 1956/1957 Robert Hofstadter discovered that the charge of the proton is not concentrated in a single point but that protons exhibit a finite charge distribution [5]. In 1968/1969 Friedman, Kendall and Taylor revealed that the proton itself is not an elementary particle but has an internal structure [6], which led to the foundation of the parton model for the description of the proton content [7, 8].

In order to describe the multitude of particles which have been found at collider experiments in the 1950s and 1960s, Murray Gell-Mann and George Zweig introduced the quark model in 1964 [9–11], which originally consisted of only three particles, namely the up, down and strange quark. To give an explanation for the suppression of so-called flavour changing neutral currents in weak decays, the charm quark has been postulated in 1970 [12, 13]. Even before the discovery of the charm quark in 1974 at SLAC<sup>(1)</sup> [14, 15], Kobayashi and Masukawa predicted the existence of a third quark generation in order to explain the CP violation of the weak interaction [16]. Whereas the bottom quark has been found only four years later at SLAC [17], it took over 20 years until the top quark has been finally discovered by the Tevatron experiments CDF and DØ [18, 19] in 1995. Since then, its characteristics have been extensively studied in great detail, also by the LHC experiments ATLAS and CMS.

Top quark properties of particular interest are its mass and its couplings to electroweak vector bosons. The top quark mass is a free parameter of the Standard Model of particle physics, which has mostly grown in the 1960s and 1970s by the pioneering work covering quantum chromodynamics and the electroweak theory [20–22] as well as the famous Higgs mechanism [23–25], which describes how particles receive their mass due to the electroweak symmetry breaking. The precise determination of the top quark mass is crucial for consistency checks [26] and may play an essential role e.g. in deciding on vacuum stability [27]. The coupling of the top quark to the recently discovered Higgs boson [28, 29] is of special interest when investigating the nature of the Yukawa coupling of fermions to the Higgs field [30] and seeking for an explanation of the mass hierarchy of the Standard Model [31]. Furthermore, the coupling of the top quark to the  $Z$  boson and the photon is not yet well constrained by direct measurements, thus allowing for various models beyond the Standard Model. In order to determine the strength of such a coupling a common approach is to measure the cross section of the corresponding physical process.

With this analysis studies on the measurement of the cross section of the associated production of top quark pairs with a  $Z$  boson, which has been radiated by one of the top quarks, are performed. In the Standard Model this is a quite rare process for which reason only two years ago the ATLAS collaboration could announce evidence with  $4.2\sigma$  [32]. It has been finally discovered in the same year with a significance of  $6.2\sigma$  in the multi-lepton channel by the CMS collaboration [33]. In

---

<sup>(1)</sup> Stanford **L**inear **A**Ccelerator

order to increase the number of signal events with respect to the multi-lepton channel, this analysis focusses on the so-called 1-lepton channel exploiting both the high production cross section of top quark pairs in the lepton+jets channel at the LHC and the high branching ratio of the hadronic decay of the  $Z$  boson. However, in contrast to the multi-lepton channel, the 1-lepton channel suffers from a vast background of top quark pair events with additional jets from gluon radiation, which exceeds the signal by a factor of 1000. In order to tackle this challenge, a dedicated event selection has been derived and the  $t\bar{t}Z$  system has been kinematically reconstructed. Great efforts have been spent on the discrimination of the signal signature from the huge amount of background events arising due to the hadronic environment of the LHC by applying multivariate techniques.

In this analysis, natural units are used meaning that the velocity of light  $c$  and the reduced Planck constant  $\hbar$  are equal to one:  $c = \hbar = 1$ . Therefore energies, momenta and masses are measured in giga-electronvolt (GeV).

In Chap. 2 an overview of the Standard Model and the physics of the top quark as well as top quarks in association with  $Z$  bosons is presented. Afterwards, Chap. 3 describes the experimental setup and introduces the LHC accelerator and the ATLAS experiment. The simulation of high-energy collisions and the reconstruction of physical objects out of the data taken with the ATLAS detector are explained in Chap. 4 and Chap. 5, respectively. The presentation of the analysis starts in Chap. 6 discussing the event selection and the reconstruction of the  $t\bar{t}Z$  system. In Chap. 7 the multivariate analysis is performed, followed by sensitivity studies regarding the whole LHC Run 2 dataset in Chap. 8. The impact of systematic uncertainties on the analysis is briefly discussed in Chap. 9. Finally, in Chap. 10 a summary and outlook of the analysis is given.

## 2. Theory

### 2.1. The Standard model of particle physics

The Standard Model of particle physics (SM), mainly developed in the 1960s and 1970s, explains the world at its most fundamental level. Since then, it has been extremely successful in describing elementary particles and its interactions. Within the field of particle physics it has met every experimental test to highest precision and, additionally, it has predicted unknown characteristics of nature which have been experimentally verified by now. Its maybe most striking confirmation was the discovery of the Higgs boson in 2012 by the LHC experiments ATLAS and CMS [28, 29] which had been theoretically predicted half a century ago. Despite its outstanding success, it suffers of a few limitations, though. The Standard Model is not able to describe the force of gravity, also the unification of all forces covered by the Standard Model has not been achieved yet. Furthermore it gives no insight in the origin of the huge asymmetry between matter and antimatter.

According to the Standard Model, elementary particles are grouped into two categories, particles with half-integer spin, the so-called fermions which are the fundamental building blocks of matter, and particles with integer spin, named bosons, mediating the fundamental forces. The interactions described by the Standard Model are the electromagnetic, the weak and the strong force. The first is mediated by the photon ( $\gamma$ ) and is a source of electricity and magnetism. The weak interaction is accountable for radioactive  $\beta$ -decays and is mediated by  $W$  and  $Z$  bosons. The strong force, mediated by gluons ( $g$ ), is responsible for the binding of protons and neutrons inside of nuclei holding them together. Within the Standard Model, the fermions can be divided into two groups, so-called quarks which interact via gluons, and leptons not participating in the strong force. A characteristic of the Standard Model is its periodic structure meaning that particles can be sorted into generations. For both quarks and leptons a total of three generations exist, each consisting of an up- and down-type quark or of a lepton and the corresponding lepton neutrino. To summarise all particles of the Standard Model, an overview of all bosons is given in Tab. 2.1 and of all fermions in Tab. 2.2, respectively.

Name	Symbol	Mediated force	Colour	Charge [ $e$ ]	Mass [GeV]
Photon	$\gamma$	electromagnetic	no	0	0 (theory)
Gluon	$g$	strong	yes	0	0 (theory)
$W^\pm$ boson	$W^\pm$	weak	no	$\pm 1$	$80.385 \pm 0.015$
$Z$ boson	$Z$	weak	no	0	$91.1876 \pm 0.0021$
Higgs boson	$H$	Higgs field	no	0	$125.09 \pm 0.24$

Tab. 2.1.: The fundamental particles of the Standard Model – section of bosons. In contrast to fermions, the gauge bosons of the Standard Model do not have an antiparticle in general, given one exception as the  $W^\pm$  bosons 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 stated in the table [34].

The Standard Model is formulated in terms of a quantum field theory. In classical mechanics, the dynamics of a system can be described by its Lagrangian  $L(q_i, \dot{q}_i)$  which is a function of a set of generalised coordinates  $q_i$  and their time derivatives  $\dot{q}_i$ . The equations of motion are determined

Generation	Name	Symbol	Colour	Charge [ $e$ ]	Mass [MeV]
1	Up quark	$u$	yes	$+2/3$	$2.2^{+0.6}_{-0.4}$
	Down quark	$d$	yes	$-1/3$	$4.7^{+0.5}_{-0.4}$
	Electron neutrino	$\nu_e$	no	0	$< 0.002$
	Electron	$e$	no	$-1$	0.511
Generation	Name	Symbol	Colour	Charge [ $e$ ]	Mass [MeV]
2	Charm quark	$c$	yes	$+2/3$	$96^{+8}_{-4}$
	Strange quark	$s$	yes	$-1/3$	$1270 \pm 30$
	Muon neutrino	$\nu_\mu$	no	0	$< 0.19$
	Muon	$\mu$	no	$-1$	105.66
Generation	Name	Symbol	Colour	Charge [ $e$ ]	Mass [GeV]
3	Top quark	$t$	yes	$+2/3$	$173.21 \pm 0.87$
	Bottom quark	$b$	yes	$-1/3$	$4.18^{+0.04}_{-0.03}$
	Tau neutrino	$\nu_\tau$	no	0	$< 0.02$
	Tau	$\tau$	no	$-1$	1.77

Tab. 2.2.: The fundamental particles of the Standard Model – section of fermions. For each particle also 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 stated 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 muon neutrino mass is given at 90 % CL [34].

by solving the Euler-Lagrange equations,

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

In quantum field theories, the system no longer comprises discrete particles but continuous fields. Thus, the Lagrangian is replaced by a Lagrangian density  $\mathcal{L}$ ,

$$L = \int \mathcal{L}(\phi_i, \partial_\mu \phi_i) d^3x \quad , \quad (2.2)$$

which is described by fields  $\phi_i$  and their derivatives  $\partial_\mu \phi_i$  with respect to the space-time coordinates  $x^\mu$ . The Euler-Lagrange equation can then be rewritten as

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

The Lagrange density of the Standard Model consists of three parts,

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} \quad , \quad (2.4)$$

where  $\mathcal{L}_{\text{QCD}}$  describes quantum chromodynamics (QCD), the theory of the strong interaction, and  $\mathcal{L}_{\text{EW}}$  represents the electroweak theory, the theoretical foundation of the electromagnetic and

the weak force. The last part,  $\mathcal{L}_{\text{Higgs}}$ , stands for the Higgs mechanism, which is responsible for the electroweak symmetry breaking giving rise to particle masses. (This section is based on Ref. [35–37])

### 2.1.1. Gauge invariance and renormalisation

The two basic principles of quantum field theories are gauge invariance and renormalisability. If a physics process is denoted gauge invariant, its underlying equations and thus itself should not change under a given transformation. This characteristic will be illustrated in the following for quantum electrodynamics (QED) [38–40], the theory of electromagnetic interactions.

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

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

acting on the Dirac spinor fields  $\psi(x)$  with the Einstein sum convention applied. The  $\gamma^\mu$  denote the Dirac matrices, a set of conventional  $4 \times 4$  matrices generating a Clifford Algebra with the following properties:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \quad (2.6a)$$

$$(\gamma^0)^2 = 1 \quad (2.6b)$$

$$(\gamma^{1,2,3})^2 = -1 \quad (2.6c)$$

with the Minkowski metric  $g^{\mu\nu}$ . They are defined in the so-called Dirac representation by the Pauli spin matrices  $\sigma_i$  in 2 dimensions,

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad i \in \{1, 2, 3\} \quad (2.7)$$

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

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0 \quad (2.8)$$

which describes the free motion of a spin-half particle with mass  $m$ , for example an electron. In quantum electrodynamics, gauge transformations are described by  $U(1)_e$  which is part of the unitary group  $U(N)$  describing the set of  $N \times N$  matrices that satisfy the unitary condition  $U^\dagger U = 1$ . The subscript  $e$  denotes the conserved quantity which in case of quantum electrodynamics is the electric charge. A possible representation is

$$U(\alpha) = e^{-i\alpha G} \quad , \quad (2.9)$$

where  $G$  is called the generator of the group. The Dirac equation (2.8) is invariant under a global  $U(1)$  phase transformation

$$\psi(x) \rightarrow \psi'(x) = e^{ie\rho}\psi(x) \quad (2.10)$$

with a constant factor  $\rho$ , meaning that the physical results are unchanged. However, it is not invariant under a local gauge transformation

$$\psi(x) \rightarrow \psi'(x) = e^{ie\rho(x)}\psi(x) \quad (2.11)$$

where  $\rho$  is now a function of the space-time coordinates. Explicitly inserting Eq. (2.11) into the

Dirac equation (2.8) then yields

$$\begin{aligned} (i\gamma^\mu \partial_\mu - m) \psi'(x) &= (i\gamma^\mu \partial_\mu - m) e^{ie\rho(x)} \psi(x) \\ &= e^{ie\rho(x)} [(i\gamma^\mu \partial_\mu - m) \psi(x) - e\partial_\mu \rho(x) \gamma^\mu \psi(x)] \end{aligned} \quad (2.12)$$

which contains the factor  $e\partial_\mu \rho(x) \gamma^\mu \psi(x) \neq 0$  thus spoiling the invariance of Eq. (2.8) under a local gauge transformation. In order to guarantee the gauge invariance of the Dirac equation, the covariant derivative  $D_\mu$  is introduced to replace the conventional one,

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ieA_\mu(x) \quad , \quad (2.13)$$

naturally introducing the principle of an interaction. The newly introduced vector gauge field  $A_\mu(x)$  corresponds to the photon, which is the mediator of the electromagnetic interaction, and must simultaneously transform like

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \rho(x) \quad (2.14)$$

in order to keep the local gauge invariance. With these transformations the Dirac equation becomes again invariant under local gauge transformations and preserves its form,

$$(i\gamma^\mu D_\mu - m) \psi(x) = 0 \quad . \quad (2.15)$$

If this is inserted in Eq. (2.5), the Lagrangian of quantum electrodynamics is obtained,

$$\mathcal{L}_{\text{QED}} = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi + e\bar{\psi} A_\mu \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad . \quad (2.16)$$

The first term in Eq. (2.16) describes the free motion of a free spin-half particle with mass  $m$ , the second term represents the interaction of such a particle with a photon and the last term is the kinetic term of the photon field which has to be additionally added to the Lagrangian. The field strength tensor  $F_{\mu\nu}$  is constructed from the vector potential  $A_\mu$ ,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \quad (2.17)$$

Any additional term like a photon mass term which would look like  $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$  again spoils the gauge invariance of Eq. (2.16), thus the condition in QED to be gauge invariant under local  $U(1)$  phase transformations forces the photon to be massless. This is in good agreement with the experimental results which set an upper limit of  $1 \cdot 10^{-18}$  eV on the photon mass. The demand of massless force carriers is a crucial prerequisite for (local) gauge invariance, which ultimately leads to the Englert-Brout-Higgs mechanism [23–25] giving rise to particle masses and the existence of a Higgs boson.

Another important subject within the Standard Model is renormalisability, for which gauge invariance is an essential precondition [41]. Renormalisability means that divergent terms occurring in theoretical calculations of physical observables can be absorbed in finite measurable quantities. Many theoretical calculations of physical observables, as cross sections of particular physics processes or decay rates, are done in terms of perturbation theory since they are not analytically solvable anymore. This can be illustrated with the electron-positron scattering, a process from quantum electrodynamics again. The calculation of the leading order process, shown in Fig. 2.1a, works fine but yields only an approximation. In order to get the exact result, all higher order corrections as vacuum polarisation presented in Fig. 2.1b have to be considered. When calculating the cross sections for this process a divergent term is inevitably picked up which in essence looks

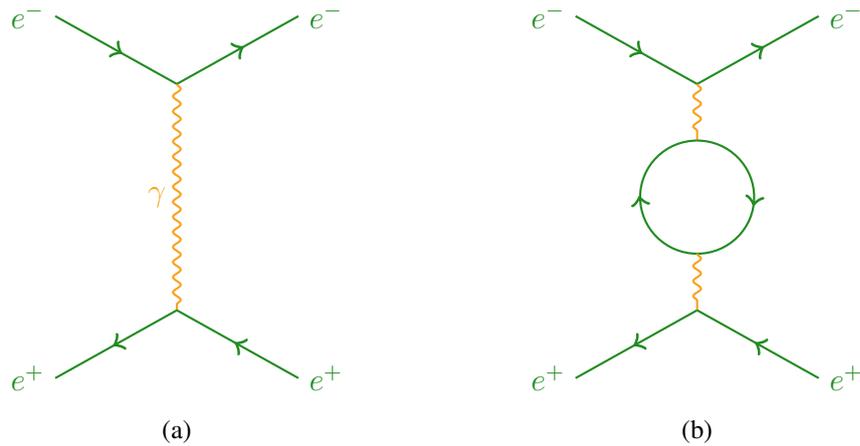


Fig. 2.1.: Electron-positron scattering via the exchange of a virtual photon. The leading order process is shown in Fig. (a), whereas a higher order correction in which a virtual photon splits up into an electron-positron pair that annihilates again, is depicted in Fig. (b).

like

$$\int_{m^2}^{\infty} \frac{1}{z} dz \quad , \quad (2.18)$$

with  $z$  denoting the momentum of the virtual particle. With this, the integral is logarithmically divergent for large  $z$  and the cross section ultimately becomes infinite which obviously contradicts experimental results. To overcome this disaster, the integral is regularised, meaning that a suitable upper cut-off mass  $M$  is introduced which in a later step is sent to infinity. Then, the integral reads

$$\int_{m^2}^{M^2} \frac{1}{z} dz = \ln \left( \frac{M^2}{m^2} \right) \quad (2.19)$$

and can be split into a calculable finite term which is independent of  $M$ , and a term which blows up as  $M \rightarrow \infty$ . Most important, all divergent terms are later absorbed in physical quantities as they appear in the final result as additions to the calculated physical observables, in this example the electron charge. This means that the measured masses  $m'$  and couplings  $g'$  are not the same as those appearing in the Lagrangians ( $m, g$ ) but rather renormalised ones containing extra corrections  $\delta m$  and  $\delta g$ ,

$$m' = m + \delta m \quad (2.20a)$$

$$g' = g + \delta g \quad . \quad (2.20b)$$

As these corrections depend on the masses and coupling of the particles involved in a physics process, the final masses and couplings become energy-dependent, also referred to as running masses or running couplings. For the here given examples, it thus has to be distinguished between a bare electric charge  $e_0$  used in theory calculations and a renormalised charge  $e$  measured in experiment. The relation between them is given by

$$e = e_0 \left( 1 - \frac{e_0^2}{12\pi^2} \ln \left( \frac{M^2}{m^2} \right) \right) \quad (2.21)$$

absorbing all infinite terms. Since the charge measured in experiment is obviously finite, the divergence in  $\left( \frac{M^2}{m^2} \right)$  must be cancelled by a corresponding divergence in the bare charge  $e_0$ . As a

direct consequence, the coupling strength is no longer constant but depends on the energy of the particular physics process. Hence, in leading order the QED coupling constant can be written as

$$\alpha(Q^2) = \frac{e^2(Q^2)}{4\pi} = \frac{\alpha(\mu^2)}{1 - \frac{\alpha(\mu^2)}{3\pi} \ln\left(\frac{Q^2}{\mu^2}\right)} \quad (2.22)$$

where  $Q^2$  denotes the momentum scale of the process and  $\mu^2$  is a reference renormalisation scale conceptionally related to the cut-off value  $M^2$ . The experimental results shown in Fig. 2.2a indeed prove that the electromagnetic coupling constant increases with larger momentum transfer, as implied by Eq. (2.22).

(For this section Ref. [36, 37, 42, 43] have been used)

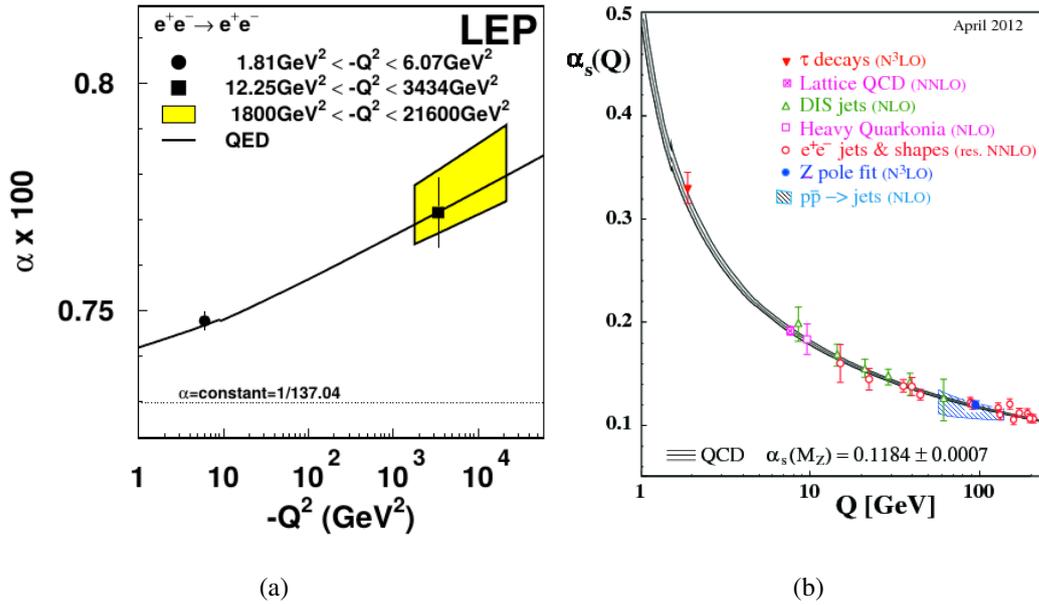


Fig. 2.2.: Running of the coupling constants for the electromagnetic and strong force obtained from experiment [44, 45]. Fig. (a) shows the energy dependence of  $\alpha$ , while Fig. (b) depicts the result for  $\alpha_s$ .

### 2.1.2. Quantum chromodynamics

Quantum chromodynamics is the theory of the strong force describing the interaction between particles which carry a colour charge, namely quarks and gluons. There are in total three different colour charges, chosen as red, green and blue, which can be considered as axes forming an abstract internal space. As no axis in colour space is distinguished, transformations between colours, corresponding to rotations in colour space, thus represent a symmetry which is described by the  $SU(3)_C$  symmetry group. The special unitary group  $SU(N)$  is a subgroup of  $U(N)$  and in addition has to provide  $\det U = 1$ . In quantum chromodynamics the colour charge is the conserved quantity, it is therefore described by the representation of  $SU(N)$  in three dimensions,  $SU(3)_C$ .  $SU(N)$  groups have in total  $N^2 - 1$  generators, which for  $SU(3)_C$  form 8 gluons and are expressed in terms of the Gell-Mann matrices  $\lambda_\alpha$ . The gluons are the force mediators of quantum chromodynamics and are represented by eight bicoloured gauge fields  $G_\mu^\alpha$ .

The Lagrangian of quantum chromodynamics can thus be written as

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} \mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu} + \sum_k \bar{q}_k (i\gamma^\mu D_\mu - m_k) q_k \quad (2.23)$$

where  $\mathbf{G}_{\mu\nu}$  represents the gluon field strength tensor and  $q_k$  denotes the different quark flavours. The covariant derivative is given by

$$D_\mu = \partial_\mu + ig_s \frac{\lambda}{2} \mathbf{G}_\mu \quad (2.24)$$

in which  $g_s$  is the strong coupling and  $\lambda$  and  $\mathbf{G}_\mu$  are vectors of the Gell-Mann matrices and gluon gauge fields, respectively. The field strength tensor is defined as

$$\mathbf{G}_{\mu\nu} = \partial_\mu \mathbf{G}_\nu - \partial_\nu \mathbf{G}_\mu - g_s \mathbf{G}_\mu \times \mathbf{G}_\nu \quad (2.25)$$

Except for the additional cross product it looks similar to the field strength tensor  $F_{\mu\nu}$  in QED (cf. Eq. 2.17). This last term is a characteristic feature of quantum field theories which are based on non-abelian groups, employing that the gauge bosons carry (colour) charge themselves and thus interact with each other. Fig. 2.3 shows two exemplary Feynman diagrams of gluon-gluon self-interactions.

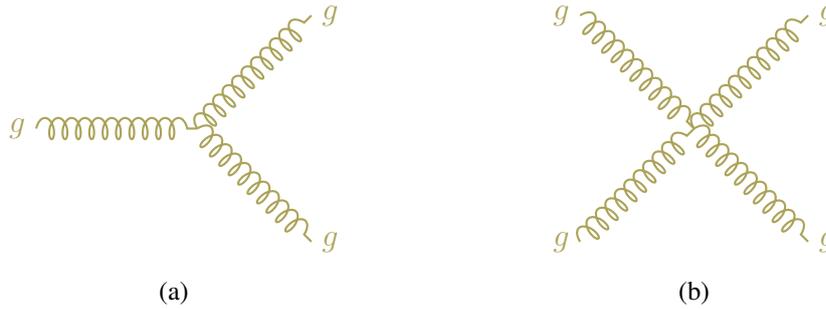


Fig. 2.3.: Gluon self-coupling vertices arising from the non-abelian structure of quantum chromodynamics. Fig. (a) shows a three-gluon vertex, whereas in Fig. (b) a four-gluon vertex is depicted.

As for QED, the QCD coupling constant is not actually constant but varies with changes of the momentum transfer scale, and in leading order is given by

$$\alpha_s = \frac{g_s^2}{4\pi} = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi} (33 - 2n_f) \ln\left(\frac{Q^2}{\mu^2}\right)} \quad (2.26)$$

where  $n_f$  is the number of quark flavours able to participate in a particular interaction. Compared to the electromagnetic coupling constant  $\alpha$  (cf. Eq. 2.22), the overall plus sign in front of the logarithm derives from the gluon self-interactions. Therefore, opposite to  $\alpha$ ,  $\alpha_s$  decreases with increasing momentum transfer and vice versa, which is shown in Fig. 2.2b, giving rise to the phenomena of asymptotic freedom [46, 47] and confinement [48, 49].

The first, asymptotic freedom describes the behaviour of quarks at very small distances, i.e. for high energies. In this regime quarks behave as quasi-free particles since there the coupling constant  $\alpha_s$  becomes increasingly small. On the contrary, the strong coupling constant increases for large distances, corresponding to small energies. Therefore, quarks can not exist as free particles any-

more but are subjected to colour confinement which forces them to form bound colourless states, the so-called hadrons. Usually, hadrons are divided into two groups, mesons and baryons. Since colour is an additive quantum number, both mesons and baryons are uncoloured objects. Mesons are combinations of a quark and an antiquark with its anticolour compensating the colour of the quark. Baryons consist either of three quarks or three antiquarks with each of the (anti)quarks carrying a different (anti)colour. However, recent measurements revealed the existence of a more exotic state consistent with a so-called pentaquark ( $uudc\bar{c}$ ), which is a bound state of five quarks [50]. Because of the low energy scales involved, the process of the formation of hadrons, referred to as hadronisation, is theoretically very challenging. Due to confinement, no single quarks but only jets, cone-shaped bundles of colourless hadrons, can be observed in experiment.

(This section is based on Ref. [36, 37, 42, 43, 51])

### 2.1.3. The electroweak theory

In the Standard Model, both the electromagnetic and the weak interaction, though appearing separately, are a low-energy representation of a single interaction described by the electroweak theory (EW). It is based on the  $SU(2)_L \otimes U(1)_Y$  gauge group and has a total of four different gauge fields. Three of them ( $W_\mu^1, W_\mu^2, W_\mu^3$ ) belong to the  $SU(2)_L$  group and couple to the weak isospin, whereas the fourth field ( $B_\mu$ ) is associated with the  $U(1)_Y$  group and couples to the weak hypercharge, defined as

$$Y = 2(Q - I_3) \quad (2.27)$$

where  $Q$  is the electric charge and  $I_3$  represents the third component of the weak isospin, respectively. As the weak force does not conserve parity, only left-handed fermions and right-handed antifermions carry a weak isospin unequal to zero forming so-called weak isospin doublets

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L \quad (2.28)$$

where the third component of the weak isospin is  $+1/2$  for the upper particle and  $-1/2$  for the lower particle. In contrast, right-handed particles and left-handed antiparticles are weak isospin singlets with  $I_3 = 0$ . The gauge bosons of  $SU(2)_L$  couple only to left-handed particles and to right-handed antiparticles, referred to as parity violation.

The Lagrangian of the electroweak interaction is given by

$$\mathcal{L}_{\text{EW}} = \sum_k i\bar{\psi}_k \gamma^\mu D_\mu \psi_k - \frac{1}{4} \mathbf{W}_{\mu\nu} \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (2.29)$$

with  $\psi_k$  denoting the different quark and lepton flavours. As the coupling to left-handed and right-handed particles is different, the covariant derivative reads

$$D_\mu = \partial_\mu + \frac{1}{2} i g \boldsymbol{\tau} \cdot \mathbf{W}_\mu + \frac{1}{2} i g' Y B_\mu \quad (2.30a)$$

$$D_\mu = \partial_\mu + \frac{1}{2} i g' Y B_\mu \quad (2.30b)$$

where Eq. (2.30a) represents the left-handed part and Eq. (2.30b) the right-handed part. Furthermore,  $g$  and  $g'$  denote the couplings of  $SU(2)_L$  and  $U(1)_Y$ , while  $\boldsymbol{\tau}$  and  $\mathbf{W}_\mu$  are vectors comprising the Pauli spin matrices and the  $SU(2)_L$  gauge fields, respectively. The field strength

tensors of the electroweak theory are given by

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.31a)$$

$$\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu - g \mathbf{W}_\mu \times \mathbf{W}_\nu \quad (2.31b)$$

In contrast to  $U(1)_Y$ , the  $SU(2)_L$  group is a non-abelian group leading to a self-interaction of its gauge boson similar to QCD (cd. Sec. 2.1.2).

In order to preserve the gauge invariance of the electroweak Lagrangian (Eq. (2.29)), the gauge bosons of the electroweak interaction have to be massless as stated in Sec. 2.1.1. However, experiments have shown that only the photon is massless, whereas the  $W^\pm$  and  $Z$  bosons are quite massive (cf. Tab 2.1). Consequently, to meet both the experimental results and the principle of gauge invariance, the electroweak symmetry has to be broken. Within the Standard Model, the electroweak symmetry breaking is described by the Brout-Englert-Higgs mechanism [23–25] by adding a new complex scalar  $SU(2)_L$  doublet field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.32)$$

to the electroweak Lagrangian. The corresponding additional term in the Lagrangian then reads

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad (2.33)$$

where the covariant derivative is simply the one used in electroweak theory and the potential is given by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (2.34)$$

with  $\lambda$  being a positive real number. The minima of this potential depend on the choice of the parameter  $\mu^2$ . If  $\mu^2$  is set to value a greater than zero, the only minimum is at  $\Phi = 0$ . However, for the case that  $\mu^2 < 0$ ,  $\Phi$  develops a ground state different from zero,

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.35)$$

for which the potential is minimised. The new minima of the potential are  $\pm v$  where

$$v = \sqrt{\frac{-\mu^2}{\lambda}} \quad (2.36)$$

is called vacuum expectation value (VEV) of the Higgs field. By choosing one of the minima the original symmetry of the potential is broken. This procedure is also referred to as spontaneous symmetry breaking meaning that the underlying symmetry is spoiled by the selection of a particular ground state. Within the freedom of gauge transformation  $\Phi$  can be formulated as an expansion around the vacuum expectation value with a scalar field  $\eta(x)$ ,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix} \quad (2.37)$$

which is later on identified with the Higgs field. Using this representation of  $\Phi$ , the Lagrangian in Eq. (2.33) can be reformulated resulting in the physical mass eigenstates of the gauge bosons,

$$\mathcal{L}_{\text{Higgs}} = \left[ \frac{1}{2} (\partial_\mu \eta) (\partial^\mu \eta) - \mu^2 \eta^2 \right] + \frac{1}{2} \frac{g^2 v^2}{4} [ |W_\mu^+|^2 + |W_\mu^-|^2 ] + \frac{1}{2} \frac{v^2}{4} |g' B_\mu - g W_\mu^3|^2 \quad (2.38)$$

## 2. Theory

This Lagrangian is given at leading order including the free propagation of the Higgs field  $\eta(x)$  with mass  $m_{\text{Higgs}} = \sqrt{2}\mu$  and the mass terms for the electroweak gauge bosons which are generated by spontaneous symmetry breaking, induced by the reformulation of  $\Phi$  as an expansion around the vacuum expectation value in combination with the requirement of local gauge invariance under  $SU(2)_L \otimes U(1)_Y$  symmetry. Higher orders not shown here include cubic and quartic gauge boson couplings to the Higgs field and self-interaction of the Higgs field. As the free propagation of the gauge bosons is already depicted in Eq. (2.29) it is not presented here. Using the electroweak mixing angle, referred to as Weinberg angle, which is defined through the couplings of the  $SU(2)_L$  and  $U(1)_Y$  groups,

$$\theta_W := \tan^{-1} \left( \frac{g'}{g} \right) \Leftrightarrow \cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (2.39)$$

and relates all couplings involved in the electroweak interaction by

$$e = g' \cos \theta_W = g \sin \theta_W \quad , \quad (2.40)$$

the electroweak gauge bosons and their acquired masses can be expressed in terms of the electroweak gauge fields as

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \quad \text{with} \quad m_\gamma = 0 \quad (2.41a)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \quad \text{with} \quad m_Z = \frac{m_{W^\pm}}{\cos \theta_W} \quad (2.41b)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \quad \text{with} \quad m_{W^\pm} = \frac{gv}{2} \quad (2.41c)$$

This means that two gauge fields of  $SU(2)_L$ ,  $W_\mu^1$  and  $W_\mu^2$ , are absorbed in the  $W^\pm$  bosons, whereas the photon and the  $Z$  boson are superpositions of  $W_\mu^3$  and  $B_\mu$ , the gauge field of the  $U(1)_Y$  gauge group. The Weinberg angle itself has been measured to be  $28.75^\circ$  which corresponds to  $\sin^2 \theta_W = 0.2313$ .

With the introduction of the Higgs mechanism not only the gauge bosons, but also quarks and leptons gain their mass by expanding the Higgs Lagrangian with an additional mass term for fermions of the form

$$\tilde{g}_f \left( \bar{\Psi}_L^f \Phi \Psi_R^f + \bar{\Psi}_R^f \Phi^\dagger \Psi_L^f \right) \quad (2.42)$$

where  $\Psi$  denote the left-handed and right handed Dirac spinors of a fermion  $f$  and  $\tilde{g}_f$  its corresponding Yukawa coupling. The mass of a fermion is thus given by

$$m_f = \tilde{g}_f \frac{v}{\sqrt{2}} \quad (2.43)$$

where the respective Yukawa coupling of a particular fermion is a free parameter in the Standard Model. The vacuum expectation value of the Higgs field is fixed by the Fermi constant  $G_F$ ,  $v = (\sqrt{2}G_F)^{-1/2}$  and has been determined to amount to 246 GeV by muon decay measurements [52–54]. Therefore, due to its particular mass the top quark has a Yukawa coupling of almost one, while all other fermions have considerably smaller couplings to the Higgs field.

(This section is adapted from Ref. [36, 42, 43, 55, 56])

### 2.1.4. The weak decay

Since the electroweak symmetry is broken, the mass eigenstates of massive strongly interacting elementary particles, i.e. quarks, and their respective weak eigenstates are not identical. Therefore, they are related to each other by the unitary  $3 \times 3$  CKM flavour mixing matrix which is named after Nicola Cabibbo, Makoto Kobayashi and Toshihide Masukawa for their outstanding pioneering work in this field [16, 57], describing the mixing between mass and weak eigenstates. It is an extension of the GIM mechanism [12], which only includes the first two quark generations and relies on the  $2 \times 2$  Cabibbo flavour mixing matrix. The mass eigenstates are a superposition of different weak eigenstates, hence quarks are allowed to change their flavour in decays mediated by the weak force via the exchange of a  $W^\pm$  boson. The relation between the weak eigenstates, denoted by a prime, and the mass eigenstates given by the CKM matrix is

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \quad (2.44)$$

The particular entries of the CKM matrix define the transition of a quark of flavour  $i$  to flavour  $j$ , where the transition probability for such a flavour change is given as  $|V_{ij}|^2$ . The actual values of the CKM matrix elements are measured to be [34]

$$V_{\text{CKM}} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & 0.00409 \pm 0.00039 \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & 0.00405 \pm 0.0015 \\ 0.0082 \pm 0.0006 & 0.04 \pm 0.0027 & 1.009 \pm 0.031 \end{pmatrix} \quad (2.45)$$

from which it can be immediately seen that transitions within the same generations are most preferred as the corresponding diagonal values are very close to unity. Transitions between the first and second generation are less preferred, followed by transitions between the second and the third generation. The the most suppressed flavour transitions are those between the first and the third generation. As an example, for the top quark only the matrix elements  $|V_{td}|^2 = 6.7 \cdot 10^{-5}$ ,  $|V_{ts}|^2 = 1.6 \cdot 10^{-3}$  and  $|V_{tb}|^2 = 1.02$  are of importance, reflecting the relative probabilities that a top quark decays into a down, strange and bottom quark, respectively.

From the structure of the matrix it can be deduced that only transitions between up-type quarks and down-type quarks are permitted. In contrast, transitions within the up-type or down-type quarks are called flavour changing neutral currents (FCNC). In the Standard Model such transitions do not exist at tree level and are moreover highly suppressed at higher orders. It was actually the suppression of FCNC that led to the prediction of the charm quark by the GIM mechanism in 1970 [13]. However, in order to explain the CP violation observed in weak interactions two quark generations had not been sufficient enough, thus a third quark generation had been postulated and experimentally established with the discovery of the bottom quark. In the Standard Model, CP violation is described by a complex phase in the CKM matrix which can mediate an asymmetry between matter and antimatter. In order to have such a complex phase, a unitary matrix needs to have at least three dimensions which in case of the CKM matrix are represented by the three quark generations.

(Furthermore, this section uses Ref. [37])

## 2.2. Parton distribution functions

In the parton model the proton content can be split into three different categories. The main constituents are three valence quarks, more precisely two up quarks and one down quark, which are

embedded in a sea composed of many temporarily created quark-antiquark pairs originating from quantum fluctuations, so-called sea quarks, and of a multiplicity of gluons. Originally introduced by Richard Feynman for the description of high-energetic collisions with hadrons involved [7, 8], even after the experimental confirmation of quantum chromodynamics the parton model remains a justifiable approximation at high energies, where both models treat partons behaving as free particles. The latter is a priori assumed in the parton model, while in QCD it is an effect referred to as asymptotic freedom arising from the anti-screening behaviour of the strong coupling constant which diminishes for decreasing distances between two coloured objects, thus allowing for perturbative calculations of collision cross sections. However, at large distances the interaction becomes stronger leading to the confinement of quarks and gluons within composite hadrons. In this regime, the strong coupling constant is typically larger than one, meaning that effects of QCD cannot be calculated perturbatively anymore at this energy scale. Instead, probability density functions have to be used to describe the momentum distribution of hadron constituents. These so-called parton distribution functions give the probability of finding a certain parton with a specific fraction  $x$  of the entire hadron's momentum at the resolution scale  $Q$ . Because of the inherent non-perturbative nature of partons which cannot be observed as free particles due to the colour confinement, the parton densities cannot be calculated analytically for a given momentum transfer scale. Thus, they have to be extracted in global fits to QCD observables from deep-inelastic scattering and hadron collider data. However, their evolution in momentum transfer from a high towards a lower scale, down to a cut-off scale of few hundreds MeV, can be described by the DGLAP evolution equations [58–60].

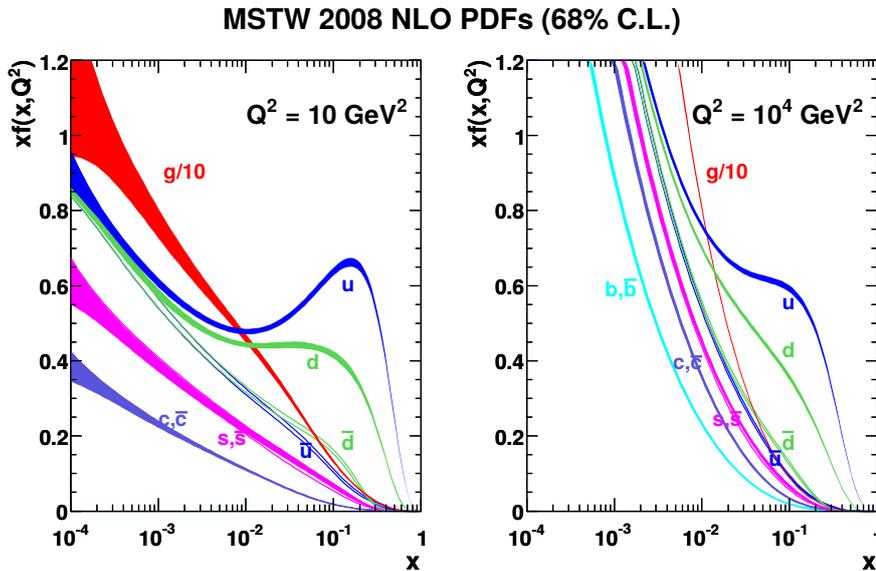


Fig. 2.4.: Parton distribution function for the proton provided by the MSTW group [61] in 2008 at next-to-leading order for two different energy scales,  $Q^2 = 10 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$ . The valence quark carry high momentum fractions of roughly  $1/3$  where the up quark component exceeds the down quark component by a factor of approximately two. For smaller  $x$ , gluons and other quark flavours become more likely. When going towards higher energies and therefore on average higher momentum transfer, more and 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 picture frame.

At hadron colliders as the LHC the incoming protons provide broad band beams of partons which carry a fraction  $x$  of the original proton momentum. Hence, parton distribution functions are used at collider experiments to describe the substructure of protons and to quantify the outcome of high-energetic proton-proton collisions. As can be seen in Fig. 2.4 the valence quarks on average carry high momentum fractions while for decreasing  $x$ , gluons and other quark flavours become increasingly likely. For a correct calculation of the total cross section of a particular physical process, an inclusive collection of all final state hadrons has to be considered. In order to overcome the difficulties arising due to the low energies involved, the factorisation theorem can be made use of, stating that the total inclusive cross section of a hadronic process  $A + B \rightarrow X$  can be written as

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_s(\mu_R^2) \hat{\sigma}_1 + \dots]_{ab \rightarrow X} \quad (2.46)$$

In this way, the calculation of the cross section is factorised into the cross sections of the leading order partonic process  $\hat{\sigma}_0$  with its higher order corrections, and the appropriate parton distribution functions  $f_{a/A}$  and  $f_{b/B}$  of the initial state parton  $a$  and  $b$  belonging to the incoming hadrons  $A$  and  $B$ , respectively. After choosing a suitable factorisation scale ( $\mu_F^2$ ) and renormalisation scale ( $\mu_R^2$ ) it is numerically integrated over the momentum fractions  $x_i$  of the involved partons and other phase space variables associated with the final state  $X$ . At the LHC the partons  $a$  and  $b$  correspond to partons originating from protons  $A$  and  $B$  out of the colliding beams. The final state could be any resonance produced through the collision of the partons  $a$  and  $b$ , for example a  $Z$  boson produced by the annihilation of a quark-antiquark pair,  $q\bar{q} \rightarrow Z$ .

### 2.3. The top quark

After a first direct experimental evidence, the top quark has been ultimately discovered in 1995 at the Tevatron experiments CDF and DØ [18, 19]. From then on a multitude of measurements have been performed to determine the top quark properties. One of the most extensively studied properties is its high mass. With  $m_{\text{top}} = 173.34 \pm 0.36 \text{ stat.} \pm 0.67 \text{ syst. GeV}$  [62] the top quark is by far the heaviest fundamental particle known<sup>(1)</sup>. Hence, it plays an important role in many fields of particle physics. The top quark mass itself is a free parameter within the Standard Model, thus it has to be determined by experiment. Since the lifetime of a particle is related to its mass as depicted by Eq. (2.47) [36], the top quark has a very short lifetime of  $\sim 5 \cdot 10^{-25} \text{ s}$ .

$$\tau_{\text{top}} = \frac{1}{\Gamma_{\text{top}}} \sim \frac{1}{m_{\text{top}}^3} \quad (2.47)$$

Therefore, the top quark is the only quark that does not form hadronic bound states since it decays before hadronisation sets in after roughly  $10^{-24} \text{ s}$ . Consequently, the bare properties of the top quark can be studied. Furthermore, the top quark likely plays a special role in electroweak symmetry breaking because its Yukawa coupling is very close to unity. Additionally, the top quark mass is, together with the Higgs boson mass, related to the question whether the electroweak vacuum is either stable or meta-stable or not stable at all. Moreover, top quark processes are major background processes to different searches of new physics, primarily Supersymmetry [31] searches. Besides the intensive work on the measurement of the top quark mass, a lot of effort is still spent on precision measurements of the top quark differential cross sections, its couplings to vector bosons and spin correlations, as these might provide deeper insights into currently not

<sup>(1)</sup> Actually, it has a mass in the order of the magnitude of a gold atom.

satisfactorily explained phenomena like flavour changing neutral currents or the coupling of the Higgs boson to fermions and therefore into the mechanism how particles receive their mass.

### 2.3.1. Top quark production

There are two mechanisms for the production of top quarks, either the single top quark production in which a top quark is produced in an electroweak process, or the top quark pair production in which a pair of top quarks is created in a strong process.

#### 2.3.1.1. Single top quark production

The electroweak production of a single top quark can be categorised into three different channels. The  $t$ -channel process describes the production of a top quark by a flavour excitation process. A virtual  $W$  boson is radiated by a quark which couples to a  $b$ -quark originating either from the quark sea of the proton or from gluon splitting, thereby producing a top quark. In the  $s$ -channel a virtual  $W$  boson is created by the annihilation of an up-type quark with a down-type quark. Then, the  $W$  boson splits up into an (anti)top quark and the corresponding down-type (anti) quark. The third production mechanism for single top quark production is the associated production of a  $W$  boson and a top quark, referred to as  $Wt$ -channel. In this channel either a down-type quark is excited by gluon and thus has sufficient energy to become a top quark by the radiation of a  $W$  boson, or a down-type quark splits up into a  $W$  boson and a virtual top quark which subsequently couples to a gluon, thereby becoming real. The corresponding Feynman diagrams are shown in Fig. 2.5.

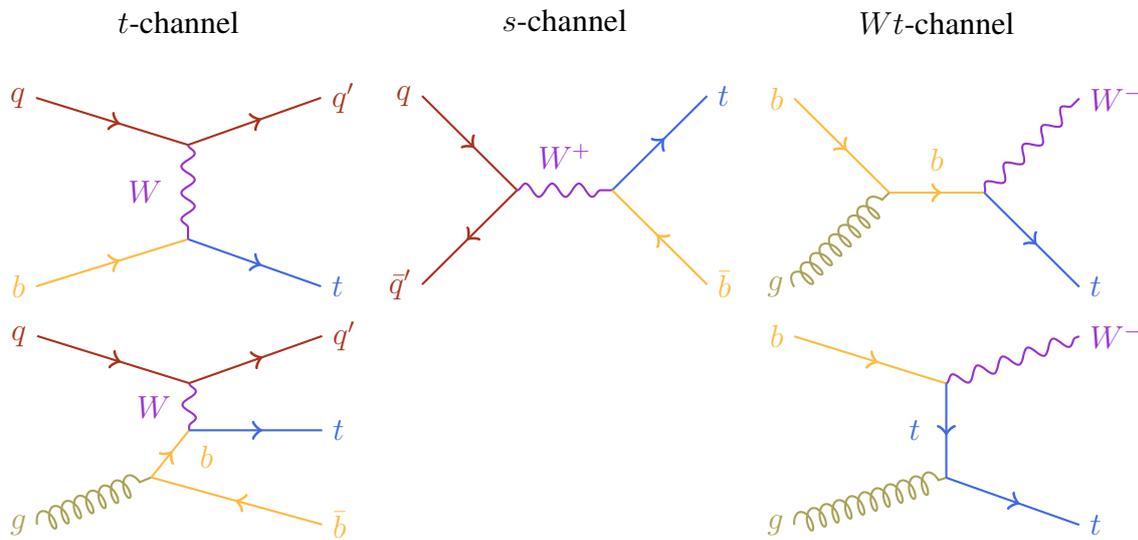


Fig. 2.5.: Exemplary production mechanisms of single top quarks via the electroweak interaction.

For single top quark production the cross sections for producing a top quark or an antitop quark is different in the  $t$ - and  $s$ -channel. This is caused by the structure of the proton itself which is composed of one down and two up valence quarks. Due to charge conservation, an up-type quark can only radiate an  $W^+$  boson, whereas a down-type quark has to emit a  $W^-$  boson. The emitted  $W$  boson annihilates either with a down-type quark producing a top quark or with down-type antiquark resulting in an antitop quark. As the number of up quarks in the proton is twice as large as the number of down-type quarks and the up-type and down-type sea quarks occur at the same

rate, the production cross section for top quark is larger than the production cross section of anti top quarks in the  $t$ -channel. The same holds for the  $s$ -channel, in which the virtual  $W$  boson will hence rather be positively than negatively charged. Consequently, the cross section for top quark production at the LHC is enhanced, whereas at proton-antiproton colliders as the Tevatron no difference in the cross sections should be recognised.

### 2.3.1.2. Top quark pair production

In contrast to the single top quark production, top quark pairs are dominantly created via the strong interaction. In Fig. 2.6 the leading order Feynman diagrams are shown, comprising the three possible categories top quark pair production is classified in, the  $s$ -,  $t$ - and  $u$ -channel. Similar to the electroweak production, the  $s$ -channel describes the annihilation of a quark and an antiquark producing a gluon which subsequently splits up into a top and an antitop quark. As discussed in Sec. 2.1.2, gluons themselves carry colour and therefore couple to each other, allowing for a complementary process in which two gluons annihilate under the emission of a virtual gluon that again splits up into top quark pair. However, at the LHC the contribution of quark-antiquark annihilation is approximately 10 %, whereas gluon-gluon fusion distinctly dominates with a share of around 90 %. A motivation for this discrepancy is given in Appendix A. In the  $t$ -channel, two gluons exchange a virtual top quark which in turn leads to the production of a top quark pair. The  $u$ -channel has exactly the same initial and final state particles as the  $t$ -channel but with an flipped assignment of the top and antitop quark to the initial state gluons.

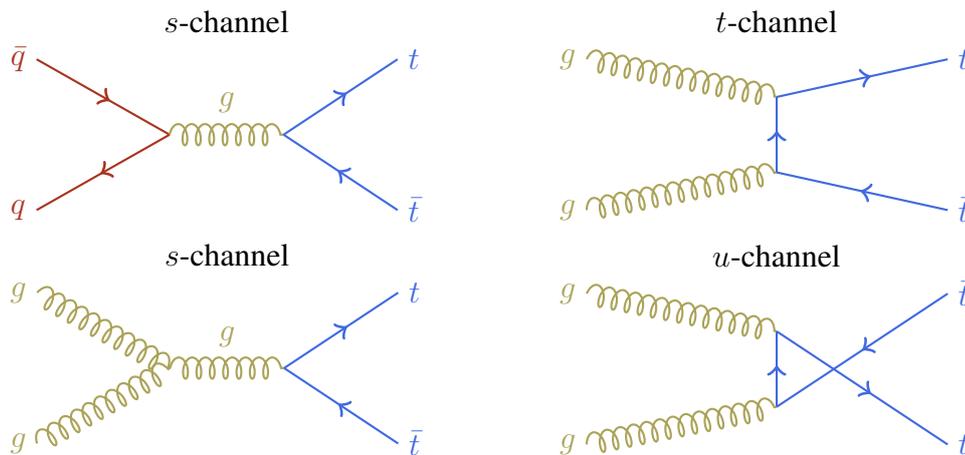


Fig. 2.6.: Exemplary Feynman diagrams for leading order top quark pair production mechanisms.

Besides the leading order processes, the top quark pair production can also involve higher order processes, as exemplary shown for next-to-leading order (NLO) in Fig. 2.7. Within strong processes, very likely an additional gluon is radiated either in the initial state or the final state, referred to as initial (ISR) and final state radiation (FSR), respectively. Another NLO correction are quantum loops describing the exchange of a purely virtual particle. Here shown are additional gluons which are radiated and absorbed by either the gluons or the top quarks of the leading order process.

In total, the theoretical prediction for the top quark pair production cross section at a proton-proton collider running at a centre-of-mass energy of 13 TeV is  $\sigma_{t\bar{t}}^{\text{NNLO+NNLL}} = 832_{-29}^{+20} (\text{scale}) \pm 35 (\text{PDF} + \alpha_s) \text{ pb}$  [63], calculated at next-to-next-to-leading order (NNLO) in perturbative QCD including soft-gluon resummation at next-to-next-to-logarithmic order (NNLL), assuming a top

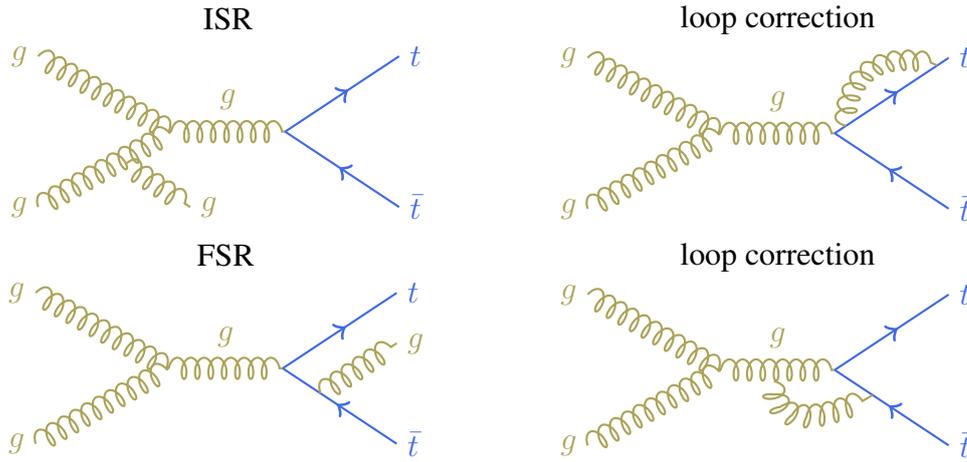


Fig. 2.7.: Exemplary Feynman diagrams for the next-to-leading order top quark pair production.

quark mass of  $m_{\text{top}} = 172.5 \text{ GeV}$ . The first uncertainty reflects uncertainties in the factorisation and renormalisation scales,  $\mu_F$  and  $\mu_R$ , while the second one is associated with possible choices in parton distributions (PDF) and the strong coupling ( $\alpha_s$ ). The various measurements on the top quark pair production measurement are in good agreement with the theoretical prediction as depicted in Fig. 2.8.

### 2.3.2. Top quark decay

Due to its high mass, the top quark has a very short lifetime as discussed in Sec. 2.3. It decays via the electroweak interaction into a  $W$  boson and a down-type quark, which is primarily a bottom quark (cf. Sec. 2.1.4). A  $W$  boson can either decay into a charged lepton and the corresponding neutrino, referred to as leptonic decay, or into an up-type quark and a down-type quark, which is named the hadronic decay. The corresponding branching fractions of the  $W$  bosons, which are  $\sim 1/3$  for the leptonic and  $\sim 2/3$  for the hadronic decay, are given in Tab. 2.3. In fact, the decay into a top and a bottom quark would be technically allowed, it is kinematically prohibited by the high rest mass of the top quark which exceeds the rest mass of a  $W$  boson more than twice. Therefore, hadronic decays of  $W$  bosons involve only quarks lighter than the top quark. The decay of top quark pairs can thus be divided into three different decay channels which are characterised by the decay mode of the involved  $W$  bosons. In the dileptonic channel both  $W$  bosons decay leptonically giving rise to two neutrinos in the final state, whereas in the fully hadronic channel both  $W$  bosons decay into quarks. The third channel is the so-called lepton+jets channel in which one  $W$  boson decays leptonically and the other one hadronically. However, the numbers given in Tab. 2.3 include the contributions from all three lepton flavours. Since  $\tau$ -leptons decay to roughly 65% [34] into hadrons and furthermore are always accompanied by a  $\nu_\tau$ , they are excluded in most top quark measurements which has a considerable impact on the branching fractions of the dileptonic and lepton+jets channel, respectively.

Decay mode	$W \rightarrow e\nu_e$	$W \rightarrow \mu\nu_\mu$	$W \rightarrow \tau\nu_\tau$	$W \rightarrow q\bar{q}'$
Branching fraction	$(10.71 \pm 0.16) \%$	$(10.63 \pm 0.15) \%$	$(11.38 \pm 0.21) \%$	$(67.41 \pm 0.27) \%$

Tab. 2.3.: Branching fractions for the  $W$  boson [34]. The leptonic branching fractions are split into the particular lepton flavours, whereas the hadronic one is combined for all possible hadronic final states, only.

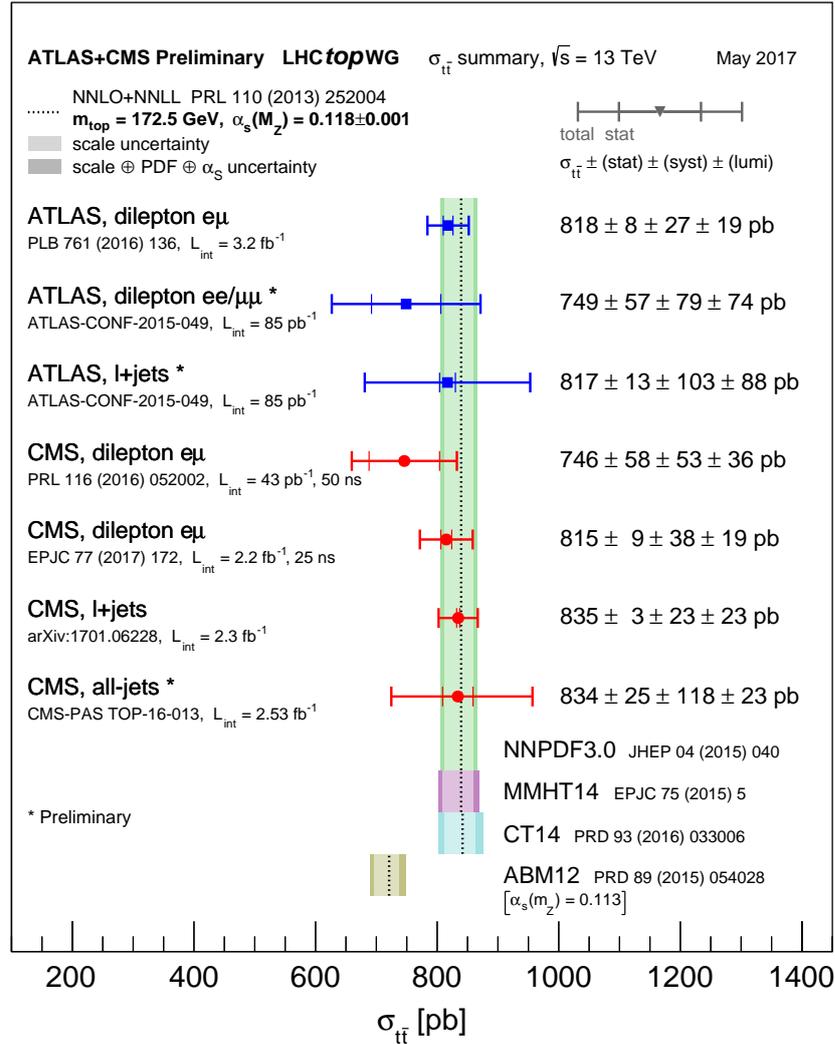


Fig. 2.8.: Summary of measurements of the top quark pair production cross section at the LHC for 13 TeV compared to the exact NNLO QCD calculation complemented with NNLL resummation [63]. The theory bands represent uncertainties due to renormalisation and factorisation scale, parton density functions and the strong coupling. The measurements and the theory calculation are quoted at  $m_{\text{top}} = 172.5$  GeV. [64]

### 2.3.2.1. The dileptonic channel

With a probability of approximately 33 % for a leptonic decay of the  $W$  boson the predicted branching ratio for the dileptonic channel is 10.6 % considering  $\tau$ -leptons and 5 % if explicitly taking only electrons and muons into account. Therefore, it has the smallest branching ratio of the three channels. The final state consists of two bottom quarks and a total of two high-energetic leptons and two neutrinos (cf. Fig. 2.9). Due to two neutrinos in the final state, the top quark pair system cannot be kinematically reconstructed so in this channel other elaborated techniques have to be applied when measuring top quark properties. However, the dileptonic channel offers a very clean event topology since it has a very low background contamination due to the four leptonic particles in the final state.

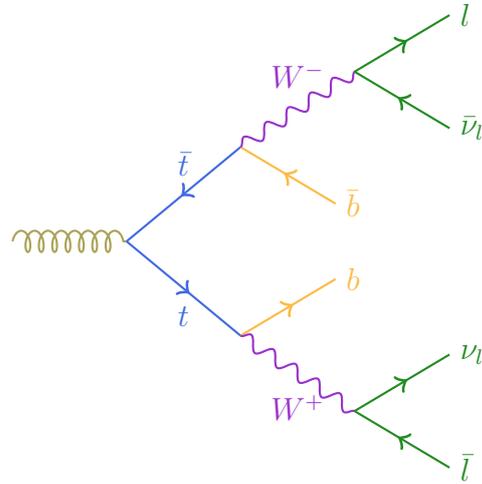


Fig. 2.9.: The dileptonic  $t\bar{t}$  decay

### 2.3.2.2. The lepton+jets channel

The predicted probability for the lepton+jets channel is 43.9 % and 30 %, respectively, dependent on whether taking  $\tau$ -leptons into account or not. Its final state is characterised by two bottom quarks, two light quarks, one high energetic charged lepton and one neutrino (cf. Fig. 2.10). Although a neutrino is present, in contrast to the dileptonic channel a kinematic reconstruction is feasible. The main background for this channel is the associated production of a  $W$  boson with jets, while a minor contribution originates from QCD multijet production. The lepton+jets channel is the preferred channel for many top quark measurements as it is well-balanced between a clean event signature and large branching fraction.

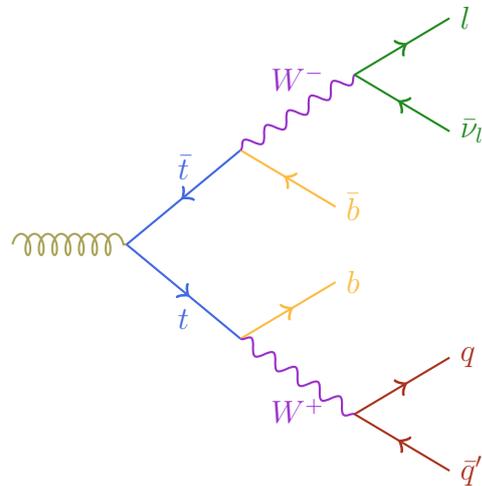


Fig. 2.10.: The lepton+jets  $t\bar{t}$  decay

### 2.3.2.3. The fully hadronic channel

In the fully hadronic channel, both  $W$  bosons decay hadronically which corresponds to a probability of 45.4%. Consequently, the fully hadronic channel has the largest branching fraction of all channels. The event signature is described by four light and two bottom quarks (cf. Fig. 2.11). In contrast to the two other channels, no high-energetic leptons or neutrinos are in the final state, allowing for a full reconstruction as all particles in the final state can be detected. The dominant background process for the fully hadronic channel is QCD multijet production. The absence of leptons and especially of neutrinos does not only have advantages as the rejection of background events containing many jets becomes very challenging.

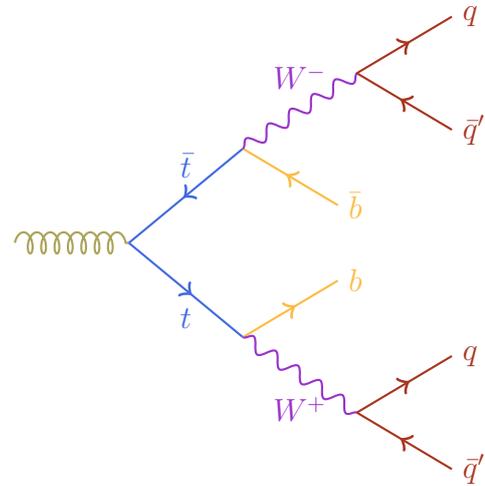


Fig. 2.11.: The fully hadronic  $t\bar{t}$  decay

## 2.4. Top quarks in association with a Z boson

Besides higher order corrections mediated by the strong force described in Sec. 2.3.1.2, particles can be also radiated via the electroweak interaction, therefore at a lower probability. For the associated production of top quarks with a  $Z$  boson, the  $Z$  boson has to be radiated from either one of the incoming quarks or from a top quark as gluons do not participate in weak interaction. Thus,  $Z$  bosons cannot be emitted by them. In case the  $Z$  boson is real, the quark it has been radiated off must be virtual as this process would be kinematically forbidden otherwise. This does not matter at all for a virtual  $Z$  boson.

The associated production of top quark pairs with a  $Z$  boson provides a direct measurement of the coupling between two of the heaviest known fundamental particles. As this coupling is not yet well constrained by direct measurements, its value can significantly vary when including many models of physics beyond the Standard Model, for example the decay of the top quark through a flavour changing neutral current mediated by a  $Z$  boson. Furthermore, it is an important irreducible background process to the associated production of top quark pairs with a Higgs boson, one of the main processes in which the Yukawa coupling of the top quark to the Higgs field is investigated, and to many searches in Supersymmetry providing a similar final state signature. The production cross section for top quark pairs in association with a  $Z$  boson at 13 TeV is  $\sigma_{t\bar{t}Z}^{\text{NLO}} = 839_{-92}^{+80} \text{ scale} \pm 25 \text{ (PDF)} \pm 25 \text{ } (\alpha_s) \text{ fb}$  [65] at next-to leading order in electroweak and QCD calculation. The most recent measurements of the ATLAS and the CMS collaboration yielding  $0.92 \pm 0.32 \text{ pb}$  and  $1.00_{-0.13}^{+0.15} \text{ pb}$ , respectively, are in quite good agreement with the theoretical prediction [66, 67]. Thus, to compare the associated production of top quarks pairs with a  $Z$  boson to bare top quark pair production, the two cross sections are divided yielding

$$\frac{\sigma_{t\bar{t}}^{\text{NNLO+NNLL}}}{\sigma_{t\bar{t}Z}^{\text{NLO}}} = 991.66 \quad , \quad (2.48)$$

stating, that the associated production of top quark pair production is almost thousandfold suppressed with respect to bare top quark pair production. A main reason is the much smaller electroweak coupling of coloured particles to  $Z$  bosons than their strong coupling to other coloured states. Furthermore, for the production of a real  $Z$  boson more energy is needed than for a virtual  $Z$  boson, which in turn occurs even less frequent. Exemplary leading order Feynman diagrams are shown in Fig. 2.12 depicting the radiation of a  $Z$  boson from a virtual top quark. In these studies the associated production of single top quarks with a  $Z$  boson is explicitly not taken as signal process but serves as an rather similar looking background process.

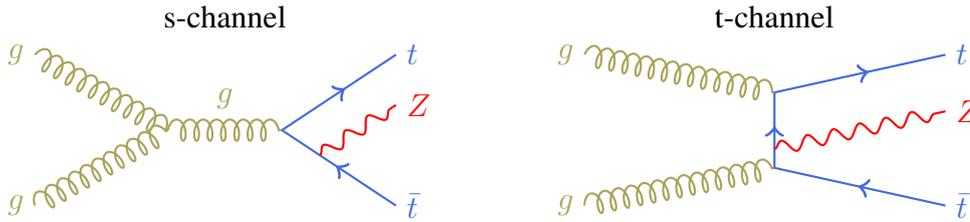


Fig. 2.12.: Exemplary Feynman diagrams for leading order production mechanisms of top quark pairs in association with a  $Z$  boson

### 2.4.1. 1-lepton channel

When investigating the associated production of top quark pairs with a  $Z$  boson, the decay of such a system can be divided into the top quark pair and the  $Z$  boson each with its respective decay products, resulting in different decay modes. First, the decay of the top quark pair can be categorised into three different decay channels according to as presented in Sec. 2.3.2. As the  $Z$  boson couples to all fundamental fermions, it can either decay in two quarks, two leptons or two neutrinos, respectively, with the respective branching ratios given in Tab. 2.4. This results in a total of five different channels categorised according to the number of leptons in the event signature. For the decay channels with two leptons the final signature is ambiguous as the leptons can either stem from a dileptonic decay of the top quark pair accompanied by a decay of the  $Z$  boson into a pair of quarks or neutrinos, or from a leptonic decay of the  $Z$  boson where the top quark pair has to decay in the fully hadronic channel. Instead, for one, three, four or even no leptons in the final state no ambiguities arise.

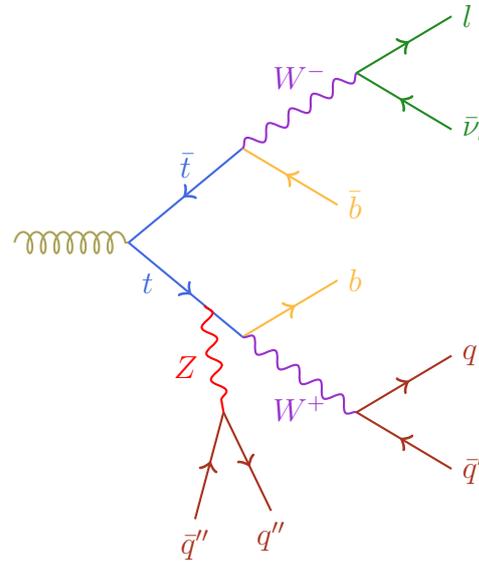


Fig. 2.13.: The  $t\bar{t}Z$  1-lepton channel

Most of the previous measurements on the associated production of top quark pairs with a  $Z$  boson have been done in multi-lepton channels [33, 68–71] as the event topology in them is much cleaner than for only one or even no lepton. For this analysis the so-called 1-lepton channel has

been chosen to investigate, consisting of a top quark pair which decays in the lepton+jets channel and a hadronically decaying  $Z$  boson (cf. Fig. 2.13). Its signature consists of a pair of opposite-sign and same-sign quarks originating from the hadronically decaying  $W$  boson and the  $Z$  boson, respectively, and of two bottom quarks from the decay of the two top quarks involved. In addition, a high-momentum lepton as well as a neutrino are contributed by the lepton+jets final state. In the 1-lepton channel, the considerably large branching fractions of the lepton+jets channel (30 %) and of the hadronic  $Z$  boson decay (70 %) are exploited in order to provide more statistics with respect to the multi-lepton channels. However, it suffers from a very large background contribution from the production of top quark pairs in association with jets arising from higher order corrections. This background process is expected to occur a thousand times more often than the actual signal as predicted by the ratio of their cross sections (cf. Eq. (2.48)). Further considerable backgrounds are the production of a  $W$  or  $Z$  boson in association with jets and the production of single top quarks. In addition, the associated production of top quarks pairs with a  $W$  or Higgs boson has the same final state, therefore called irreducible.

Decay mode	$Z \rightarrow l^+l^-$	$Z \rightarrow \nu\bar{\nu}$	$Z \rightarrow q\bar{q}$
Branching fraction	$(3.3658 \pm 0.0023) \%$	$(20.00 \pm 0.06) \%$	$(69.91 \pm 0.06) \%$

Tab. 2.4.: Branching fractions for the  $Z$  boson [34]. The leptonic branching fraction is averaged over the particular lepton flavours, whereas the one for hadrons and neutrinos is combined for all possible final states.



## 3. Experimental setup

### 3.1. Experiments at CERN

CERN<sup>(1)</sup> is the European Organisation of Nuclear Research founded in 1954 with its head and an accelerator and detector complex near Geneva in Switzerland. It operates the largest particle physical laboratory in the world and pursues a variety of experiments in different areas of particle physics. During the last 60 years the contributions of scientist from all over the world have led to remarkable progress in particle physics of which the probably most outstanding incident in the recent past was the discovery of a new boson matching to a Standard Model Higgs boson [28, 29]. Nowadays being one of the most important institutes in the world, various accelerators of different type delivered data for research over the years. It started with linear proton accelerators (LINAC 1 & 2) before moving over to synchrotron colliders as the Proton Synchrotron (PS) or the Large Electron Positron Collider (LEP) step by step. Besides the currently running experiments, most notably are the Super Proton Synchrotron (SPS) which enabled the discovery of the  $W$  and  $Z$  bosons [72–74], and LEP, the highest-energetic accelerator for leptons until today with a centre-of-mass energy of  $\sqrt{s} = 209$  GeV, providing high-precision measurements of the electroweak interaction and excellent confirmation of the Standard Model [75]. An overview of the CERN accelerator complex with all its current experiments at and besides the LHC is shown in Fig. 3.1.

### 3.2. The Large Hadron Collider

The Large Hadron Collider (LHC) is an international particle physics experiment located at CERN near Geneva, Switzerland, which was built to probe the Standard Model and to search for new physics phenomena with unprecedented energies. Since the centre-of-mass energy for LEP was limited by synchrotron radiation losses, an upgraded electron-positron accelerator turned out not to be feasible anymore and it was decided to install a hadron collider, the LHC. As charged particles radiate when being accelerated, they lose energy in each circulation. Therefore, the beams need to be constantly provided with energy. The energy loss per turn  $\Delta E$  of a particle with electric charge  $Ze$  and mass  $m$  is

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

with the beam energy  $E$ , the Lorentzfactor  $\gamma = E/m$  and  $R$  as the radius of the circuit. As the proton has a mass ( $m_{\text{proton}} = 938.272$  MeV [34]) approximately 2000 higher than the electron mass ( $m_{\text{electron}} = 0.511$  MeV [34]), proton-proton synchrotrons suffer much less from synchrotron radiation losses than lepton-lepton-colliders and therefore enable significantly higher beam energies. However, due to their high centre-of-mass energy hadron colliders are often called discovery machines, whereas lepton colliders provide high-precision measurements as there are no hadronic remnants within an event<sup>(2)</sup>. Operating since 2008, during its first data taking period (Run 1) lasting from 2009 to 2013, on the 30th of November 2009 the LHC became the world's highest-energy particle accelerator beating previous record of 0.98 TeV per beam held for eight years by the proton-antiproton collider Tevatron at Fermilab [77].

<sup>(1)</sup> The acronym stems from the original french name **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire.

<sup>(2)</sup> An event is defined as the crossing of two proton beams with a hard scattering interaction occurring.

### 3. Experimental setup

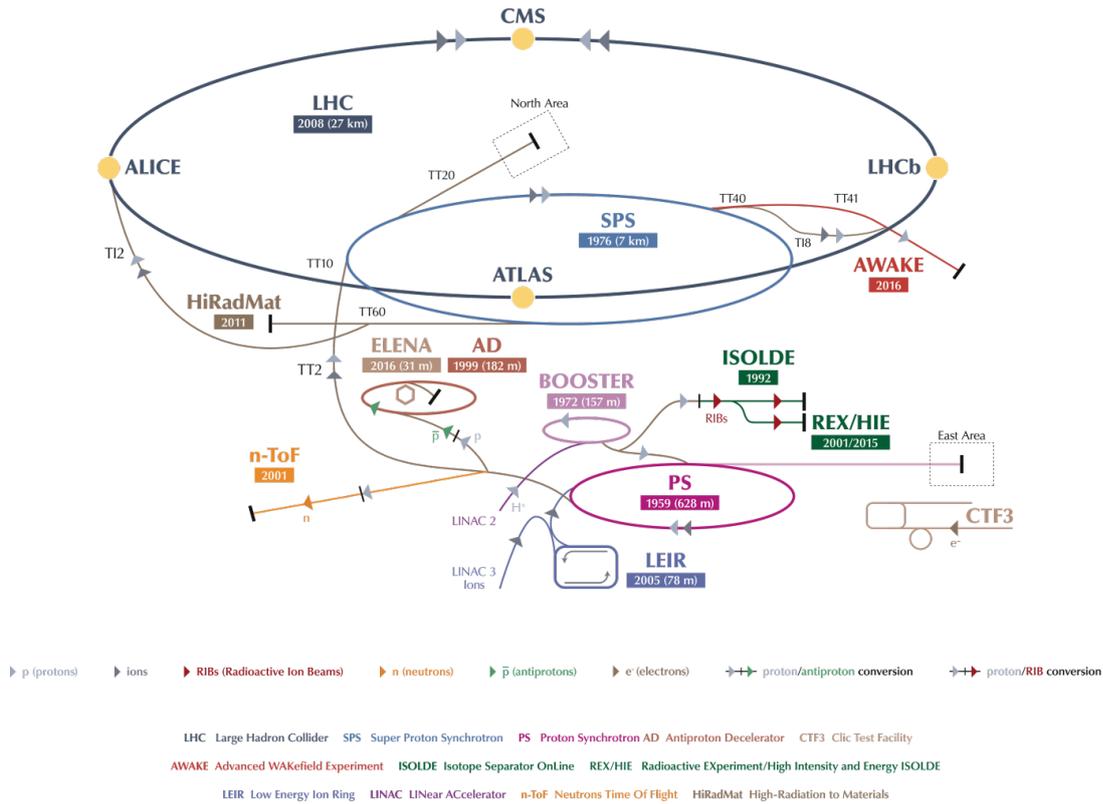


Fig. 3.1.: Schematic view of the CERN accelerator complex. The LHC ring is depicted in dark blue with its four major experiments indicated as yellow dots. Below the pre-accelerator chains for protons is shown including the LINAC 2, BOOSTER, PS and SPS, as well as for heavy ions comprises LINAC 3 and LEIR. Furthermore a few smaller experiments not primarily being concerned with the LHC are sketched [76].

The LHC is a circular machine with a circumference of 27 km and housed underground in the same tunnel as its predecessor LEP, therefore following the LEP tunnel layout with eight alternating straight sections and circular arcs. Designed to accelerate protons to a centre-of-mass energy of up to 14 TeV, it is also able to accelerate heavy ion nuclei to centre-of-mass energies of 2.76 TeV per nucleon. However, in order to reach such high energies the protons cannot be directly injected into the LHC, but have to run through a complex pre-accelerator chain employing already existing infrastructure at CERN of former experiments. In a first step, protons out of a bottle of hydrogen are accelerated to an energy of 50 MeV by the LINAC2 accelerator. Subsequently, the particles are inserted into the Proton Synchrotron Booster and the Proton Synchrotron to be further accelerated to 1.4 GeV and 25 GeV, respectively. In the last step of the pre-acceleration process the proton energy gets increased to 450 GeV by the Super Proton Synchrotron from where they are injected into the great LHC accelerator ring. There the protons get accelerated to their final collision energy which amounts to 6.5 TeV during the 2015 and 2016 data taking period, corresponding to a proton-proton centre-of-mass energy of 13 TeV [78].

After the pre-acceleration, the protons are filled into two parallel vacuum-kept beam pipes of the LHC where the proton beams rotate clockwise and counter-clockwise, respectively. The particles are accelerated to almost velocity of light by eight radiofrequency (RF) cavities per beam, which also compensate the energy loss due to synchrotron radiation by oscillating at a frequency of 400 MHz. Furthermore, these cavities conserve the bunch structure of the proton beams due to

a local acceleration or deceleration depending on whether the particles are too slow or too fast in the electric field of the cavities. In order to bend particles on a circular path, in total 1232 15 m long dipole magnets are employed generating magnetic fields peaking at 8.33 T. These magnets use niobium-titanium cables that become superconductive at temperatures below 10 K. The steering, adjustment and focussing of the proton beams is done by using a multitude of differently designed magnets, such as quadrupole magnets or higher-order multipoles. To achieve such high performance both the radiofrequency cavities and the magnets installed along the beam line have to be superconductive. Therefore, the radiofrequency cavities are cooled down to 4.5 K in order to prevent energy loss from electrical resistance, whereas the magnets are even further cooled to 1.9 K by a gigantic cryogenic system filled with liquid helium, for only superconductive magnets enable currents of up to 12 000 A needed for those strong magnetic fields.

The proton beams cross each other at several points around the LHC accelerator ring, on the one hand without being collided in order to correct for small spatial displacements due to the slightly different radii of the inner and out beam line, and on the other hand to collide them in a controlled manner at four points within one of the detectors at the LHC. At these interaction points, the proton beams are magnetically deflected such that they collide almost head-on. Concurrently, the beam diameter is kept as small as possible in order to increase the instantaneous luminosity. The total rate of proton-proton collisions  $f$  is given by

$$f = \sigma \cdot L \quad , \quad (3.2)$$

with the instantaneous luminosity  $L$  describing the number of particle collisions per time and per area. For particle colliders as the LHC it can be written as

$$L = \frac{nN_1N_2}{4\pi\sigma_x\sigma_y} \cdot \nu \quad , \quad (3.3)$$

where  $\nu$  denotes the revolution frequency of the  $n$  proton bunches which are brought to collision [79]. At the LHC the bunch-spacing is limited to 25 ns which corresponds to a collision rate of 40 MHz in the centre of each experiment.

$N_1$  and  $N_2$  are the number of protons the colliding bunches consist of per beam, and  $\sigma_x$  and  $\sigma_y$  are the widths of the proton beams, determined in so-called van-der-Meer scans [80]. In such a scan the two beams are moved through each other in order to determine the size of the beams in the interaction region. The cross section  $\sigma$  is a measure of the probability of a collision occurring at a bunch crossing of protons with a certain energy, to be interpreted as an effective area of a proton that can be hit by another proton. It is the sum of all cross sections corresponding to the processes that can happen in proton-proton collisions. The total number of events of a process  $i$  is

$$N_i = \sigma_i \cdot \mathcal{L} \quad , \quad (3.4)$$

where the time-integrated luminosity

$$\mathcal{L} = \int L dt \quad (3.5)$$

represents the amount of data taken within a certain time interval  $t$  [34].

In order to adapt the calculation of the instantaneous luminosity (cf. Eq. (3.3)) to real life, some correction factors have to be applied. Besides the relativistic  $\gamma$ -factor accounting for the relativistic proton velocities, the normalised transverse emittance  $\varepsilon$  measures the beam quality and amplitude function  $\beta^*$  quantifies the squeezing of the beam optics at the interaction point, indicated by the star. Additionally, as the beams usually collide at an angle  $\theta_c \neq 0$ , the luminosity has to be corrected with respect to actual head-on collisions in terms of the geomtric luminosity

reduction factor

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

which is composed of  $\theta_c$ , the root mean square of the bunch length  $\sigma_z$  and the root mean square of the transverse beam size at the interaction point  $\sigma^*$ . Taking into account all these corrections, Eq. (3.3) can be rewritten as

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

using the beam width can be expressed in terms of emittance and amplitude function,

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

Besides two smaller experiments, LHCf<sup>(3)</sup> and TOTEM<sup>(4)</sup>, four major experiments are associated to the LHC accelerator. The smallest detector at the LHC is LHCf, designed to study particles generated almost directly in line with the colliding proton beams and look for cosmic rays. It shares its location with the ATLAS experiment which will be explained later on. ALICE<sup>(5)</sup> is specifically designed to study quark-gluon plasma, the state of matter straight after the big bang in heavy-ion collisions. The other big special-purpose detector at the LHC is LHCb<sup>(6)</sup> which explores the origin of the matter-antimatter imbalance by studying the properties of  $B$ -hadron decays. A rather new supplement to LHCb is the MoEDAL<sup>(7)</sup> detector which basically looks for magnetic monopoles and microscopic black holes. ATLAS<sup>(8)</sup> and CMS<sup>(9)</sup> are the two general-purpose detectors located diametrically opposite to each other at the LHC ring. They cover a broad field of particle physics and benefit from cross-checking each other's results. Near to CMS the TOTEM experiment has its place at the beam pipe. It measures remnants from deep-inelastic particle collisions scattered along the beam pipe and performs studies on the measurement of the proton size [78].

### 3.3. The ATLAS experiment

The ATLAS experiment [81] is a forward-backward symmetric multi-purpose particle detector designed for hadron collisions at high luminosities. It is located underground in one of the experimental caverns of the LHC ring, more precisely at LHC point 1, one of the four sites at the LHC where protons and heavy ions are brought to collision. Designed to cover a wide range of particle physics topics like top quark and Higgs boson properties or the search for new physics<sup>(10)</sup> phenomena, e.g. Supersymmetry (SUSY), in the most possible comprehensive way, it has to fulfil a number of various requirements whereof a few will be mentioned in the following. Besides an excellent electromagnetic calorimeter for electron and photon identification as well as a full-coverage hadronic calorimeter for the measurement of jets and missing transverse energy, a high-precision

---

<sup>(3)</sup> Large Hadron Collider Forward located

<sup>(4)</sup> TOTAL Elastic and diffractive cross section Measurement

<sup>(5)</sup> A Large Ion Collider Experiment

<sup>(6)</sup> Large Hadron Collider Beauty

<sup>(7)</sup> MONopole and EXotics Detector At the LHC

<sup>(8)</sup> A Toroidal LHC Apparatus

<sup>(9)</sup> Compact Muon Solenoid

<sup>(10)</sup> Physics not described (beyond) the Standard Model of Particle Physics (BSM)

momentum measurement for both muons and other charged particles must be guaranteed. Similarly, an almost full coverage of the solid angle as well as an efficient tracking for both high and low transverse momenta is essential. In addition, a reasonable power for identifying  $\tau$ -leptons and heavy flavour particles is supplied, even a full event reconstruction capability. Its overall concept is to provide a full coverage for all electromagnetically and strongly interacting final state particles and their kinematics.

As depicted in Fig. 3.2, the ATLAS detector has a cylindrical shape situated along the beam pipe. With a length of 46 m and a diameter of 25 m, weighing approximately 7000 tonnes it is the biggest but by far not the heaviest of the LHC experiments. Surrounding the interaction point it is divided into a central part, the barrel, where most of the components are aligned concentrically around the beam pipe, and two disk-shaped parts at both ends of the cylinder, the so-called end-caps. The ATLAS detector is separated into three different detector layers, the inner detector (ID), the calorimeter (CAL) and the muon spectrometer (MS), which are concentrically placed around the interaction point. Superconductive magnets provide magnetic fields curving the tracks of charged particles and thus providing a measurement of the momentum. Data recorded from the various subdetectors is used to identify particles by reconstructing their tracks and measuring their energy. A specifically designed trigger and data acquisition system has to be made use of in order to handle the enormous amount of data taken.

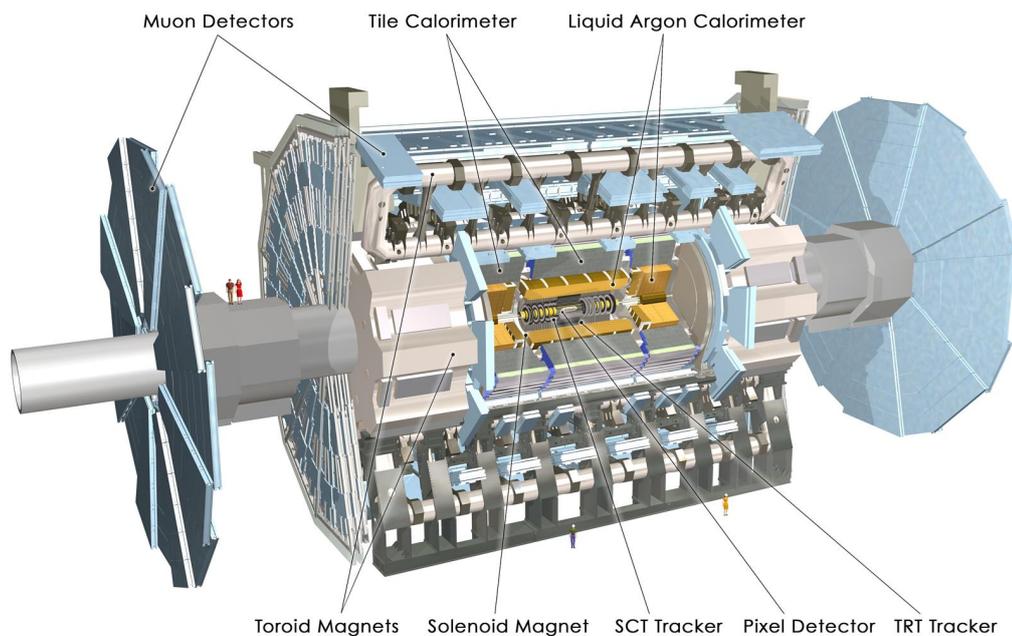


Fig. 3.2.: Computer generated image of the whole ATLAS detector showing its main components, the inner detector, the calorimeter system, the muon spectrometer and the magnet system. [82]

### 3.3.1. The ATLAS coordinate system

The ATLAS detector uses a right-handed coordinate system with the nominal interaction point in the centre of the detector defined as origin and the  $z$ -axis oriented along the beam line of the LHC. In the  $x$ - $y$ -plane perpendicular to the beam line, the positive direction of the  $x$ -axis is pointing towards the centre of the LHC ring and the positive  $y$ -direction pointing upwards. The transverse

components of momentum and energy,  $p_T$  and  $E_T$ , as well as the missing transverse energy  $E_T^{\text{miss}}$  (cf. Sec. 5.6) are defined as their respective projections to the  $x$ - $y$ -plane,

$$p_T := \sqrt{(p_x)^2 + (p_y)^2} \quad (3.9a)$$

$$E_T := \sqrt{m^2 + (p_T)^2} \quad . \quad (3.9b)$$

Due to its shape, a much more common and convenient way to describe objects' locations and particle trajectories in the detector are cylindrical coordinates. The azimuthal angle  $\phi$  is measured around the beam axis in the transverse plane, with  $\phi = 0$  corresponding to the positive  $x$ -axis and increasing clock-wise looking into the positive  $y$ -direction<sup>(11)</sup>, while the polar angle  $\theta$  defines the separation angle from the beam line. However, instead of the polar angle, a quantity defined as

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

is often used as a measure of the forward direction the pseudorapidity  $\eta$  [81]. Therefore, the coverage of each subdetector is commonly given in ranges of  $|\eta|$ , where  $\eta = 0$  is equivalent to  $\theta = \pi/2$  (transverse plane) and  $\eta = \pm\infty$  corresponds to  $\theta = 0^\circ$ , which means parallel to the beam pipe. High values of  $|\eta|$  are also called the forward region. In case of massive objects the Lorentz invariant rapidity  $y$  [81], defined as

$$y = \frac{1}{2} \ln \left[ \frac{E + p_z}{E - p_z} \right] \quad (3.11)$$

is used instead. In the limit of high energies ( $p \gg m$ ) the pseudorapidity  $\eta$  approximately equals the rapidity  $y$  [34]. Differences in rapidity (and therefore in the limit equivalently in pseudorapidity) are invariant under boosts in  $z$ -direction. The distance  $\Delta R$  in the angular space ( $\eta$ - $\phi$ ) between two objects in the detector is hence expressed as [81]

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad . \quad (3.12)$$

#### 3.3.2. The ATLAS magnet system

In order to measure the charge of particles travelling through the detector, a hybrid system of four large superconductive magnets provide several magnetic fields [83]. As shown in Fig. 3.3, it consist of the central solenoid (CS) magnet enclosing the inner detector in the barrel, and three toroid magnets around the muon spectrometer, located in both the barrel region and the end-caps. The magnetic system of ATLAS as a whole is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ. It is made up of aluminium-stabilised cryogenically cooled niobium-titanium alloy and cooled down to 4.5 K.

The central ATLAS solenoid extends over a length of 5.3 m and has a bore of 2.44 m. Designed to provide a homogeneous magnetic field of 2 T in the central tracking volume of the inner detector, a peak value of 2.6 T can be maximally reached [84].

The ATLAS toroid magnets consist of eight large air-core coils each, assembled radially and symmetrically around the beam axis, delivering a peak magnetic field of 3.9 T and 4.1 T in the barrel and end-cap toroids, respectively. The barrel toroid (BT) with a length of 25.3 m, an outer diameter of 20.1 m and an inner diameter of 9.4 m is by far the biggest component of the ATLAS magnet system. In contrast, the end-cap toroid magnets are much more compact, exhibiting a

---

<sup>(11)</sup>  $\phi \in [-\pi, +\pi]$

length of 5 m, an outer diameter of 10.7 m and an inner diameter of 1.65 m. They are inserted in the barrel toroid at each end and line up with the central solenoid [85]. The end-cap toroid (ECT) system is rotated by  $22.5^\circ$  with respect to the barrel coil system in order to provide radial overlap and to optimise the bending power in the interface regions of both coil systems. However, the bending power decreases in the transition regions, where the two magnets overlap, which corresponds to an  $\eta$ -region of  $(1.3 < |\eta| < 1.6)$  [86].

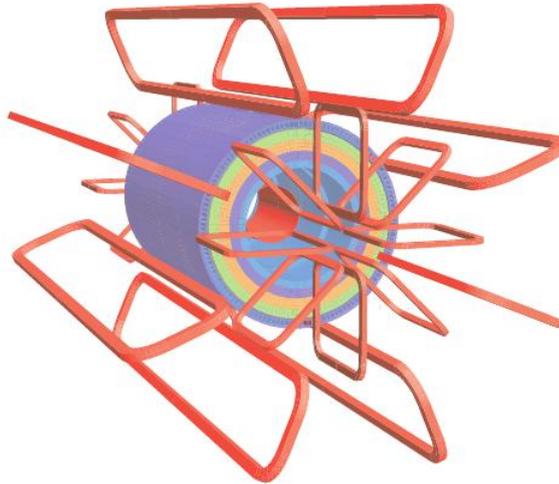


Fig. 3.3.: The ATLAS magnet system coloured in red [81]. The picture shows the barrel coils, the end-cap toroidal coils and the central solenoid embedded in the calorimeter.

The magnetic fields provided by the ATLAS magnet system are essential for the momentum measurements in the inner detector and the muon spectrometer. According to the Lorentz force

$$\mathbf{F}_L = q \cdot (\mathbf{v} \times \mathbf{B}) \quad (3.13)$$

charged particles moving in a magnetic field are deflected, hence their tracks left in the detector are bent from a straight path. As the angular deflection of a particle is proportional to its momentum and the sign of its charge is defined by the direction of the curvature with respect to the polarity of the magnetic field, the momentum and sign of charged particles can be therefore determined by measuring their tracks. Using this method of momentum measurement, the relative resolution is proportional to the momentum and inverse proportional to the magnetic field [87]. In order to keep the resolution small enough when dealing with particle energies as delivered at the LHC, the magnetic fields have to be so strong that superconductive magnets have to be used to provide them.

### 3.3.3. The ATLAS inner detector

The ATLAS inner detector has been designed to provide precise vertex identification and high-precision momentum measurements in order to meet the requirements imposed by the benchmark physics processes and high-precision measurements, given the very large track density at the LHC<sup>(12)</sup>. These features are provided by semiconducting pixel and silicon microstrip layers, used in conjunction with a straw tube tracker. The layout of the inner detector is shown in Fig. 3.4. Enclosing the beam pipe within a range of  $|\eta| < 2.5$  it is the most central element of the ATLAS

<sup>(12)</sup> Every 25 ns approximately 1000 particles emerge from the collision point within  $|\eta| < 2.5$ .

detector starting only 3 cm from the beam axis and has an overall length of 6.2 m as well as an outermost radius of 105 cm. The precision tracking elements are located within a radius of 56 cm, followed by detector modules providing continuous tracking.

The ATLAS inner detector is immersed in a 2 T magnetic field generated by the central solenoid (cf. Sec. 3.3.2) in order to perform tracking and identification of charged particles as well as to determine primary<sup>(13)</sup> and secondary<sup>(14)</sup> vertices, all with high spatial and momentum resolution. From the beam pipe outwards, it consists of three independent parts, the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). In the barrel region the different detector parts are arranged in concentric cylinders around the beam axis, whereas in the end-caps the detector elements are mounted on disks perpendicular to the beam line, as shown in Fig. 3.5 depicting both the barrel and the end-cap region of the inner detector separately. This layout allows a full tracking coverage over  $|\eta| < 2.5$ , including impact parameter measurements and vertex identification for heavy-flavour or  $\tau$ -tagging [88, 89].

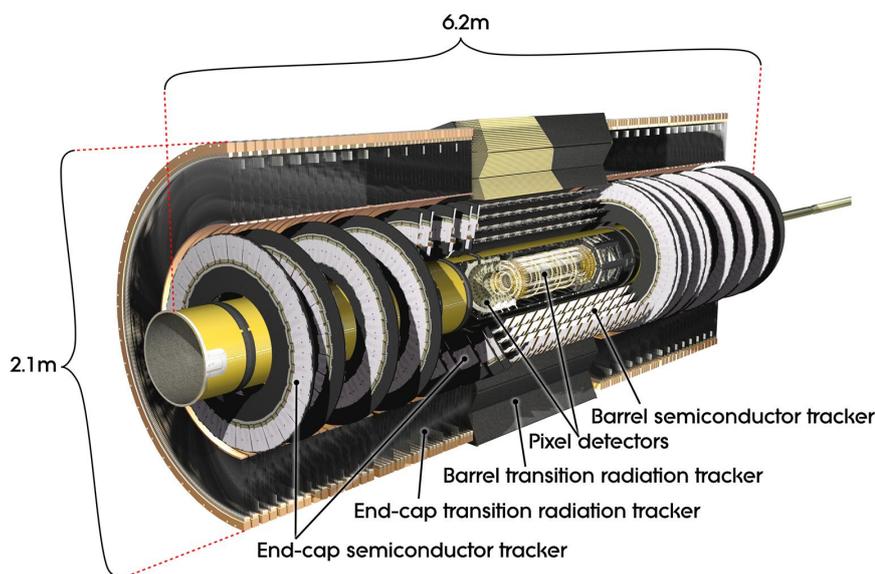


Fig. 3.4.: Computer generated cut-away view of the ATLAS inner detector. [90]

#### 3.3.3.1. The ATLAS pixel detector

The pixel detector consists of four, originally three layers of silicon pixel detectors placed around the beam axis at roughly 3 cm, 5 cm, 9 cm, and 13 cm in the barrel covering  $|\eta| < 1.7$ , and of five disks of silicon pixel detectors between radii of 11 and 20 cm in the end-caps, covering  $1.7 < |\eta| < 2.5$  (cf. Fig 3.5). Thus the pixel detector provides at least four high-precision measurements with fine granularity for each charged particle as close to the interaction point as possible over the full acceptance. Therefore, an important task of the pixel detector is the reconstruction of secondary vertices which are crucial in tagging jets containing  $B$ -hadrons. The pixel layers are segmented in  $R$ - $\phi$  and  $z$  direction, respectively, and have a minimum pixel size in  $R$ - $\phi \times z$  of  $50 \times 400 \mu\text{m}^2$ .

<sup>(13)</sup> Point situated on the beam line which indicates, from extrapolation of the measured tracks, a parton-parton interaction within the single proton-proton collision of interest.

<sup>(14)</sup> Crossing of tracks which is displaced with respect to the beam line and thus can not be associated to a primary parton-parton collision, as it mostly stems from secondary decays of particles with a high lifetime, e.g. bottom quarks, within the decay chain.

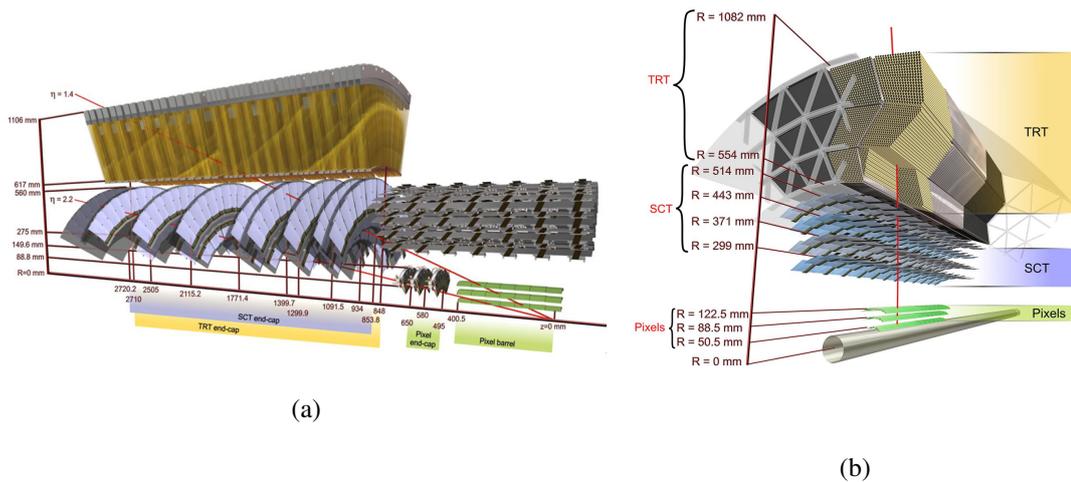


Fig. 3.5.: Computer generated schematic view of the end-cap (Fig. (a)) and barrel (Fig. (b)) part of the ATLAS inner detector [90]. Note that in these images the new IBL layer is missing.

For both barrel and end-caps the spatial resolution of the pixel detector is approximately  $10\ \mu\text{m}$  in the  $R$ - $\phi$ -plane and  $115\ \mu\text{m}$  in  $z$ -direction [91].

During the preparation of the LHC Run 2 operation, an additional pixel layer, the so-called insertable B-layer (IBL) has been installed directly on the beam pipe with a coverage of  $|\eta| < 2.5$ . The reason was the increase in track and vertex reconstruction resolution in expectation of the increase in luminosity, which came along with the raise of the centre-of-mass energy from  $\sqrt{s} = 8\ \text{TeV}$  to the current  $\sqrt{s} = 13\ \text{TeV}$  after Run 1. The intrinsic resolution of the newly installed IBL is  $8\ \mu\text{m}$  ( $R$ - $\phi$ ) and  $40\ \mu\text{m}$  ( $z$ ) [92].

### 3.3.3.2. The ATLAS semiconductor tracker

Based on the same technology, the semiconductor tracker is concentrically arranged around the pixel detector, and therefore completely surrounding it. It consists of eight silicon strip layers around the barrel to be crossed by each track at distances between 30 cm and 51 cm providing four precision space point measurements in the  $R$ - $\phi$  and  $z$ -coordinates, and nine disks in each end-cap at distances from the interaction point of 85 cm to 273 cm. In the barrel, the detector is composed of four stereo silicon strip modules each consisting of two layers rotated by 40 mrad to each other in order to obtain the  $z$ -measurement. The coverage of the barrel part of the semiconductor tracker is  $|\eta| < 1.4$  while with the end-cap modules the remaining solid angle is covered, i.e.  $1.4 < |\eta| < 2.5$ . The intrinsic accuracy is both for barrel and end-caps  $17\ \mu\text{m}$  in the  $R$ - $\phi$ -plane and  $580\ \mu\text{m}$  for the  $z$ -direction in the barrel and the  $R$ -direction in the end-cap disks, respectively.

As both the pixel detector and the semiconductor tracker are made of silicon, they are prone to irradiation. Consequently, electronic noise from radiation damage is suppressed by cooling the detector modules down to below  $-5\ ^\circ\text{C}$ .

### 3.3.3.3. The ATLAS transition radiation tracker

The third part of the ATLAS inner detector is the transition radiation tracker which can operate at the very high rates expected at the LHC. It is based on 4 mm thin-walled, gas-filled proportional drift-tubes, so-called straws, which are equipped with a  $30\ \mu\text{m}$  diameter gold-plated tungsten wire

leading to fast response and good mechanical and electrical properties. A large number of hits<sup>(15)</sup> enables track-following up to  $|\eta| = 2.0$ . In the barrel region the straws are aligned parallel to the beam pipe and are 144 cm long with their wires divided into halves at approximately  $\eta = 0$ , while in the end-cap regions, the 37 cm long straws are arranged radially in a total of 18 wheels (cf. Fig. 3.5). The transition radiation tracker provides tracking information in the transverse plane only, with an intrinsic accuracy of 130  $\mu\text{m}$  per straw.

A main purpose of the ATLAS transition radiation tracker is the discrimination between electrons and more heavy particles as pions and kaons by employing xenon gas due to the detection of transition-radiation photons created in radiators between the straws.

#### 3.3.4. The ATLAS calorimeter system

The ATLAS calorimeter system consists of two main parts: The electromagnetic calorimeter (ECAL) which is designed to measure the energy of mainly electromagnetically interacting particles like photons, positrons and electrons, and the hadronic calorimeter (HCAL) constructed to determine the energy of hadrons as protons, neutrons or  $\pi$ -mesons which interact via the strong force. Both calorimeters are so-called sampling type calorimeters meaning that they consist of alternating layers of high-density active material in which incoming particles induce showers, and readout layers where these showers are measured. The ATLAS calorimeter system covers a range of  $|\eta| < 4.9$  using different techniques suited to the widely varying requirements of the physics processes of interest and of the radiation environment in this large  $\eta$ -range. Due to the full coverage in the azimuthal angle it also allows for the measurement of missing transverse energy (cf. Sec. 5.6). Besides good containment for electromagnetic and hadronic showers, calorimeters must limit the punch-through into the muon system. A schematic view of the ATLAS calorimeter system is presented in Fig. 3.6.

##### 3.3.4.1. The ATLAS electromagnetic calorimeter

The ATLAS electromagnetic calorimeter [94] uses lead as absorber material and liquid argon (LAr) as active medium. Its accordion-shaped kapton electrodes and lead absorber plates provide full coverage in  $\phi$  and include a total pseudorapidity region of up to  $|\eta| < 3.2$ . Incoming particles passing through the detector produce electromagnetic showers which ionise the argon atoms in the active layers. The energy of the incoming particles is then proportional to the number of created charges. The energy of electromagnetically interacting particles can be measured by counting the number of ionised atoms within the active material. Position and energy of the incoming particle as well as the shape of the shower can be determined by the hits in the individual calorimeter cells.

The electromagnetic calorimeter is divided into a barrel part extending up to  $|\eta| < 1.475$ , and two end-cap components ranging from  $1.375 < |\eta| < 3.2$ , each housed in their own cryostat with the overlap between the modules ensuring a continuous coverage in  $\eta$ . The barrel calorimeter consists of two identical half-barrels separated by a small gap of 4 mm at  $z = 0$ . Each end-cap calorimeter is divided into two coaxial wheels, where the inner (outer) wheel covers the region  $2.5 < |\eta| < 3.2$  ( $1.375 < |\eta| < 2.5$ ). Over the pseudorapidity range of  $|\eta| < 1.8$  the actual calorimeter is preceded by an instrumented layer of liquid argon, the so-called presampler, which is used to correct for energy losses prior to the calorimeter.

The total thickness of the EM calorimeter is  $> 22$  radiation lengths ( $X_0$ ) in the barrel and  $> 24$  radiation lengths in the end-caps. The barrel electromagnetic calorimeter is embedded in the vacuum of the barrel cryostat surrounding the inner detector (cf. Fig. 3.3) and hence shares its vacuum

---

<sup>(15)</sup> typically 36 hits per track

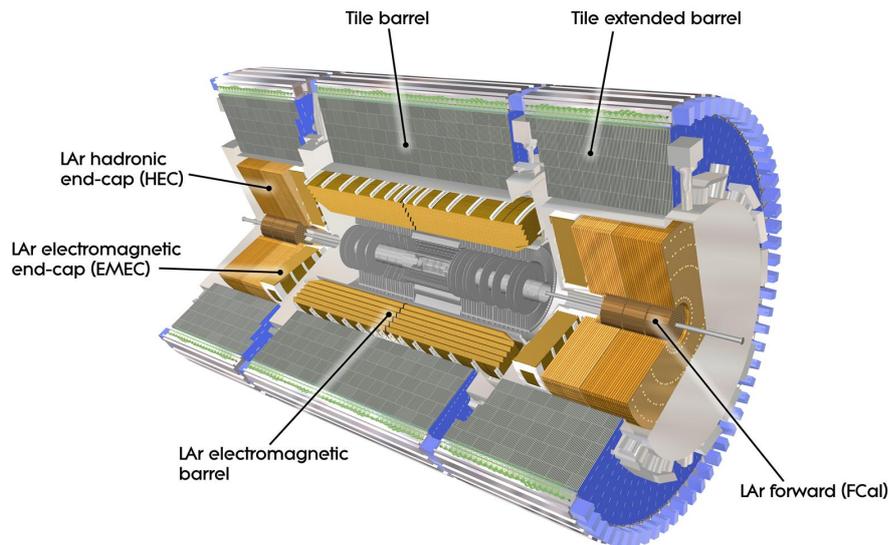


Fig. 3.6.: Computer generated schematic view of the ATLAS calorimeter [93]. The electromagnetic calorimeter encloses the inner detector in the barrel region which itself is embedded in the tile calorimeter. The end-caps comprise the forward calorimeters which are surrounded by the electromagnetic and hadronic end-cap calorimeters. Together they are encased by the extended tile calorimeters.

vessel with the central solenoid, while two end-cap cryostats house the end-cap electromagnetic and hadronic calorimeters as well as the integrated forward calorimeter.

### 3.3.4.2. The ATLAS hadronic calorimeters

The ATLAS hadronic calorimeter is subdivided into three components, the tile calorimeter, the hadronic end-cap calorimeter and the forward calorimeter, which unlike the electromagnetic calorimeter are built from different materials. With all these parts it covers the range  $|\eta| < 4.9$ .

#### The tile calorimeter

The main component of the ATLAS hadronic calorimeter is the tile calorimeter [95]. It totally surrounds the electromagnetic calorimeter and uses steel absorber plates and scintillating tiles as the active material. Attached to the tiles are wavelength-shifting fibres which convert the light induced by ionising particles into visible light which is detected by photomultiplier tubes. The energy of impinging particles can then be deduced from the total amount of light collected.

The tile calorimeter is composed of one barrel and two extended barrels. The barrel cylinder covers a pseudorapidity of  $|\eta| < 1$  which overlaps with the coverage of the extended barrels,  $0.8 < |\eta| < 1.7$ , preventing gaps in  $\eta$ . The tile calorimeter extends radially from an inner radius of 2.28 m to an outer radius of 4.25 m and is segmented in depth into three layers at approximately 1.5, 4.1 and 1.8 interaction lengths ( $\lambda$ ) for the barrel and 1.5, 2.6 and 3.3 interaction lengths for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is  $9.7\lambda$  at  $\eta = 0$ .

#### The liquid-argon hadronic end-cap calorimeters

The ATLAS hadronic end-cap calorimeter (HEC) [94] consists of two independent wheels of outer radius 2.03 m per end-cap, both located directly behind the end-cap electromagnetic calorimeter and sharing the same LAr cryostats. In contrast to the tile calorimeter, the hadronic end-cap calorimeter uses copper as absorber medium and, as the electromagnetic calorimeter, liquid argon for ionisation. Covering a range of  $1.5 < |\eta| < 3.2$  it overlaps with the tile and the forward calorimeter.

#### The liquid-argon forward calorimeter

The ATLAS forward calorimeter (FCal) [94] encloses the beam pipe with its front face at approximately 4.7 m away from the interaction point, and is integrated into the end-cap cryostats. In order to avoid back-scattering of neutrons it is shifted with respect to the electromagnetic calorimeter front face by about 1.2 m. This severely limits the depth of the calorimeter and therefore requires a high-density design. With a length of approximately 10 interaction lengths it covers the range in pseudorapidity of  $3.1 < |\eta| < 4.9$ . Each end-cap module consists of three parts which all use liquid argon as the active medium. The first part closest to the interaction point is a copper calorimeter mainly designed to measure electromagnetic interactions followed by two calorimeters made out of tungsten predominantly intended for the measurement of the energy of hadronically interacting particles.

##### 3.3.4.3. The calorimeter resolution

The granularity of the ATLAS calorimeter system is defined by the size of the respective calorimeter cells in the different detector regions and calorimeter parts. In the electromagnetic barrel calorimeter the cell size ranges from  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  to  $0.025 \times 0.1$  and  $0.075 \times 0.025$ . In the end-cap regions the electromagnetic calorimeter has a cell size of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  up to  $0.1 \times 0.1$ , whereas in the hadronic calorimeter cells extend from  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  to  $0.2 \times 0.2$ . In the central and extended barrel regions cells of the hadronic tile calorimeter extend from  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  to  $0.2 \times 0.1$ . The granularity of the forward calorimeter is between  $\Delta X \times \Delta y = 3.0 \text{ cm} \times 2.6 \text{ cm}$  and  $5.4 \text{ cm} \times 4.7 \text{ cm}$ .

##### 3.3.5. The ATLAS muon spectrometer

An important and therefore very sophisticated part of the ATLAS detector is the ATLAS muon spectrometer or muon system [96], a set of detectors specifically designed to both trigger on muons and to measure their tracks and therefore momenta within  $|\eta| < 2.7$ . With a mean lifetime of about  $2.2 \mu\text{s}$  [34] muons are the only detectable particles traversing the so far explained parts of the detector since they are minimally ionising particles and therefore lose only little energy in the inner detector and calorimeters. Hence, their measurement is performed, based on the deflection of their tracks in the magnetic fields provided by the large air-core toroid magnets in the outermost layer of the detector. Up to  $|\eta| = 1.4$ , magnetic bending is provided by the large barrel toroid. In the range of  $1.6 < |\eta| < 2.7$  the muon tracks are bent by the two smaller end-cap toroids inserted into both ends of the barrel toroid. Within  $1.4 < |\eta| < 1.6$ , usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. The magnetic field is mainly oriented perpendicularly to the flight path of the muons in order to maximise deflection and thus resolution.

The overall layout of the ATLAS muon spectrometer is shown in Fig. 3.7. In the barrel region three concentrically mounted cylindrical layers of high-precision tracking chambers, so-called

monitored drift tubes (MDT) and resistive plate chambers (RPC), surround the beam axis at radii of about 5 m, 7.5 m and 10 m, providing a coverage of  $|\eta| < 1.05$ . Additionally, at large pseudorapidities close to the interaction point, cathode strip chambers (CSC) are installed on the innermost wheel. With a coverage of  $1.0 < |\eta| < 2.7$ , at both end-caps monitored drift tubes in conjunction with thin gap chambers (TGC) are arranged on four wheels, located at distances of 7.4 m, 10.8 m, 14 m and 21.5 m from the interaction point. The individual modules are arranged symmetrical in  $\phi$  and split up into eight sections.

On top of that, the ATLAS muon spectrometer has its own trigger system covering the pseudorapidity range of  $|\eta| < 2.4$ , including resistive plate chambers in the barrel and thin gap chambers in the end-caps. The muon trigger system serves a threefold purpose: It has to identify bunch-crossings requiring a time resolution better than the LHC bunch spacing of 25 ns, further it must provide well-defined thresholds on the transverse muon momentum, and measure the muon coordinate in the direction orthogonal to that determined by the high-precision tracking chambers [97].

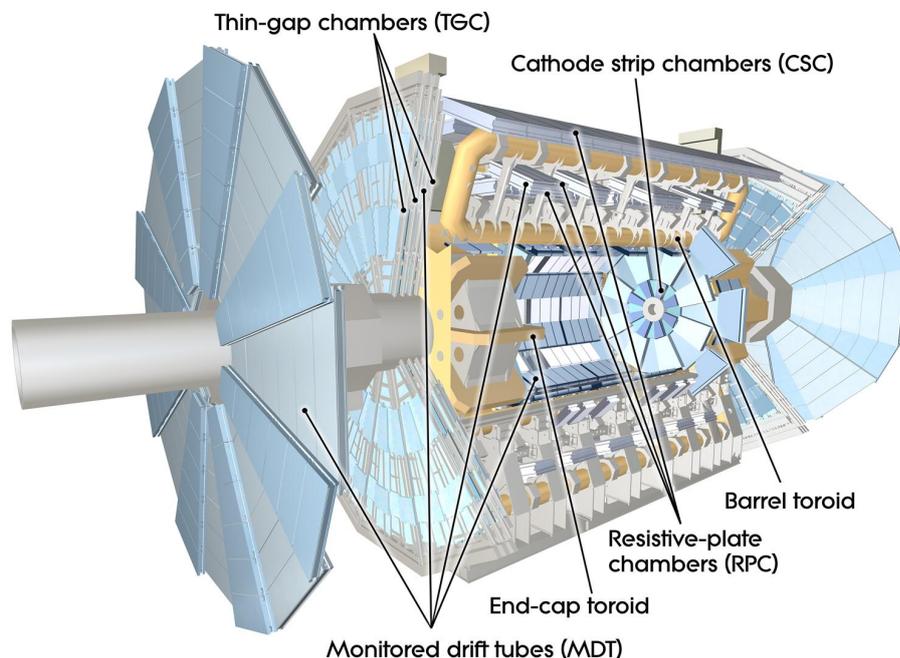


Fig. 3.7.: Computer generated schematic view of the ATLAS detector focussing on the muon spectrometer.

### 3.3.5.1. The monitored drift tubes

Over the full acceptance range of the muon system, the monitored drift tubes provide high-precision tracking measurements in the  $\eta$ -direction, but do not deliver any measurement in  $\phi$ . The chambers consist of three to eight layers of aluminium drift tubes varying in length between 70 cm and 630 cm and filled with a gaseous mixture of argon and carbon dioxide at high pressure, trespassing muons create electron-ion pairs by ionising the gas. Then, the electrons drift towards the central gold-plated tungsten rhenium wire held at high voltage and create a measurable electron avalanche. With this setup, an average  $z$ -resolution of 80  $\mu\text{m}$  per tube is achieved which combines to approximately 35  $\mu\text{m}$  for each chamber.

#### 3.3.5.2. The cathode strip chamber

Due to the high particle rate close to the beam axis, the innermost disk closest to the interaction point, covering a pseudorapidity range of  $2.0 < |\eta| < 2.7$ , is equipped with high rate capable cathode strip chambers instead of monitored drift tubes. These are multiwired proportional chambers following the principle of the muon drift chambers. Instead of being segmented into different tubes with a single wire, multiple wires pointing radially outward within a single module are used. Both readout cathodes are segmented, one parallel and the other perpendicular to the wires, in order to provide two-dimensional positional information. Thus, each chamber has an  $\eta$ -resolution of  $40 \mu\text{m}$  and of  $40 \text{ mm}$  in the transverse plane.

#### 3.3.5.3. The resistive plate chambers

In order to trigger on muons, the high-precision tracking chambers are complemented with additional fast-readout modules. In the barrel region ( $|\eta| < 1.05$ ) three layers of resistive plate chambers are used for triggering. They are gaseous parallel electrode-plate detectors with a gas-filled gap of  $2 \text{ mm}$  and no wire in between. The signal readout is segmented into strips, which are capacitively coupled to the plates. Due to the small size they allow for a time resolution of less than  $2 \text{ ns}$  which makes them usable as triggers.

#### 3.3.5.4. The thin gap chambers

In the end-cap regions ( $1.05 < |\eta| < 2.7$ ), so-called thin gap chambers, slim multiwire proportional chambers are used for triggering on muons. Besides a fast response promoting their usages as triggers, they provide a rough  $\phi$ -position measurement perpendicular to the MDT in the endcaps with a resolution of about  $5\text{--}10 \text{ mm}$ .

#### 3.3.6. The ATLAS forward detectors

Besides the main detector parts there are three smaller detector modules covering the ATLAS forward regions. The main purpose of two systems is to measure the luminosity delivered to the ATLAS detector. At  $\pm 17 \text{ m}$  from the interaction point LUCID<sup>(16)</sup> [98], the main online relative-luminosity monitor for ATLAS, measures inelastic proton-proton scattering in the forward direction. The other detector, ALFA<sup>(17)</sup> [99], located at  $\pm 240 \text{ m}$  from the interaction point, consists of scintillating fibre trackers which lie inside so-called Roman pots designed to approach as close as  $1 \text{ mm}$  to the beam line. The third forward detector module, ZDC<sup>(18)</sup>, plays a key role in determining the centrality of heavy-ion collisions. It is located just beyond the point where the common straight-section vacuum pipe divides back into two independent beam pipes at  $\pm 140 \text{ m}$  from the interaction point. It consists of layers of alternating quartz rods and tungsten plates and measures neutral particles pseudorapidities of  $|\eta| \geq 8.2$ .

#### 3.3.7. The ATLAS trigger system and data acquisition

Triggers are experimental tools to restrict a large amount of continuously taken data to a manageable rate. They are commonly used in high energy physics experiments to pick out the events of interest discarding the rest, as it is not feasible to record the whole overwhelming rate of particle collisions induced at a hadron collider experiment.

<sup>(16)</sup> LUMinosity measurement using Cerenkov Integrating Detector

<sup>(17)</sup> Absolute Luminosity For ATLAS

<sup>(18)</sup> Zero-Degree Calorimeter

At the ATLAS detector a multi-stage trigger system [100, 101] is employed in order to identify interesting collision events. In 2012 during Run 1 (2009-2013) the LHC delivered an integrated luminosity of  $20 \text{ fb}^{-1}$  at a centre-of-mass energy of 8 TeV with a bunch spacing of 50 ns. In the Run 2 data taking period, started in 2015, the centre-of-mass energy has been increased to 13 TeV leading to a 20 % higher proton-proton cross section and a doubled hard interaction cross section for many physics processes of interest. Due to the halved bunch spacing of 25 ns, also the luminosity increased with respect to Run 1. These higher cross sections, in particular for processes dominated by the strong interaction, as well as the higher instantaneous luminosities of up to the machine's design luminosity of  $1 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and therefore higher pile-up (cf. Sec. 3.4) lead to a rise of the total interaction rate by a factor of six compared to Run 1. In order to cope with these harsher conditions, the ATLAS trigger system has been upgraded during the LHC long shutdown 1 (2013-2014) from a three-stage to a two-stage system, consisting of a hardware level-1 trigger (L1) [102, 103] and a software-based high level trigger (HLT) [104]. The latter is a comprised system of the Run 1 level-2 trigger and event filter, reducing the complexity and allowing for dynamic resource sharing between them.

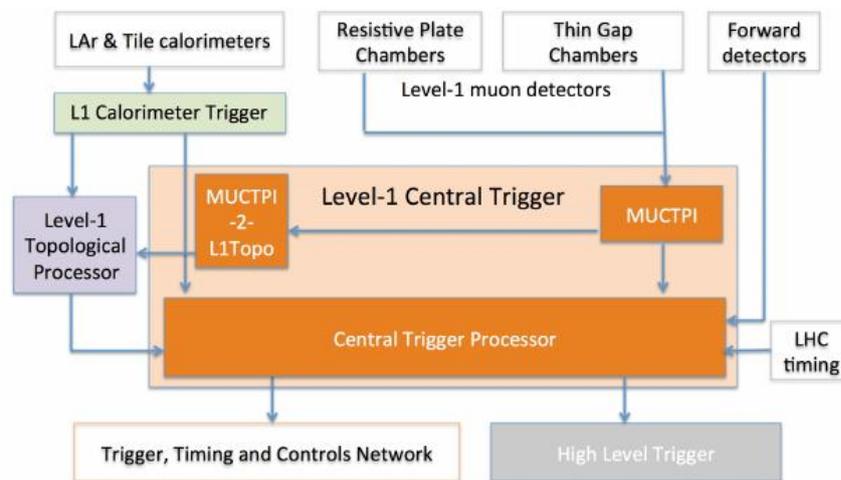


Fig. 3.8.: Schematic view of the ATLAS level-1 trigger system with its components for the various subdetector systems. The information from the muon spectrometer is first gathered by an interface module before being transmitted to the central processor where it gets combined with the information from the other subdetectors into regions of interest. These are then sent to the high-level trigger [102].

The first stage, the level-1 trigger, applies custom electronics to determine regions of interest (RoI), by using preprocessed information from the calorimeters, dedicated muon detectors and forward detectors close to the beam pipe, thereby reducing the event rate from the LHC bunch crossing rate of approximately 40 MHz to a rate of up to 100 kHz. The trigger decision is formed by the central trigger processor (CTP) [105] within time intervals of  $2.5 \mu\text{s}$ . The first part of the level-1 trigger is the level-1 calorimeter trigger (L1Calo), which uses coarse granularity information from the liquid-argon and tile calorimeters to identify energy clusters and particle types, and then provides a count of objects above a certain energy threshold as well as the missing transverse energy and dedicated  $\tau$ -lepton triggers. The L1Muon system employs resistive plate chambers and thin gap chambers to identify muons in the barrel and the end-caps, respectively (cf. Sec. 3.3.5). The information is combined in the Muon-to-CTP interface (MUCPTI) [105] and transmitted to the central trigger processor. The newly introduced level-1 topological trigger (L1Topo) allows

for the evaluation of topological selection criteria based on calorimeter and muon spectrometer information such as angles between trigger objects and the invariant mass of two or more trigger objects.

The regions of interest formed in level-1 triggers are sent to the high level trigger in which sophisticated selection algorithms are run using full granularity detector information in either the regions of interest or of the whole event, reducing the level-1 output rate of 100 kHz to a permanently recorded rate of up to 1 kHz within a processing time of about 200 ms [104].

After passing this second trigger stage the selected events are fully read out of the detector and written into permanent storage as raw data at a local computing center at CERN, referred to as Tier-0. There the raw data undergoes some reconstruction procedures, interpreting the electrical signals from the different subdetectors as physical objects like electrons or jets. In order to handle the large amount of data produced by the ATLAS detector (several Petabyte a year), the Worldwide LHC Computing Grid (WLCG) [106] has been created. It consists of more than 170 computing centres in over 40 countries provided by universities and other scientific research institutes. Out of the Tier-0 copies of raw and reconstructed data are transferred to thirteen large computing centres of sufficient capacity in terms of storage, computing power and data transmission bandwidth, the so-called Tier1. Associated to each Tier-1 centre are in total 160 Tier-2 centres, which handle simulated data generation and reconstruction. The local workstations and personal computers of scientists who access the grid infrastructure are referred to as Tier-3.

#### 3.4. Pile-up and underlying event

At high-luminosity colliders, additional collisions typically take place besides the hard proton-proton interaction and are most likely detected during the interaction of interest. They are referred to as pile-up (PU) [107]. All the additional (soft) contributions have to be corrected for, because usually only the hard interaction of a single pair of partons is of interest.

A common source of additional collisions are multiple proton-proton collisions taking place in a bunch crossing, denoted as in-time pile-up, leading to a multitude of vertices (cf. Fig. 3.9) [108]. Following a poisson distribution with mean value  $\mu$ , the average number of interactions per bunch crossing  $\langle \mu \rangle$  is used as a measure of the hard scatter activity not associated to the main primary vertex as well as of further soft contributions in a given event, and was about 14 and almost 25 for the 2015 and 2016 data taking period, respectively (cf. Fig. 3.10). To isolate the event of interest the primary vertex of the interactions is reconstructed and tracks pointing to the unwanted vertices are discarded. A similar strategy is used for the calorimeters, but the pointing is considerably less precise. All contributions not removed will deteriorate the detector resolution.

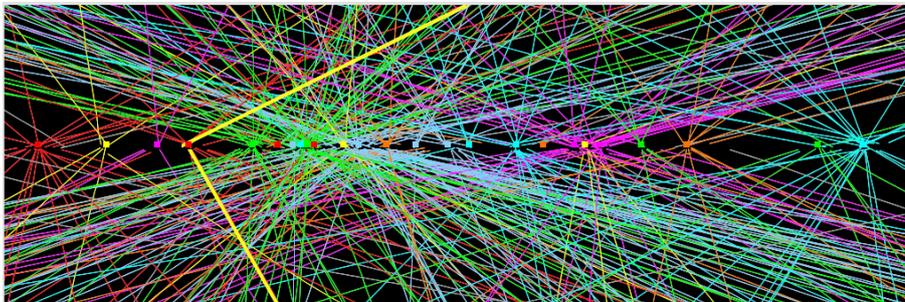


Fig. 3.9.: A candidate  $Z$  boson event in the dimuon decay mode recorded on April 15th 2012. It shows a high pile-up rate with 25 reconstructed vertices [109].

Another source of additional collisions is the so-called out-of-time pile-up, which is a detector-dependent effect caused by readout times longer than the collision rate. Therefore, collisions from bunch crossings before and after the interaction of interest may also give contributions to a calorimeter measurement and thus increase the noise. When reducing the bunch spacing from 50 ns down to 25 ns for Run 2, the contribution of out-of-time pile-up obviously increased.

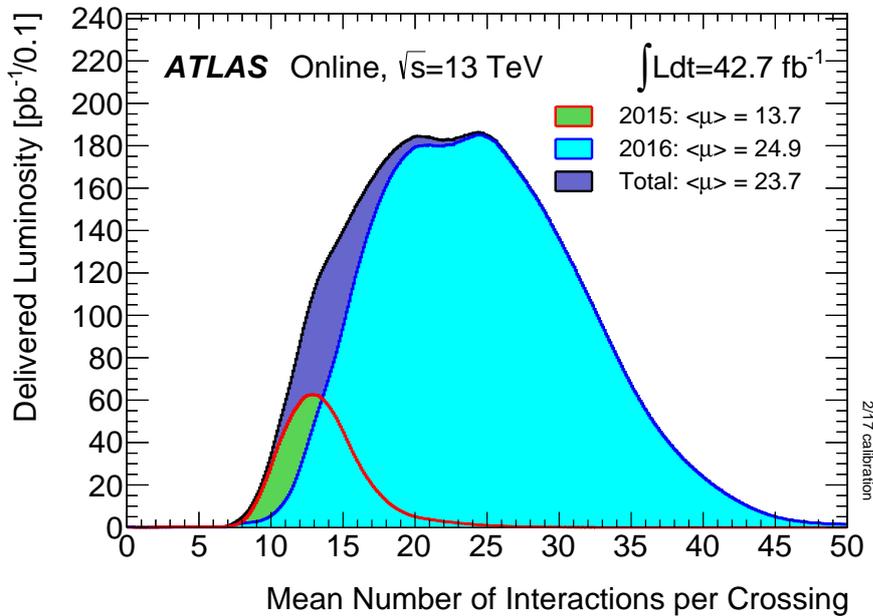


Fig. 3.10.: Luminosity-weighted distribution of the mean number of interactions per bunch crossing  $\langle\mu\rangle$  at the ATLAS detector for the 2015 and 2016 proton-proton collision data at a centre-of-mass energy of 13 TeV. The integrated luminosity shown in this picture corresponds to all data delivered to ATLAS during stable beam periods for both 2015 and 2016. The red curve shows the 2015 and the blue curve the 2016 data taking period, whereas their combination results in the dark blue shape with a black border. For all distributions the corresponding value of  $\langle\mu\rangle$  is also given [110].

Especially at hadron colliders where the colliding particles are not elementary as for lepton colliders but have a substructure, the parton-parton interaction of interest is very likely polluted by many processes contributing to the event of interest except the hard scattering itself, called the underlying event [111] (UE). It is mainly composed of soft and semi-hard multiple parton interactions within a single proton-proton collision, but gluon and photon radiation might be considered as well depending on the particular definition of the underlying event .



## 4. Data and Monte Carlo Samples

### 4.1. Monte Carlo event generation

In high energy physics the comparison of experimental data to theoretical expectations is an essential part in almost every physics analysis. Ideally, a detailed description of the final state is provided such that any experimental observable or combination of observables can be predicted. Furthermore, a proper understanding of the signal and background processes is important in order to be able to separate them. All these aspects are (almost) perfectly covered by so-called Monte Carlo (MC) generators, dedicated software tools modelling step-by-step a real-life physics process as authentically as possible and thus predicting the outcome of collider experiments. They have a large field of application, ranging from detector and trigger optimisation studies to feasibility studies of dedicated physics processes or general analysis strategies. Though, despite the great efforts people have put in to continuously improve the Monte Carlo generation and simulation, they are not absolutely perfect. Suffering from the broad range of physics they have to describe, the most common strategy is to combine components of many different simulation programs, each of them (more or less) well modelling one specific physics process or step.

In quantum mechanics calculations only provide a probability for different outcomes of a measurement but not the actual result. Thus, it is impossible to know beforehand what will happen for each event as anything that is not forbidden for some reason could take place. However, when averaging over a large sample of events, the expected probability distribution will emerge, provided that the calculations indeed describe real nature. Therefore, events from a particular particle physics process do not always look the same but are continuously distributed over its kinematically allowed phase space, following a probability density function. In Monte Carlo generators random numbers are used to make choices intended to reproduce the quantum mechanical probabilities for different outcomes at various stages of a physical process.

Generally, the term Monte Carlo method [112] refers to a wide range of numerical techniques that calculate probability-related quantities making use of random numbers. A series of random numbers, typically uniformly distributed between 0 and 1, is mapped into a series of numbers which follow the desired probability distribution function  $f(x)$ . Treating the values of  $x$  as a set of simultaneous measurements, thereof the probability for  $x$  being in a certain region of phase space can be deduced. This is effectively equivalent to the calculation of the integral of the probability distribution  $f(x)$ . Thus, Monte Carlo techniques are widely used to integrate high-complex probability density functions in a multi-dimensional phase space which in most cases cannot be done analytically anymore.

Two essential concepts of Monte Carlo generators are perturbation theory and the factorisation theorem. Generally, if no exact solution of a given problem can be calculated, perturbation theory is used to approximate the result. In order to be able to use it, the problem must provide the possibility to formulate it in terms of a power series,

$$f(x) = a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 + \mathcal{O}(x^4) \quad , \quad (4.1)$$

where higher orders of  $x$  add only increasingly small corrections to the final result if  $x < 1$ . Thus, a problem can be approximated by taking into account only leading order contributions while neglecting corrections of higher orders. This is a commonly used approach within many quantum

field theories, such as quantum chromodynamics where the series is developed in powers of the strong coupling constant  $\alpha_s$ , to approximate complex equations as a cross section calculation with power series.

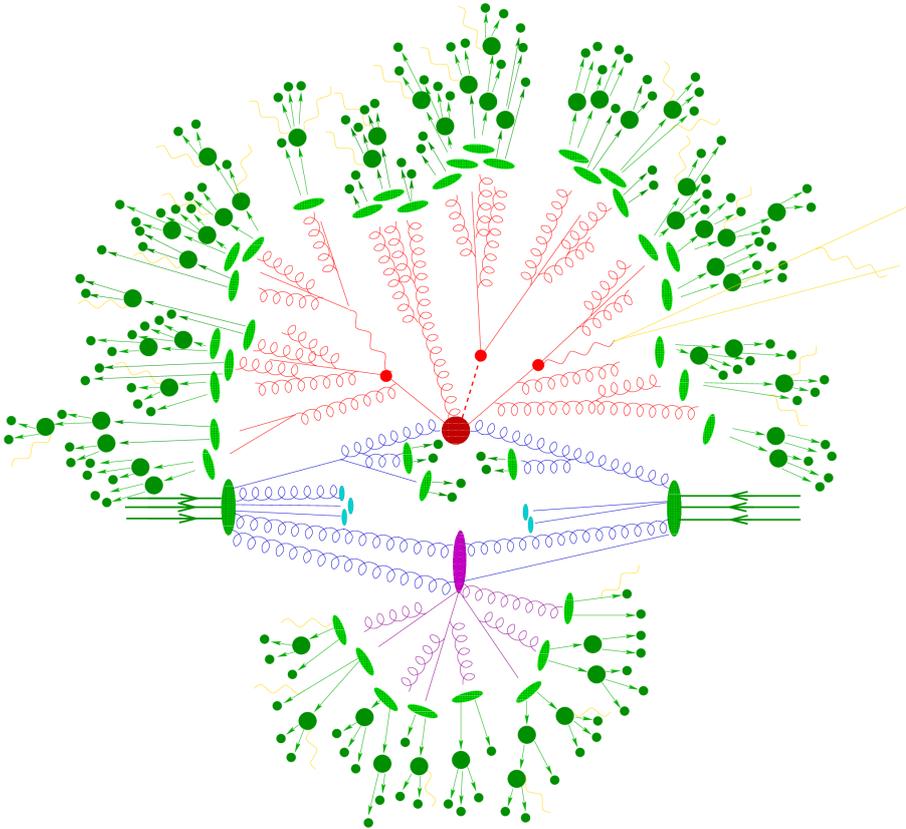


Fig. 4.1.: Schematic overview of the structure of a proton-proton collision as simulated in Monte Carlo event generation. The red blob in the centre represents the hard collision, surrounded by a tree-like structure representing bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation [113].

The aim of a Monte Carlo event generation is not only to model the so-called hard interaction but also to give an estimate as accurate as possible of all features occurring in a real-life event. Due to the factorisation theorem, the Monte Carlo method allows to divide the generation of a proton-proton collision into different stages, where in each step a set of rules is defined that can be iteratively used to construct a more and more complex state. A schematic overview of the different steps of the event simulation with Monte Carlo generators is shown in Fig. 4.1. Since hadrons are composite particles, the initial state of the hard scattering process has first of all to be determined with the help of parton distribution functions, which describe the momentum fractions of partons within the incoming protons and therefore govern the probability of being involved in the hard scattering process for the different proton constituents as explained in Sec. 2.2. The hard scattering process and its cross section is defined by the matrix element of the process of interest within the proton-proton collision. It includes both initial and final state partons as well as short-lived resonances produced in the hard process such as a top quark or a  $Z$  boson which transfers their properties, e.g. spin, to the final state partons. As already stated, the matrix element can be

expressed and calculated in terms of a series expansion.

Since the underlying theory for proton-proton collisions is mainly the strong force described by QCD<sup>(1)</sup>, the calculation of the matrix element can be expressed as an expansion in orders of the strong coupling constant  $\alpha_s$ . The leading order process is described by the tree level, i.e. the process corresponding to the lowest order Feynman diagram possible, which in case of  $t\bar{t}$ -production can be seen in Fig. 2.6. Every additional leg or loop to the diagram means the incorporation of a higher order correction as depicted in Fig. 2.7. Including higher orders, the theoretical prescription of the experimental data increases in precision but also in difficulty. Using this Feynman amplitude the differential cross section of the process of interest can be calculated, which is then used as probability distribution in the phase space of the particles associated to this process. The phase space is spanned by all degrees of freedom of the process of interest, usually given by kinematic and angular quantities of the emanating particles. With a scan through the phase space a multitude of so-called candidate events is created, each defined by the choice of the free parameters. The differential cross section of a specific candidate event, typically referred to as event weight, is directly related to the probability for this event to occur. Averaging over the event weights of all candidate events of a process gives an approximation of the total cross section.

The parton shower describes the radiation of additional gluons from initial and final state partons, referred to initial state radiation (ISR) and final state radiation (FSR), respectively. Both mechanisms add extra partons to the process, in case of initial state radiation mainly in forward direction close to incoming protons. Another source of initial or final state radiation to be taken into account is bremsstrahlung from accelerated electromagnetically charged particles involved or produced in the collision.

Due to the principle of asymptotic freedom, partons can be thought of as freely moving particles at short distances. Because of the non-abelian nature of the associated gauge theory,  $\alpha_s$  increases at low momentum scales (cf. Sec. 2.1.2), leading to a break down of perturbative QCD for energies of the order  $\Lambda_{\text{QCD}}$ . Therefore, parton showers can only be used down to a cut-off energy of  $\mathcal{O}(1 \text{ GeV})$ . If the energy scale further decreases with larger distances, confinement forces coloured partons to build colourless bound states. The formation of colourless hadrons from coloured particles cannot be described perturbatively anymore, thus phenomenological QCD-inspired models have to be used instead in order to characterise the hadronisation process. A common approach are the so-called string models [114], in which a separate colour field is assumed to be stretched between each colour and its matching anticolour. When a colour-anticolour pair is moving apart, the energy stored in the colour field in between grows until at some point it becomes energetically more favourable to produce new particles than further move away from each other, thus the field lines (“strings”) can break up by the production of new quark-antiquark pairs that screen the end-point colours. Then, these newly emerged colour charges combine with other coloured particles in their neighbourhood in order to produce colour neutral hadrons, resulting in cone-shaped bundles of uncoloured particles, so-called jets. In order to stay reasonably predictive when dealing with soft, non-perturbative QCD effects, additional parameters have to be introduced whose optimal values can be only determined by comparison to data, referred to as tuning of a Monte Carlo generator.

Complementary, cluster models [115] have been invented which employ an intermediate stage in the parton shower, namely the clustering of partons in groups of colourless objects. Here the hadronisation follows a confined colour flow where gluons are considered as a simple sum of a colour and an anticolour, and all colours distinguishable from each other. The clusters then decay according to their available phase space thus making free tuning parameters obsolete. Many of

<sup>(1)</sup> Of course, other non-strong processes can happen, for example mediated by the electroweak interaction. However, they occur less often and are typically accompanied by higher order QCD effects.

the hadrons created during the hadronisation process are unstable particles and decay further at various timescales. Some of them are sufficiently long-lived so that their decays are visible in the detector or are almost stable<sup>(2)</sup>. Hence, subsequent hadron decays have to be incorporated in the generation, whereas for sufficiently long-lived hadrons the event generator description has to be matched to a detector simulation.

So far only processes connected to the hard scattering process have been mentioned with only one parton from each proton participating. As the incoming protons are made up of a multitude of partons, further parton pairs may collide within a single proton-proton collision. These so-called multiple parton interactions will create further jets in the forward direction since the proton remnants themselves carry colour and therefore hadronise, too. Within the multiple parton interactions colour-reconnections can induce cross-talks between the hard process and the beam remnants leading to colour-space correlations which further impede a proper event simulation. Each additional parton collision might of course be associated with initial and final state radiation as well. This further activity not belonging to the hard interaction is referred to as underlying event, not to be confused with pile-up. While the first describes the evolution of the proton remnants, the latter refers to multiple proton-proton collisions within a bunch crossing (cf. Sec. 3.4). However, both phenomena share certain analogies which makes it difficult to distinguish between them.

Merging the individual components in a consistent manner is a critical task as, for example, double counting and mismatching of partons have to be avoided. In order to overcome these chinks, matrix element calculations and parton showers are typically modelled separately and reunited afterwards employing jet-parton matching algorithms. Two different approaches can be followed. The first is the so-called CKKW matching [116, 117], where multijet matrix elements are merged with the shower development by reweighting the matrix elements with Sudakov form factors and vetoing shower emissions in regions of phase space already covered by the parton level configurations. The second is called MLM matching [118] in which matrix element partons are matched to parton jets. Events are rejected if there are extra jets which fail to match to the light partons generated at the matrix elements level or if there are missing jets. Due to computational reasons, a small amount of additional partons is usually added to the matrix element calculations. Similarly, for the parton showers mostly just the lowest order matrix elements are implemented. However, combining the different steps of Monte Carlo generation leads in most of the cases to a sufficiently realistic description of a physical process benefitting from the strengths of each individual simulation program. Finally, the last step of the event generation is the detector simulation in which the detector response is modelled. The output of the detector simulation then undergoes the same reconstruction algorithms as data and therefore can be compared to experimental measurements.

## 4.2. Monte Carlo event generators

In order to model signal and background processes and to evaluate systematic uncertainties, among a multitude of Monte Carlo simulations has been chosen to suit requirements for the respective physical process. For the generation of the hard scatter interaction the generators POWHEG-BOX, MADGRAPH5\_AMC@NLO and SHERPA are made use of. POWHEG-BOX [119, 120] performs the calculations of the matrix elements (ME) at next-to-leading order (NLO) and matches them to the particle shower (PS) simulation of MC shower generator according to the POWHEG method [121, 122]. MADGRAPH5\_AMC@NLO [123] merges the features of the MADGRAPH5 [124] generator and the AMC@NLO [125] tool. The first allows for event generation of any physical model that can be written in form of a Lagrangian, whereas the latter facilitates merging NLO matrix element calculations with parton shower simulations. SHERPA [126]

---

<sup>(2)</sup> A particle is considered stable if  $c\tau \geq 1$  cm.

is a general-purpose event generator which allows for matrix element calculations of multi-parton processes up to NLO. In addition, also the simulation of the parton shower, the underlying event (UE) and the hadronisation is performed by SHERPA, as well as the merging of NLO matrix element calculations with parton shower. Another tool for simulating a large range of collider physics processes is the PYTHIA [127, 128] generator. It is used as an interface to hard scatter generators, except SHERPA, to model parton shower, underlying event and hadronisation. The latest version 8 of the PYTHIA generator is a rewrite in C++ of the previous version which are based on Fortran and still used provided they have a well developed tuning. HERWIG++ [129, 130] is additionally used for the evaluation of systematic uncertainties as an alternative generator for the parton shower simulation. The properties of heavy flavour decays, particularly important to this analysis, have been simulated with EVTGEN [131], except for processes modelled using SHERPA. In all samples the top quark mass and the Higgs boson mass is set to 172.5 GeV and 125 GeV, respectively. In case of the explicit generation of  $b$ -quarks their mass is constrained to 4.75 GeV.

### 4.3. Monte Carlo simulation samples

The associated production of a top quark pair with one vector boson ( $t\bar{t}Z$ ,  $t\bar{t}W$ ) is generated at NLO with MADGRAPH5\_AMC@NLO interfaced to PYTHIA 8 [128]. For the matrix element, the parton distribution function (PDF) is set to NNPDF3.0NLO [132]. The A14 tune [133] is applied together with the NNPDF2.3LO PDF set [134, 135]. The production of  $t\bar{t}$  in association with two  $W$  bosons ( $t\bar{t}WW$ ) is simulated the same way but only with LO matrix element calculations.

The  $t$ -channel production of a single top quark in association with a  $Z$  boson ( $tZ$ ) is generated and normalised at leading order (LO) using MADGRAPH5\_AMC@NLO and PYTHIA 6 [127] with the CTEQ6L1 PDF set [136] and the PERUGIA2012 tune [137, 138].

For the  $Wt$ -channel production of a single top quark in association with a  $Z$  boson ( $tWZ$ ), instead of PYTHIA 6 its successor PYTHIA 8 is used, together with the NNPDF2.3LO PDF set and the A14 tune. The matrix element is calculated at NLO. In order to remove the overlap and interferences with  $t\bar{t}Z$  and with  $t\bar{t}$  followed by a three-body decay ( $t \rightarrow WZb$ ), diagram removal (DR) [139] is applied. The  $tWZ$  sample are normalised to their NLO cross section.

Both  $t\bar{t}$  and single top quarks are simulated at NLO with POWHEG-BOX 2.0 and the CT10 PDF set [140]. The parton shower, hadronisation and underlying event are simulated using PYTHIA 6 with the CTEQ6L1 PDF set and the corresponding PERUGIA2012 tune. The  $t\bar{t}$  samples are normalised to their next-to-next-to-leading order (NNLO) cross section including the resummation of soft gluon emission at next-to-next-to-logarithmic (NNLL) accuracy using TOP++2. [141]. For the single top quark  $Wt$ -channel also diagram removal is employed.

As the joint production of  $t\bar{t}$  in association with a pair of  $b$ -quarks ( $t\bar{t}b\bar{b}$ ) [142] is not well simulated by the nominal  $t\bar{t}$  Monte Carlo generator, an extra simulated sample is taken into account for this specific process using the SHERPA 2.1 generator for the matrix element calculation together with the CT10 and the LHA[CT10nlo\_nf4] [143, 144] PDF sets.

The associated production of a top quark pair and a Higgs boson ( $t\bar{t}H$ ) is generated at NLO matrix element calculations with MADGRAPH5\_AMC@NLO using the NNPDF3.0NLO PDF set. The showering is done with PYTHIA 8. As for  $t\bar{t}Z/W$ , the A14 tune and the NNPDF2.3LO PDF set are used.

In addition the decay of the Higgs boson into a pair of  $b$ -quarks ( $H \rightarrow b\bar{b}$ ) is simulated separately but with the same generator settings as  $t\bar{t}H$ . The Higgs boson production via vector boson fusion (VBF) and gluon fusion is produced with the POWHEG-BOX 2.0 generator using the CT10 PDF set. PYTHIA 8 is used together with the CTEQ6L1 PDF set and the AZNLO [145] tune for the parton shower, fragmentation and underlying event.

The production of Higgs boson in association with vector bosons ( $VH$ ) is generated at LO using PYTHIA 8 with the CTEQ6L1 PDF set.

Diboson processes with only leptonic final states ( $llll$ ,  $lll\nu$ ,  $ll\nu\nu$ ), as well as diboson processes with hadronic contributions ( $qqll$ ,  $qq\nu\nu$ ,  $qql\nu$ ,  $ggllll$ ,  $ggll\nu\nu$ <sup>(3)</sup>) are generated with the SHERPA 2.2.1 generator using the NNPDF3.0NLO PDF set and normalised to the NNLO cross section. The matrix elements are calculated for up to two additional partons at LO as well as at NLO and merged afterwards with the SHERPA parton shower using the ME+PS@NLO prescription [146].

The production of three massive vector bosons with subsequent leptonic decays is modelled at LO with the SHERPA 2.1 generator using the CT10 PDF set. By means of the the COMIX [147] and OPENLOOP [148] matrix element generators, at LO either two or three additional partons are included, whereas at NLO only one additional parton is considered in the first case and none in the second case.

Events containing  $Z$  or  $W$  bosons associated with jets ( $Z/W$ +jets) are simulated with the SHERPA 2.2.1 generator using the NNPDF3.0NLO PDF set. The samples are normalised to the NNLO cross section. For both processes the matrix elements are calculated up to two additional partons at LO and up to three partons at NLO.

The production of four top quarks is generated with MADGRAPH5\_AMC@NLO at LO, using the A14 tune together with the NNPDF2.3LO PDF set.

Theoretical systematics have been estimated for  $t\bar{t}$  regarding the hard scattering process, the hadronisation behaviour and additional radiation. The latter has been investigated with two Monte Carlo samples both generated with POWHEG-BOX 2.0, where a different shower radiation behaviour has been simulated using the PERUGIA2012 radHi and radLo tune, respectively.

As for the nominal  $t\bar{t}$  sample, the hadronisation and underlying event has been modelled using PYTHIA 6 with the CTEQ6L1 PDF set and the corresponding PERUGIA2012 tune. In contrast to the nominal  $t\bar{t}$  samples, the factorisation has been modified as well as the renormalisation scale ( $\times 2$  and  $\times 0.5$ ) and the NLO radiation scheme. Though, despite of these changes the same combination of Monte Carlo generators had been used. For the hard interaction of  $t\bar{t}$  two Monte Carlo samples, one using POWHEG-BOX and the other MADGRAPH5\_AMC@NLO for the matrix element calculation, have been chosen. In both cases the CT10 PDF set is taken for the matrix element and CTEQ6L1 for the simulation of the showering, which is done with HERWIG++, using the UEEE5 tune [149] together with the CTEQ6L1 PDF set.

The impact of different showering algorithms have been studied with another two Monte Carlo samples. The hard interaction is modelled with POWHEG-BOX each time, whereas the hadronisation is modelled with either PYTHIA 6 or HERWIG++. More precisely, two different hadronisation models are compared, namely the Lund string model [114, 150] implemented in PYTHIA 6 and the cluster fragmentation model [115] used in HERWIG++. As PDF set the HERWIG++ sample uses the CT10 PDF set for the matrix element calculation and CTEQ6L1 for the shower modelling. Therefore, not only the hadronisation model is different which makes it much more difficult to directly compare the hadronisation models. Additionally, the UEEE5 tune together with the CTEQ6L1 PDF set is applied. On the contrary, the other sample goes with the CTEQ6L1 PDF set for the hard scattering process and the PERUGIA2012 tune. In fact, the latter sample is the nominal  $t\bar{t}$  sample used in this analysis.

Another feature which was taken into account is the difference between diagram removal and diagram subtraction [151] (DS) procedure applied in the  $Wt$ -channel of single top quark events. These samples share all other generator characteristics as those with diagram removal.

A summary of the basic parameters of the Monte Carlo datasets, nominal as well as systematic samples, is shown in Tab. 4.1. To account for additional proton-proton interactions from the

---

<sup>(3)</sup>  $qq$ ,  $ll$  and  $\nu\nu$  can either stem from a  $Z$  or a  $W$  boson

same or close-by bunch crossing, a set of minimum-bias pile-up events, simulated with PYTHIA 8 and the MSTW2008LO PDF set [61], is superimposed to the hard scattering processes. In order to match the pile-up profile measured in data, the distribution of simulated pile-up events is reweighted.

#### 4.4. Preprocessing and normalisation of simulated events

After the Monte Carlo generation step all events are passed through a full simulation of their interaction with all components of the ATLAS detector [153] using GEANT4 [154]. Then, the simulated signatures are evaluated with the ATLAS digitisation software in order to emulate the respective detector response to impinging stable particles combined with an estimate of the detector noise [153]. Afterwards, the same reconstruction and pattern recognition procedure is applied on the Monte Carlo datasets as for real data in order to make them comparable [155, 156].

The simulated events are corrected such that the object identification, reconstruction and trigger efficiencies, energy scales and energy resolutions match those determined from data control samples. Since a full simulation of the detector response is very intensive in computing power and time, alternatively fast simulation approaches can be used if a large amount of simulated data statistics is required [153]. Therefore, the commonly used ATLFAST-II simulation directly simulates the input of the reconstruction software, using both a simplified description of the detector geometry and a simplified shower parametrisation for the calorimeters [157]. However, to stay on a reasonably high level of accuracy, the option to perform a full simulation with GEANT4 for any subdetector system is still provided, albeit with a small performance degradation in terms of physics description. Within this analysis only fully simulated Monte Carlo datasets have been used.

Events from Monte Carlo simulation which pass all selection criteria have to be normalised to the integrated luminosity  $(\int L dt)_{\text{data}}$  of the recorded data considered for an analysis. The normalisation can be described by a scaling factor  $f_{\text{lumi}}$ , defined as

$$f_{\text{lumi}} = k \cdot \varepsilon_{\text{filter}} \cdot \frac{(\int L dt)_{\text{data}} \cdot \sigma_{\text{MC}}}{N_{\text{MC}}}, \quad (4.2)$$

where the k-factor  $k$  accounts for higher order corrections on the process cross section or other correction factors evaluated after the event simulation and  $\sigma_{\text{MC}}$  represents the cross section with what the respective process has been produced. The filter efficiency  $\varepsilon_{\text{filter}}$  is a scaling factor in order to properly reflect the selection rate of a events from a filtered sample. Filters are used already on generator level to enhance the statistics of Monte Carlo datasets by selecting specific decay topologies at the level of event simulation, either by the selection of explicit decay channels of a particular physics process or by the limitation of the generated phase space. The latter is a common procedure within many analyses related to physics beyond the Standard Model, e.g. Supersymmetry, as such signals are predicted to lie within particular regions of phase space. This reduces the amount of resources spent on the subsequent computing intensive detector simulation and particle reconstruction. The number of Monte Carlo events  $N_{\text{MC}}$  is calculated during the actual analysis when no event selection has been applied yet, including the generator specific per event Monte Carlo weight.

#### 4.5. Data recorded by ATLAS

For this analysis solely Monte Carlo generated events are used, scaled to the luminosity of the data from proton-proton beam collisions with a bunch spacing of 25 ns at a centre-of-mass energy of

13 TeV, recorded with the ATLAS detector at the LHC in 2015 and 2016, is used. During the data recording the instantaneous luminosity had been steadily increased peaking at  $5 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  in 2015 and  $13.8 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  for 2016 data. As a consequence, the mean number of interactions per bunch crossing ascends for 2015 with respect to 2016 (cf. Fig. 3.10). Only data passing certain data quality requirements<sup>(4)</sup> is used corresponding to an integrated luminosity of  $3.2 \text{ fb}^{-1}$  and  $32.9 \text{ fb}^{-1}$  for the 2015 and 2016 data recording periods, respectively. These values sum up to a combined integrated luminosity of  $36.1 \text{ fb}^{-1}$  for the full dataset from 2015 and 2016 suited for this analysis [110].

---

<sup>(4)</sup> The so-called good run lists (GRL) used can be found in [158] and [159] for the 2015 and 2016 data taking period, respectively.

Sample	Generator	ME PDF	Shower	Normalisation	$\sigma$ [pb]
$t\bar{t}Z$	AMC@NLO	NNPDF3.0 @NLO	PYTHIA 8	NLO	0.8783
$t\bar{t}W$	AMC@NLO	NNPDF3.0 @NLO	PYTHIA 8	NLO	0.6008
$t\bar{t}WW$	MADGRAPH	NNPDF3.0 @LO	PYTHIA 8	NLO	0.0099
$tZ$	MADGRAPH	CTEQ6L1 @LO	PYTHIA 6	LO	0.2401
$tWZ$	AMC@NLO	NNPDF2.3 @NLO	PYTHIA 8	NLO	0.0156
$t\bar{t}$	POWHEG	CT10 @NLO	PYTHIA 6	NNLO+NNLL	831.1145
$t\bar{t}b\bar{b}$	SHERPA 2.1	CT10 @NLO	SHERPA	NLO	0.3151
$t (Wt)$ (DR)	POWHEG	CT10 @NLO	PYTHIA 6	aNNLO	71.6699
$t$ ( $s$ -channel)	POWHEG	CT10 @NLO	PYTHIA 6	aNNLO	3.3499
$t$ ( $t$ -channel)	POWHEG	CT10 @NLO	PYTHIA 6	aNNLO	70.4280
$4t$	MADGRAPH	NNPDF2.3 @LO	PYTHIA 8	NLO	0.0092
$t\bar{t}H$	AMC@NLO	NNPDF3.0 @NLO	PYTHIA 8	NLO	0.5065
$VH$	EVTGEN	CTEQ6L1 @LO	PYTHIA 8	NNLO	2.2496
$VBF H$	POWHEG	CT10 @NLO	PYTHIA 8	NNLO	3.7480
$gg \rightarrow H$	POWHEG	CT10 @NLO	PYTHIA 8	NNLO	43.9200
$W$ +jets	SHERPA 2.2.1	NNPDF3.0 @NLO	SHERPA	NNLO	59 676.4694
$Z$ +jets	SHERPA 2.2.1	NNPDF3.0 @NLO	SHERPA	NNLO	17 506.0349
Diboson	SHERPA 2.2.1	NNPDF3.0 @NLO	SHERPA	NNLO	187.7751
Triboson	SHERPA 2.1	CT10 @LO	SHERPA	LO	0.0151
$t\bar{t}$ radHi	POWHEG	CT10 @NLO	PYTHIA 6	NNLO+NNLL	831.7534
$t\bar{t}$ radLo	POWHEG	CT10 @NLO	PYTHIA 6	NNLO+NNLL	831.7510
$t\bar{t}$	AMC@NLO	CT10 @NLO	HERWIG++	NLO	831.7529
$t\bar{t}$	POWHEG	CT10 @NLO	HERWIG++	NNLO+NNLL	831.0250
$t (Wt)$ (DS)	POWHEG	CT10 @NLO	PYTHIA 6	aNNLO	68.2370

Tab. 4.1.: A summary of the Monte Carlo datasets used in this analysis together with their basic generator parameters. The generator type for the hard interaction and the PDF set are listed for each process, as well as the showering algorithm and the order to which the samples are normalised. The cross sections  $\sigma$  depicted in the last column are given in picobarns. If no further specification is made, for all samples except those produced with SHERPA on top of the matrix element calculation EVTGEN has been used to model heavy flavour decays. For the samples generated with SHERPA a dedicated parton shower tuning developed by the SHERPA authors is applied. The Monte Carlo samples for single top production are normalised to approximate NNLO (aNNLO) [152].



## 5. Physics objects

The Monte Carlo simulations as well as the raw data kept in permanent storage are subjected to various reconstruction procedures translating the signals from either the detector simulation or the detector itself into physical objects. This analysis considers the  $t\bar{t}Z$  1-lepton final state which includes electrons, muons, ( $b$ -tagged) jets and the presence of missing transverse energy. The physics object definitions follow the official ATLAS recommendations whereof the main reconstruction and identification criteria applied for each object is given below. Since, for simplicity,  $\tau$ -leptons are not dedicatedly considered but only picked by chance through their decay products, they will not be separately mentioned. In Fig. 5.1 the signatures of different particles which very likely show up in high-energy collisions of hadrons are schematically sketched for a sector of the ATLAS detector.

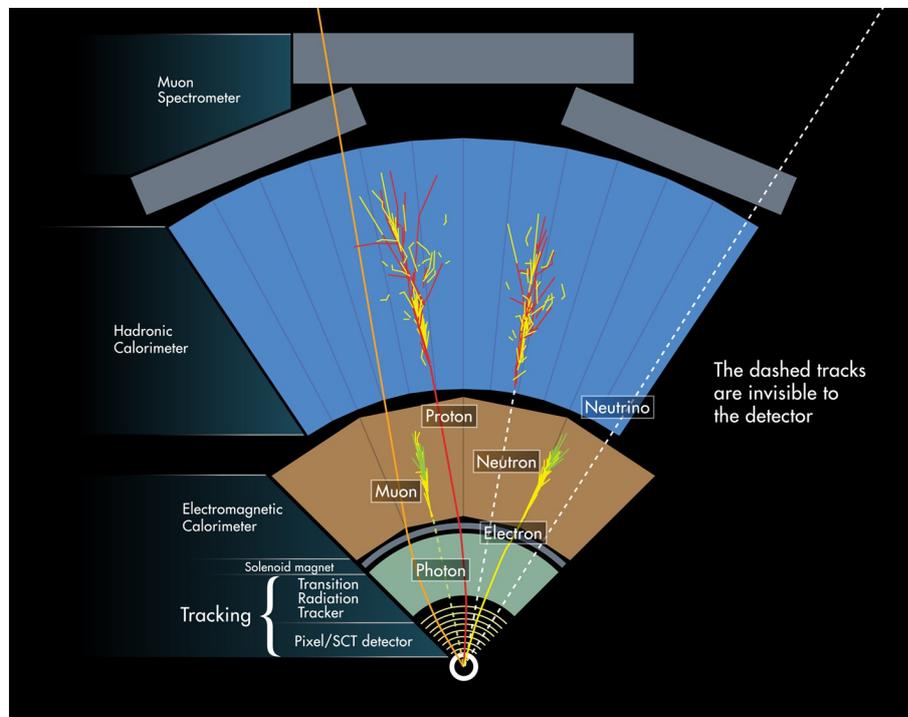


Fig. 5.1.: Schematic view of the signatures of different particles for a sector of the ATLAS detector [160].

### 5.1. Primary vertex

As mentioned in Sec. 3.3.3, trajectories of charged particles are reconstructed as tracks in the inner detector using spatial measurements of the different subdetector systems and the magnetic field of the central solenoid [108, 161, 162]. In the pixel detector and the first layer of the silicon strip tracker seed hits are identified, which point to the interaction region. The seeds are then extended outwards throughout the layers of the semiconductor tracker and the transition radiation tracker to form an entirely reconstructed track by iteratively fitting the hits and applying a set of quality

criteria. In addition, tracks are reconstructed by extrapolation of unused track segments in the transition radiation tracker inwards to the semiconductor tracker and the pixel detector. Similarly to this, tracks of charged particles not identified in both the inner detector and the calorimeter system are reconstructed in the muon spectrometer [162, 163].

Vertices of particle decays which are compatible with originating from the hard scatter interaction, are referred to as primary vertices. They are reconstructed with iterative vertex fitting and finding algorithms [108, 164]. First, a vertex seed candidate is selected according to the reconstructed tracks originating from the nominal interaction point. Then, the vertex position is fitted with the position of the seed and its associated tracks. Tracks which are incompatible with this vertex are used as seeds for new vertices until no further tracks are present in the event.

The full reconstruction of all physical objects used in this analysis requires the identification of the hard scatter interactions they originate from, i.e. the primary vertices. Events must have at least one primary vertex which is associated to two or more reconstructed tracks in the inner detector. Tracks not considered for the vertex reconstruction have to fulfil further requirements as a transverse momentum of at least 400 MeV and  $|\eta| < 2.5$  plus further criteria on the number of hits in the inner detector. If more than one primary vertex has been reconstructed in a given event, the one with the highest sum of squared transverse momenta  $\sum p_t^2$  of the associated tracks is chosen as nominal interaction point for the subsequent reconstruction of physical objects in the event.

## 5.2. Electrons

Electron candidates [165] are reconstructed from isolated energy deposits (clusters) in the electromagnetic calorimeter which can be associated to a track of a charged particle in the inner detector. The associated calorimeter cluster energy is calibrated using Monte Carlo simulation based on multivariate analysis techniques. Furthermore, uniformity corrections are applied to data in order to equalise the response of the longitudinal layers of the electromagnetic calorimeter between data and Monte Carlo simulations, whose residual disagreements are corrected with in-situ energy scale measurements in  $Z \rightarrow e^+e^-$  events. The four-momenta of the electrons are reconstructed with the energy of the calibrated clusters, whereas the angular information is taken from the track in the inner detector which best matches the selected cluster [166, 167].

Energy clusters in the electromagnetic calorimeter are reconstructed with the “sliding window” clustering algorithm [168] which first searches for a local maximum of deposited energy above a threshold of  $E_T > 2$  GeV by systematically scanning the electromagnetic calorimeter with a sliding window of size  $3 \times 5$  in units of calorimeter cells in the  $\eta$ - $\phi$ -plane. The cell size corresponds to the granularity of the electromagnetic calorimeter middle layer, which is  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ . The actual clusters are then formed out of these seed clusters where double counting of calorimeter cells has to be avoided. Afterwards, the reconstructed cluster is matched to a track wherefore a minimum number of hits in the inner detector layers is required [168]. In order to discriminate from pions, a pattern recognition algorithm has been applied before which models the energy loss of pions to either confirm or to discard the pion hypothesis [166].

Electron candidates are further required to have their associated tracks to be compatible with the primary vertex in the event in order to ensure their relation to the hard scattering process and to reduce contributions from background processes as photon conversion to an electron-positron pair or secondary vertex decays. Several conditions are applied to the longitudinal ( $z_0$ ) and transverse ( $d_0$ ) impact parameter<sup>(1)</sup>, reflecting the closest separation of the track to the primary vertex in

---

<sup>(1)</sup> The impact parameter is defined as the perpendicular distance between the primary vertex and a track associated to a secondary vertex.

the respective direction. Electrons are required to satisfy  $|z_0 \sin \theta| < 0.5$  mm and  $|d_0|/\sigma_{d_0} < 5$ , where  $\sigma_{d_0}$  represents the estimated uncertainty on  $d_0$  [166].

In order to further discriminate electron candidates from backgrounds from non-prompt leptons<sup>(2)</sup> or misidentified leptons from multijet production, they have to pass additional identification criteria. For this purpose, a likelihood-based method utilising multivariate analysis techniques is used, which simultaneously exploits several properties of the reconstructed object at once, related to the shape of the electromagnetic showers, tracking and the track-to-cluster matching. Out of these input variables signal and background probability densities  $P_{S/B}$  are built, from which likelihood functions  $\mathcal{L}_{S/B}$  and a discriminant  $d_{\mathcal{L}}$  are constructed:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B} \quad (5.1a)$$

$$\mathcal{L}_{S/B} = \prod_{i=1}^n P_{S/B,i}(x_i) \quad (5.1b)$$

The  $x_i$  denote the values of the input variables which serve as arguments for the probability density functions  $P_{S/B,i}$  of the  $n$  input variables to the likelihood discriminants. Thus, three working points with increasing requirements on the likelihood-based identification criteria can be defined, namely LooseLH, MediumLH and TightLH, which correspond to a selection efficiency of 95 %, 90 % and 80 % for an electron with  $E_T \approx 40$  GeV, respectively. As the shower shapes of electrons vary with pseudorapidity and energy, the identification working points are optimised in different  $\eta$  and  $E_T$ -bins, thus leading to slightly differing values [165, 167].

The contamination with leptons originating from hadron decays is further suppressed by requiring reconstructed leptons to be isolated from other particles in their close proximity. This is achieved by satisfying requirements on dedicated isolation variables which are based on track momenta and calorimeter energy deposits in defined distances of  $\Delta R$  around the lepton, excluding contributions from the lepton itself and correcting for effects from pile-up and underlying event [163, 166].

For this analysis only electron candidates with a transverse energy of  $E_T > 25$  GeV or  $E_T > 27$  GeV are selected, depending on the object definition of the 2015 and 2016 data taking period, respectively. Furthermore, the electron candidates have to lie within  $|\eta| < 2.47$  where the transition region between the barrel and the end-cap electromagnetic calorimeters (“crack region”) is excluded. To further reduce backgrounds from jet misidentification, electron candidates have to fulfil the TightLH requirements. For the isolation the criteria defined by the Gradient isolation selection have to be met, which employs gradually increasing requirements depending on the transverse momentum of the lepton [163].

### 5.3. Muons

Muon candidates [163] are reconstructed from the information of the inner detector and the muon spectrometer, complemented with the measurements in the calorimeter system although muons are minimally ionising particles. The track reconstruction in the inner detector is performed as it is done for electrons. In the muon spectrometer, hit patterns are searched for in each muon chamber separately. Within the monitored drift tubes segments are formed fitting a straight line to the detected hits which yields particle trajectories in the bending plane. Complementary, the trajectory parts in the orthogonal plane are determined in the resistive plate chambers and the thin gap chambers, respectively. In the cathode strip chambers a separate combinatoric algorithm is

<sup>(2)</sup> Leptons not associated to the primary vertex in a given event.

applied to build segments. After all found segments are fit together resulting in a muon track, an overlap removal procedure is applied preventing the usage of one segment for more than one track candidate. According to the separate reconstruction approach that has been applied in the subdetector systems muon candidates are categorised into four different types. If a track can be fully reconstructed in both the inner detector and the muon spectrometer, so-called combined muons are formed by globally refitting a track to hits in both subdetectors. Most of the muons are first reconstructed in the muon spectrometer and then extrapolated to a track in the inner detector, while the inverted case only amounts to a small fraction of the combined muons. If a muon candidate leaves a signature in the muon spectrometer which is insufficient for a full reconstruction therein, it can be reconstructed as a segment-tagged muon by extrapolating a track in the inner detector to the muon spectrometer, where at least one local segment in either the monitored drift tubes or the cathode strip chambers must be present. Another attempt is to match tracks in the inner detector to energy deposits in the calorimeter system which are in agreement with those of minimally ionising particles. Although such calorimeter-tagged muons have less purity, they emerge to be useful as they compensate for decreased acceptance in regions not covered by the muon spectrometer. Eventually, the fourth type of muon candidates is the so-called extrapolated muon, which are only based on a fully reconstructed track in the muon spectrometer the matching to a track in the inner detector has less priority. However, in this case muon candidates are required to pass at least two (three) layers of monitored drift tubes in the barrel (end-cap) region of the detector. In order to prevent an overlap of these different muon types, a type hierarchy is established. If a reconstructed track fulfils the requirements of two categories, priority is first given to combined muons, followed by segment-tagged and calorimeter-tagged muons. If an overlap with extrapolated muons occurs, the reconstruction providing the better fit quality and larger number of hits is selected [163].

The muon momentum is calibrated to correct for discrepancies between data and Monte Carlo simulation. Therefore, calibration constants related to the momentum scale and resolution in both the inner detector and the muon spectrometer are obtained from data using a binned maximum-likelihood fit with templates derived from simulation, comparing dimuon invariant mass distributions in  $Z \rightarrow \mu^+\mu^-$  and  $J/\Psi \rightarrow \mu^+\mu^-$  events from data and Monte Carlo simulations. The corrected transverse momentum of the muons is then evaluated by combining the weighted average of the corrected momenta from both subdetectors and thus used to derive momentum corrections which are dependent on both the transverse momentum and the pseudorapidity of the muons [163].

The main background for muon identification stems from decaying pions and kaons. This background is suppressed by imposing quality requirements on the muon candidates in order to target only muon candidates from hard scatter interactions. Since a hadron decay would yield a kink in the track topology and the measured momenta in the inner detector and the muon spectrometer would differ significantly, the most important discriminants for combined muons are the fit quality and the difference of transverse momenta measured in the inner detector and the muon spectrometer. Therefore, three variables are used as measures for these properties. The first variable is the ratio of the charge and the momentum measured in both the inner detector and the muon spectrometer, respectively, divided by the uncertainties on these measurements. As second variable the absolute value of the difference between the momenta measured in the inner detector and the muon spectrometer is taken. The third distinctive feature is the normalised  $\chi^2$  of the combined track fitting procedure. According to cuts on these requirements, muon candidates can be defined as Loose, Medium, Tight and high- $p_T$  with increasing stringency on the criteria, which correspond to selection efficiencies of 97 % (98 %), 95 % (96 %), 78 % (80 %) and 90 % (92 %) for muons with  $p_T < 20$  GeV ( $20 < p_T < 100$  GeV), respectively. The fourth working point, the so-called high- $p_T$  regime has been defined in order to provide a selection that allows for a good resolution for the momentum measurement at high transverse momenta ( $p_T > 100$  GeV), at cost

of a smaller selection efficiency [163].

For this analysis only combined muons within a pseudorapidity of  $|\eta| < 2.5$  are taken into account. As for electrons, the requirement on the transverse momentum is dependent on the applied triggers, namely  $p_T > 25$  GeV and  $p_T > 27$  GeV for the 2015 and 2016 data taking period, respectively. To reduce the contamination with muons originating from hadron decays, all selected muons have to fulfil Medium quality criteria. The isolation criteria to be met are those of the Gradient isolation selection and are thus the same as for electrons. To ensure their association to the primary vertex of the event, muon candidates are required to satisfy  $|d_0|/\sigma_{d_0} < 3$  and  $|z_0 \sin \theta| < 0.5$  mm. In addition, cosmic muons<sup>(3)</sup> are vetoed, i.e. muon candidates with  $|z_0| > 1$  mm and  $|d_0| > 0.22$  mm are discarded [163].

## 5.4. Jets

Colour-charged particles as quarks and gluons cannot be detected individually as confinement forces them to form colourless bound states leading to cone-shaped showers of uncoloured particles, the jets, via hadronisation. When interacting with the detector material, they manifest themselves as distinct decay cascades in the hadronic calorimeters.

For the jet reconstruction several methods have been developed which can be separated into two classes. The first and more intuitive approach is an iterative procedure starting with a cone of a defined radius  $R$  around an energy deposit above a given threshold [169–171]. After calculating the sum of the momenta of all cone constituents, the resulting direction is taken as the centre of the next cone in the iteration, until all stable cones around the initial energy deposit are found and combined to one single jet. However, these techniques are found to be infrared and collinear unsafe, i.e. jets containing particles with a lower momentum or particles close to each other and pointing in the same direction are not correctly identified, thus spoiling the accuracy of the so-called cone algorithms. In order to overcome these chinks, the second class of jet reconstruction algorithms has been developed called sequential recombination algorithms [172], which reconstruct jets by iteratively recombining the jet constituents until the remaining objects are too far away from each other. Thus, the defining parameter is the distance between the particles, which is specified in the plane spanned by the azimuthal angle  $\phi$  and the rapidity  $y$ :

$$d_{ij}^2 = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \left[ \frac{(\Delta y_{ij}^2 + \Delta \phi_{ij}^2)}{R^2} \right] \quad (5.2)$$

The  $k_{t,i/j}$  denote the transverse momenta of the jet constituents  $i$  and  $j$ , respectively, while the distance parameter  $R$  has to be given to the algorithm and defines the width of the reconstructed jets. The parameter  $p$  can be freely chosen resulting in different recombination sequences for each value of  $p$ . The choice of  $p = 1$  corresponds to the  $k_t$  algorithm [173, 174] which is hierarchical in relative  $k_t$ , and  $p = 0$  is used for the Cambridge-Aachen algorithm [175, 176] which is hierarchical in the angle what makes it useful to distinguish jet substructures. However, the most common choice is  $p = -1$  corresponding to the so-called anti- $k_t$  algorithm. It has the advantage of being both collinear and infrared (ICR) safe [177] and also very fast in computing time. In contrast to Cambridge-Aachen and the  $k_t$  algorithm, it also provides jets with a circular profile since soft particles are recombined with hard particles before being recombined with other soft particles [178].

<sup>(3)</sup> Muons generated by cosmic rays hitting the atmosphere and whose tracks generally do not cross the nominal interaction point.

Due to these outstanding benefits in ATLAS the anti- $k_t$  algorithm is used by default with a distance parameter of  $R = 0.4$ . As input to the jet reconstruction algorithm serve topological energy clusters [179] which are reconstructed by collecting calorimeter cells in three dimensions around an initially selected seed cell which is required to have a significant signal-to-noise ratio [162, 180]. Cells directly adjacent to the seed are collected in the cluster, and two additional layers of neighbouring cells with a slightly reduced signal-to-noise ratio threshold are added. Afterwards, a splitting algorithm searches for local maxima and splits the cluster if more than one local maximum has been found. Jets which are within the acceptance of the inner detector are associated to tracks according to the ghost association procedure [180, 181]. These tracks are assigned with infinitesimal momentum and included in the clustering sequence of the anti- $k_t$  algorithm [178], allowing for a unique association of tracks to the respective jets without altering its reconstructed momentum.

In order to account for the difference in detector performance for electromagnetically and hadronically particles, the measurements of the energy deposits are calibrated for each cell to the energy scale of the electromagnetic calorimeter, called local cell reweighting [179]. Furthermore, jet energy scale (JES) calibration [180, 182] is applied. This procedure includes the correction of the jet four-momentum to make it point to the identified primary vertex which does not affect the energy of the jet. In addition, corrections which account for pile-up effects have to be done. To correct for the detector response in different  $\eta$ -regions and within different calorimeter technologies, a  $p_T$ - and  $\eta$ -dependent calibration, derived from Monte Carlo simulations, are applied. To reduce the dependence of the jet energy measurement on the directional structure of the jets, further corrections have to be applied, as well as corrections for jets which are not fully contained in the calorimeter. The global sequential correction (GSC) [183] is designed to reduce the jet response dependence on the flavour of the parton which initiated the jet, by using global jet properties as the portion of measured jet energies in particular parts of the calorimeter or track multiplicities. Finally, in situ corrections account for differences in the measurement of the jet transverse momentum between data and Monte Carlo simulation. Such differences can originate from limitations in the simulation of the underlying event, physics of jet formation, pile-up activity, the physics of electromagnetic and hadronic interactions in the detector, and the description of the detector material. The corrections are calculated using  $\gamma/Z$ +jets and multijet events in data and Monte Carlo simulations [184–186].

A common procedure is the categorisation of jets according to their flavour, i.e. the flavour of the parton the jets originate from. Hence, jet flavour distinguishes between  $b$ -quarks,  $c$ -quarks and light quarks, referring to  $u$ ,  $d$ ,  $s$ -quarks, and gluons.

The contamination with jets arising from pile-up is mitigated using the jet vertex tagger (JVT) discriminant [187]. The jet vertex tagger algorithm performs a two-dimensional likelihood evaluation employing variables targeting the association of non-pile-up jet tracks to the primary vertex, based on information from the calorimeter system and the inner detector. For jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$  the resulting JVT output is required to be larger than 0.59 [187]. In order to correct for discrepancies between data and Monte Carlo simulation, momentum-dependent efficiency corrections on the JVT selection, evaluated in  $Z$ +jets events measuring the number of jets passing the JVT criterion, are applied. Further suppression of jets stemming from background processes such as showers from cosmic rays, calorimeter noise or beam-induced secondary cascades, additional criteria have to be fulfilled by the jet candidates [188].

Besides the criteria just mentioned, in this analysis only jets with  $p_T > 25$  GeV and  $|\eta| < 2.5$  are considered.

## 5.5. Flavour tagging

Hadrons which consist of at least of one  $b$ -quarks are referred to as  $B$ -hadrons. Jets containing such  $B$ -hadrons, so-called  $b$ -jets, can be tagged by taking advantage of the distinct properties of  $B$ -hadrons, therefore simplifying the discrimination from hadrons with a lighter quark content. The relatively high mass of  $B$ -hadrons results in large transverse momenta of the decay products with respect to the jet axis, thus leading to a large angular distance between them. In addition, the  $B$ -hadron keeps a large fraction of the momentum of the original  $b$ -quark. As the decay of a  $b$ -quark into a top quark is kinematically forbidden and as the decay into a quark of the first or second generation highly suppressed,  $B$ -hadrons exhibit a relatively long lifetime compared to hadrons consisting of light quarks, resulting in travelling lengths in the detector of several millimetres for sufficiently high momenta. This leads to distinct secondary decay vertices which can be reconstructed separately from the primary vertex of the hard scattering. Therefore, several algorithms are used in ATLAS to target the different properties of  $b$ -quark decays and are afterwards combined into a multivariate  $b$ -tagging evaluation [189].

In this analysis,  $b$ -jets are identified using the MV2c10 algorithm [189] which is based on a boosted decision tree (cf. Sec. 7.1) using the output weights of the IP3D [189] and SV1 [190] algorithms as well as of JetFitter [191]. While IP3D is an impact parameter algorithm, with the SV1 algorithm secondary vertices are determined. Thus, both algorithms make use of the long lifetime of  $B$ -hadrons. JetFitter serves for the reconstruction of the full  $B$ -hadron decay chain exploiting the topological structure of weak  $b$ - and  $c$ -hadrons decays inside the jet [189]. A cut on the MV2c10 output at 0.6459 has been applied corresponding to a  $b$ -jet selection efficiency of 77% as determined in simulated  $t\bar{t}$  events. The related factors for the rejection of charm- and light-flavoured jets are then 6 and 134, respectively, meaning that only 1 out of 6 (134)  $c$ -jets (light jets and gluons) is misidentified as a  $b$ -jet. The rejection for  $\tau$ -leptons has a value of 22 [189]. These numbers correspond to an efficiency of 83% for  $c$ -jet rejection, 99% for light jet rejection and 95% for the rejection of  $\tau$ -leptons, respectively. The  $b$ -tagging procedure has been applied to all jets satisfying the requirements stated in Sec. 5.4.

To account for small discrepancies between the  $b$ -tagging performance in data and Monte Carlo simulations, simulated events selected by  $b$ -jet identification are calibrated according to their jet flavour with scale factors dependent on the transverse momentum and the pseudorapidity of the jets. Therefore,  $b$ -tagging scale factors for  $b$ -jets are determined in  $t\bar{t}$  events [192], while for  $c$ -jets the number of reconstructed  $D^{*+}$  decays within jets is compared before and after applying the tagging requirements [193]. Finally, the misidentification efficiency of light jets is evaluated in dijet samples using a negative tag method. As  $B$ -hadrons exhibit a large branching fraction to leptons, it is particularly important to apply overlap removal techniques wiping off these soft leptons (cf. Sec. 5.7) in order to avoid double counting.

## 5.6. Missing transverse energy

The momenta of incoming particles within a proton-proton collision at the ATLAS experiment are almost entirely aligned in  $z$ -direction. Protons are no elementary particles but made up of numerous partons, therefore momentum conservation can only be applied in the transverse plane assuming that before the collision the momenta of the proton constituents are negligibly small perpendicular to their movement.

The missing transverse momentum  $\mathbf{p}_T^{\text{miss}}$  is defined as the negative vector sum of all selected,

reconstructed and calibrated objects in an event,

$$\mathbf{p}_T^{\text{miss}} := - \sum_{i=1}^{N_{\text{objects}}} \vec{p}_{T,i} \quad (5.3)$$

with its magnitude denoted as the more familiar missing transverse energy  $E_T^{\text{miss}}$  [182, 194, 195]. Transverse momentum measurements which cannot be associated to a fully reconstructed object, including tracks originating from the primary vertex and calorimeter clusters, are used to form so-called soft term contributions. These are either derived from energy deposits in the calorimeter or as track-based soft terms (TST) from track measurements in the inner detector, where the latter provides a better performance under the Run 2 pile-up conditions. The resolution of missing transverse momentum reconstruction can be further improved by using the jet vertex tagging technique which associates tracks from jets to primary vertices, thus enabling the removal of pile-up jets.

Stable particles which hardly interact with matter, so-called weakly interacting particles, traverse the detector without leaving significant signatures. Hence, they cannot be reconstructed by direct measurements. However, the kinematic properties of weakly interacting particles can be quantified exploiting momentum conservation in the transverse plane of the detector. If in a given event no such particle appears, the total momentum of the outgoing particle will be zero in the transverse plane. Consequently, weakly interacting particles manifest themselves as an imbalance of visible momenta in the transverse plane which is reconstructed as missing transverse energy. Additionally, the azimuthal angle  $\phi^{\text{miss}}$  specifies its angular position in the transverse plane. With only one weakly interacting particle the event is still fully kinematically reconstructable despite the  $z$ -component of the missing transverse momentum. Though, if more than one weakly interacting particle appears in the event, it is not possible to distinguish them anymore as their momentum vectors are added up resulting in a completely new momentum vector. In this case other highly sophisticated techniques have to be used. In the Standard Model, the only particles that can cause missing transverse energy are neutrinos, but extended theories like SUSY predict additional particles which are invisible to the detector.

## 5.7. Overlap removal

At the ATLAS detector, for the reconstruction of different physical objects, e.g. jets, electrons and muons, specialised algorithms for each object type are used. Since the various algorithms rely on shared information from the various subdetectors, one actual physics object might be reconstructed multiple times as a different kind. In order to avoid double counting of a single final state object, like e.g. an isolated electrons being reconstructed both as electron and as a jet with the requirements above, a procedure called overlap removal [196] is applied. If two reconstructed objects lie too close to each other, one is removed whereas the other is kept, measured by the distance in the  $\eta$ - $\phi$ -plane which is defined by  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ .

Firstly, electron candidates sharing a track in the inner detector with muon candidates are removed. To prevent double counting of electron energy deposits as jets, the closest jet within a distance of  $\Delta R < 0.2$  to a reconstructed electron is removed. In contrast, if the distance between the nearest jet and the electron candidate is  $0.2 < \Delta R < 0.4$ , the electron is discarded to ensure it is sufficiently separated from the nearby jet activity. To reduce the background from heavy flavour decays inside jets, muon candidates are removed if they are separated from the nearest jet by  $\Delta R < 0.4$ . However, if this jet has fewer than three associated tracks, the muon is kept and the jet is dropped instead in order to avoid an inefficiency for high-energy muons undergoing a significant energy loss in the calorimeter system.

## 6. Event selection and reconstruction

In order to extract the  $t\bar{t}Z$  1-lepton signal process from the tremendous amount of data collected at the LHC, dedicated selection criteria have to be applied. The purpose of these criteria is not only to choose events of interest but also to reject background contributions. On top, the selected events are used to reconstruct the original  $t\bar{t}Z$  system.

### 6.1. Event selection

As described in Sec. 2.1.4, a top quark is assumed to exclusively decay into a  $W$  boson and a  $b$ -quark. Hence, the  $t\bar{t}Z$  a-lepton final state is characterised by one charged light lepton<sup>(1)</sup>, one neutrino and six jets originating from bottom or light quarks, respectively. In order to pick such events, multiple selection criteria have been applied.

First, to guarantee data quality, only data events listed in the good run lists mentioned in Sec. 4.5 are selected ensuring a fully operational detector during the data taking. Also events in which the liquid argon or tile calorimeters are flagged to be in error state are removed from the analysis.

Depending on the flavour of the light lepton, the events are split into two categories, the electron and the muon channel. Only events in which either a single-electron or a single-muon trigger has fired are selected. In this analysis, different sets of single-lepton high level triggers with different criteria on momentum/energy and identification, used in a logical “or”-disjunction, are applied. In Tab. 6.3 an overview of the triggers which have been applied in this analysis is given. From the names the requirement imposed on the identification criterion can be determined as well as the threshold above which the energy/momentum of the selected objects must lie. To be selected, the higher the energy/momentum of leptons, the less tighter are the identification criteria they have to fulfil and vice versa. Additionally, for both electrons and muons the triggers with the lowest energy/momentum threshold demand dedicated isolation criteria.

To account for the high-energetic charged lepton from one  $W$  boson decay, exactly one electron (muon) with  $E_T > 25$  GeV ( $p_T > 25$  GeV) which has to be matched to a given trigger decision for the data collected in 2015, and with  $E_T > 27$  GeV ( $p_T > 27$  GeV) and  $|\eta| < 2.47$  ( $|\eta| < 2.5$ ) for data from 2016 is required. In addition, for electrons the so-called crack region ( $1.37 < |\eta| < 1.52$ ) is excluded. Since the  $Z$  boson is assumed to decay hadronically, at least six jets with  $p_T > 25$  GeV and  $|\eta| < 2.5$  have to be reconstructed in an event. In order to suppress background contributions from the associated production of  $W$  bosons with jets at least one jet has to be  $b$ -tagged employing the MV2c10  $b$ -tagging algorithm (cf. Sec. 5.5) at a working point with an efficiency of 77%. The background from non-prompt and falsely identified leptons is suppressed by applying requirements in a two-stage way on the missing transverse energy and the transverse  $W$  boson mass, defined as

$$m_T^W := \sqrt{2p_T^l E_T^{\text{miss}}(1 - \cos \Delta\phi)} \quad , \quad (6.1)$$

where  $\Delta\phi$  is the angular difference between the charged lepton  $l$  and the missing transverse momentum. In the electron channel, first both requirements  $E_T^{\text{miss}} > 30$  GeV and  $m_T^W > 30$  GeV have to be fulfilled. In a second stage, the events have to satisfy the inclusive condition of hav-

<sup>(1)</sup> Because of the rather short lifetime of  $\tau$ -leptons, only electrons or muons are considered in this analysis.

## 6. Event selection and reconstruction

ing either  $E_T^{\text{miss}} > 40 \text{ GeV}$  or  $m_T^W > 50 \text{ GeV}$ . For the muon channel a single step is sufficient requiring  $E_T^{\text{miss}} + m_T^W > 60 \text{ GeV}$ .

At this stage of the analysis no data-driven background estimation has been done. The contribution of fake and non-prompt lepton backgrounds is intended to be estimated using the data-driven matrix method [197, 198]. The number of background events originating from the production of  $W$  bosons in association with jets will be evaluated using both data and Monte Carlo simulation. For the shape of this background simulations are sufficient, while the normalisation of this background is supposed to be estimated by exploiting the expected asymmetry of the charge of leptons originating from  $W$  boson decays [199].

Process	Electron channel		Muon channel		Combined $e + \mu$	
$t\bar{t}Z$ signal	732 ±	5	771 ±	5	1503 ±	8
$t\bar{t}$ +jets	152655 ±	249	174828 ±	258	327469 ±	358
Single top	7049 ±	49	7774 ±	50	14823 ±	70
$W$ +jets	14075 ±	110	18837 ±	145	32914 ±	182
$Z$ +jets	2777 ±	47	2577 ±	31	5354 ±	56
Diboson	1249 ±	35	1387 ±	35	2635 ±	49
$t\bar{t}H$	563 ±	2	595 ±	2	1158 ±	3
$t\bar{t}W$	517 ±	2	562 ±	2	1079 ±	3
$t\bar{t}+b\bar{b}$	121 ±	2	206 ±	3	326 ±	4
$tZ$ +jets	41 ±	1	43 ±	1	84 ±	1
$4t$	31 ±	0	30 ±	0	61 ±	0
$t\bar{t}WW$	22 ±	0	22 ±	0	44 ±	1
$tWZ$	15 ±	1	12 ±	1	27 ±	1
$VH$	11 ±	2	11 ±	3	22 ±	4
$gg \rightarrow H$	13 ±	13	0 ±	0	13 ±	13
$H \rightarrow b\bar{b}$	2 ±	1	1 ±	1	3 ±	2
VBF $H$	0 ±	0	2 ±	1	2 ±	1
Triboson	1 ±	0	0 ±	0	1 ±	0
Total background	179141 ±	283	206887 ±	303	386015 ±	415
Total expected (MC)	179873 ±	283	207658 ±	303	387518 ±	415

Tab. 6.1.: Final event yields after all selection criteria have been applied. The expected contributions from Monte Carlo simulations have been scaled to an integrated luminosity of  $36.1 \text{ fb}^{-1}$  in order to match the data luminosity. The numbers given in the table are rounded to next whole number. The associated uncertainties correspond to the statistical uncertainties only.

The final event yields after applying all selection criteria are given in Tab. 6.1. Event yields and uncertainties given as zero are, except for VBF  $H$  in the electron channel, not equal to zero but fairly smaller than one. Due to rounding effects the sum of both channels might differ from the numbers given in the fourth column. Both signal and background are taken from Monte Carlo simulation. As expected, the main background contribution is the production of top quark pair events in association with additional jets, followed by single top quark events and the associated production of a vector boson ( $W/Z$ ) with jets. Moderate background contributions are diboson production and the associated production of top quark pairs with a  $W$  or Higgs boson which are approximately of the same amount as the signal contribution. All other backgrounds contribute

to a smaller extend. For all upcoming analysis steps the background processes from Higgs boson production via vector boson fusion, the production of Higgs boson with their subsequent decay to  $b$ -quark pairs and triboson events are not taken into account anymore as their contribution is smaller than 1 % of the signal amount.

However, with the applied selection cuts the ratio of  $t\bar{t}$  events to  $t\bar{t}Z$  signal events is reduced from the theoretical value 991.66 to 208.56, 226.73 and 217.87 for the electron channel, the muon channel and the combination of both, respectively, which means a reduction of about 78 %. For each channel the corresponding statistical significance  $Z$ , i.e. the probability that an observed excess is caused by the presence of signal and is not originating from statistical fluctuations of the background, has been calculated according to

$$Z = \frac{S}{\sqrt{S + B + \Delta B}} \quad , \quad (6.2)$$

where  $S$  and  $B$  denote the number of signal and background events and  $\Delta S$  and  $\Delta B$  represent their statistical errors, respectively. The results are summarised in Tab. 6.2, together with the event ratios of  $t\bar{t}/t\bar{t}Z$  and its decrease with respect to the theoretical value of 991.66. In order to estimate the impact of both a statistical and a systematic uncertainty, the significance values are calculated three times, with either no error, only a statistical error or with both a statistical and systematic error on the input values. As no systematic uncertainties are estimated at this stage of the analysis, it is assumed to amount  $\sim 30\%$ . As it can be clearly seen, while the statistical uncertainties do almost not affect the significance values, the incorporation of a systematic uncertainty considerably reduces the sensitivity by a factor of around 12 % in all three channels. Due to higher statistics, the combination of the electron and the muon channel seems to deliver the best sensitivity so far, for which reason in the subsequent analysis steps no explicit separation of the two channels is done anymore.

	Electron channel	Muon channel	Combined
$t\bar{t}/t\bar{t}Z$	208.56	226.73	217.87
decrease w.r.t. theory [%]	78.97	77.14	78.01
significance (no err.)	1.726	1.692	2.415
significance (stat. err.)	1.724	1.691	2.413
significance (stat. & syst. err.)	1.514	1.485	2.119

Tab. 6.2.: Summary of the gain in sensitivity achieved by applying the selection criteria given above, listed for both the electron and the muon channel separately and for the combination of them, respectively. The calculation of the statistical significance according to Eq. (6.2) has been done with either no uncertainty, only the statistical error or both the statistical and the systematic uncertainties, assuming a systematic uncertainty of 30 % at this stage of the analysis. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

## 6. Event selection and reconstruction

Data taking period	Single-electron trigger	Explanation
2015	HLT_e24_lhmedium_L1EM20VH	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 24</math> GeV</li> <li>• Medium likelihood identification</li> <li>• Additional requirements on level-1 trigger: electromagnetic cluster with <math>E_T &gt; 20</math> GeV, thresholds on energy variation with <math> \eta </math>, veto on hadronic core</li> </ul>
	HLT_e60_lhmedium	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 60</math> GeV</li> <li>• Medium likelihood identification</li> </ul>
	HLT_e120_lhloose	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 120</math> GeV</li> <li>• Loose likelihood identification</li> </ul>
2016	HLT_e26_lhtight_nod0_ivarloose	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 26</math> GeV</li> <li>• Tight identification criteria</li> <li>• No information from impact parameter used</li> <li>• Loose isolation criteria based on track information</li> </ul>
	HLT_e60_lhmedium_nod0	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 60</math> GeV</li> <li>• Medium likelihood identification</li> <li>• No information from impact parameter used</li> </ul>
	HLT_e140_lhloose_nod0	<ul style="list-style-type: none"> <li>• <math>E_T &gt; 120</math> GeV</li> <li>• Loose likelihood identification</li> <li>• No information from impact parameter used</li> </ul>
Data taking period	Single-muon trigger	Explanation
2015	HLT_mu20_iloose_L1MU15	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 20</math> GeV</li> <li>• Loose isolation criteria based on calorimeter information</li> <li>• Additional requirement on level-1 trigger: track with <math>p_T &gt; 20</math> GeV</li> </ul>
	HLT_mu50	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 50</math> GeV</li> </ul>
2016	HLT_mu26_ivarmedium	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 26</math> GeV</li> <li>• Medium isolation criteria based on track information</li> </ul>
	HLT_mu50	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 50</math> GeV</li> </ul>

Tab. 6.3.: Overview of the single-lepton triggers which have been applied in this analysis. Since the object definitions are not equal for the 2015 and 2016 data taking period, a slightly different set of triggers has to be used for 2015 and 2016 data, respectively. In the right column a list of requirements the objects have to fulfil in order to make the trigger fire is given. The here presented triggers are high-level triggers (HLT).

## 6.2. Monte Carlo distributions

In this section few distributions from Monte Carlo simulations are shown representing some basic variables after the event selection. For all images the signal component is enhanced by a factor of 100 and placed in front of the stacked background, in order to make it visible against the overwhelming background amount reflecting the huge overspill of  $t\bar{t}$  events depicted in green. Some kinematic distributions as the missing transverse energy and the transverse  $W$  boson mass or lepton-related quantities are shown in Fig. 6.1, whereas Fig. 6.2 and 6.3 present distributions of jet-related and composite kinematic variables as well as of counting variables. All distributions depicted in Fig. 6.1 show an almost perfect agreement of the shapes of signal and background. A rather similar behaviour can be noticed for the variables shown in Fig. 6.2 except for the transverse momentum of the leading jet in the upper right image. Also the distributions of  $H_T$  and  $S_T$  on the left side and the counting variables on the right side of Fig. 6.3 exhibit considerable separation power in their shapes between signal and background. However, despite the need of enhancing the signal purity, the signal can still not be separated using simple cuts on the different distributions. Hence, in order to search for variables with more separation power, the  $t\bar{t}Z$  system has been fully reconstructed using the detected final state objects as explained in the following section.

## 6. Event selection and reconstruction

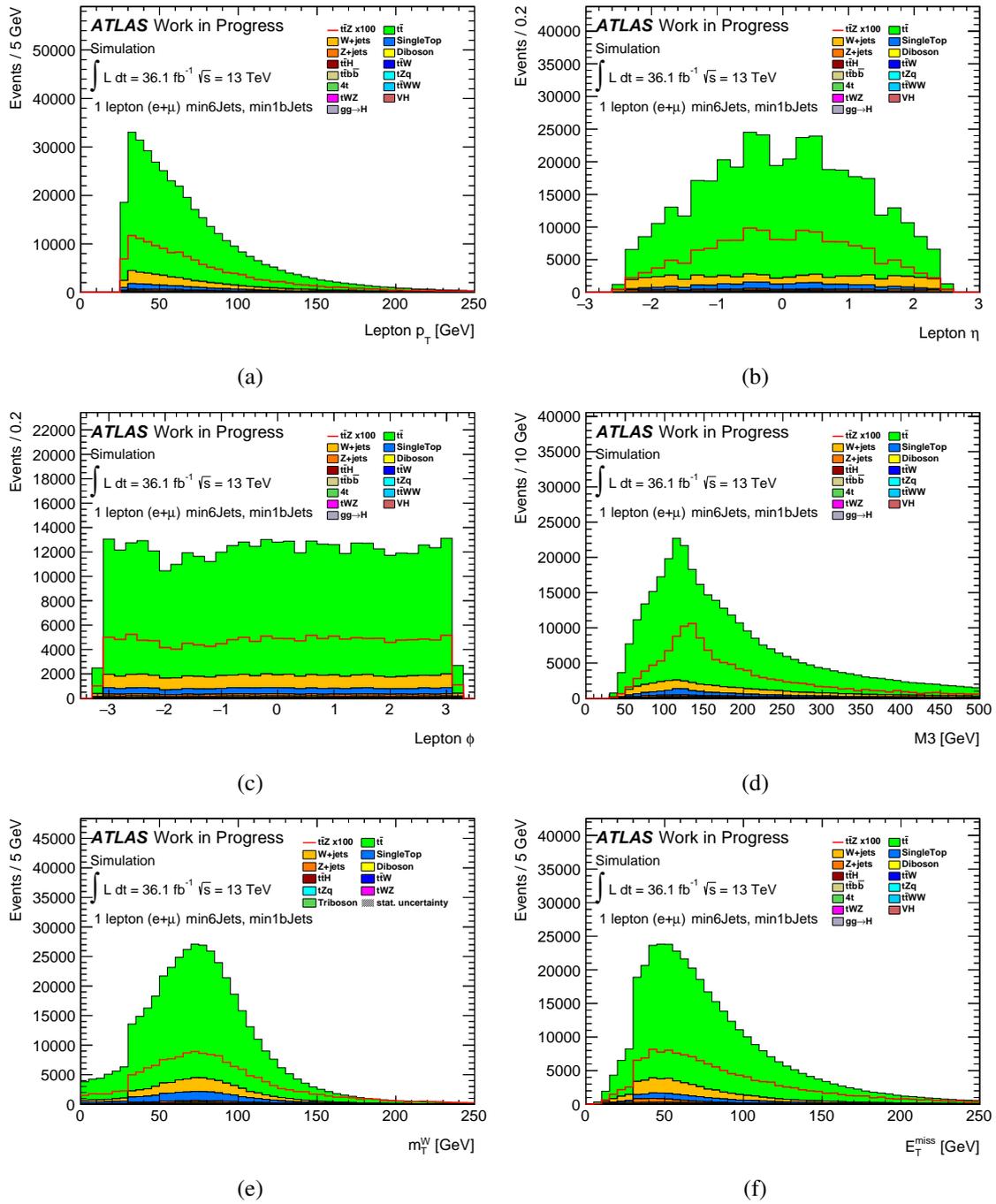


Fig. 6.1.: Presentation of kinematic distributions for the selected events using Monte Carlo simulations. Fig. (a)–(c) show the lepton transverse momentum, pseudorapidity and azimuthal angle, respectively. In Fig. (d) the  $M3$  variable is presented, referring to the mass of the three jets with the highest vectorially summed transverse momentum, while Fig. (e) depicts the transverse  $W$  boson mass and Fig. (f) the missing transverse energy. Here, the selection cuts can be clearly spotted indicated from the sharp vertical lines at 30 GeV.

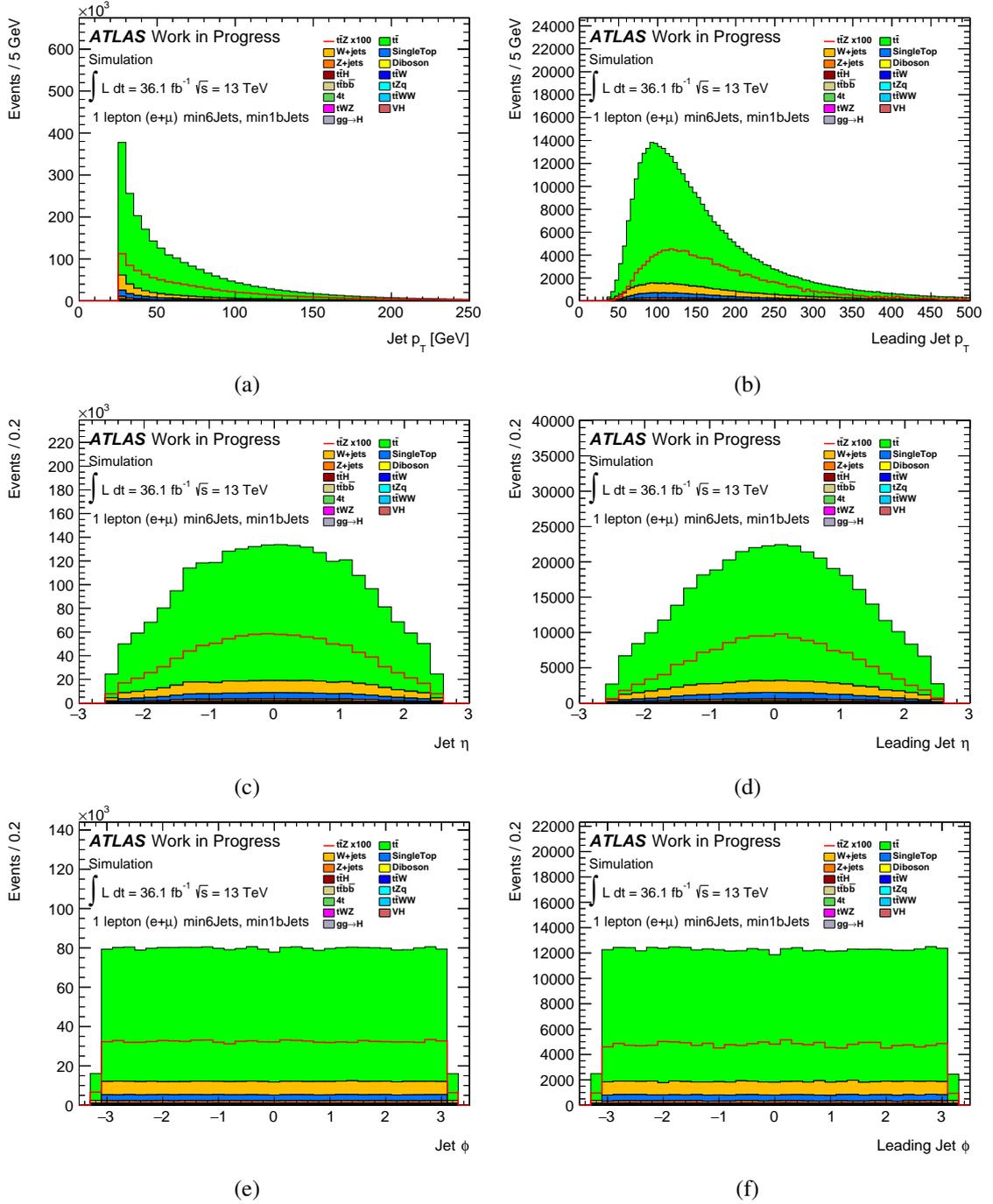


Fig. 6.2.: Presentation of jet-related distributions for the selected events using Monte Carlo simulations. Fig. (a)–(c) show the transverse momentum, pseudorapidity and azimuthal angle of all jets in a given event, whereas in Fig. (d)–(f) the transverse momentum, pseudorapidity and azimuthal angle of the leading jets in a given event is depicted, respectively.

## 6. Event selection and reconstruction

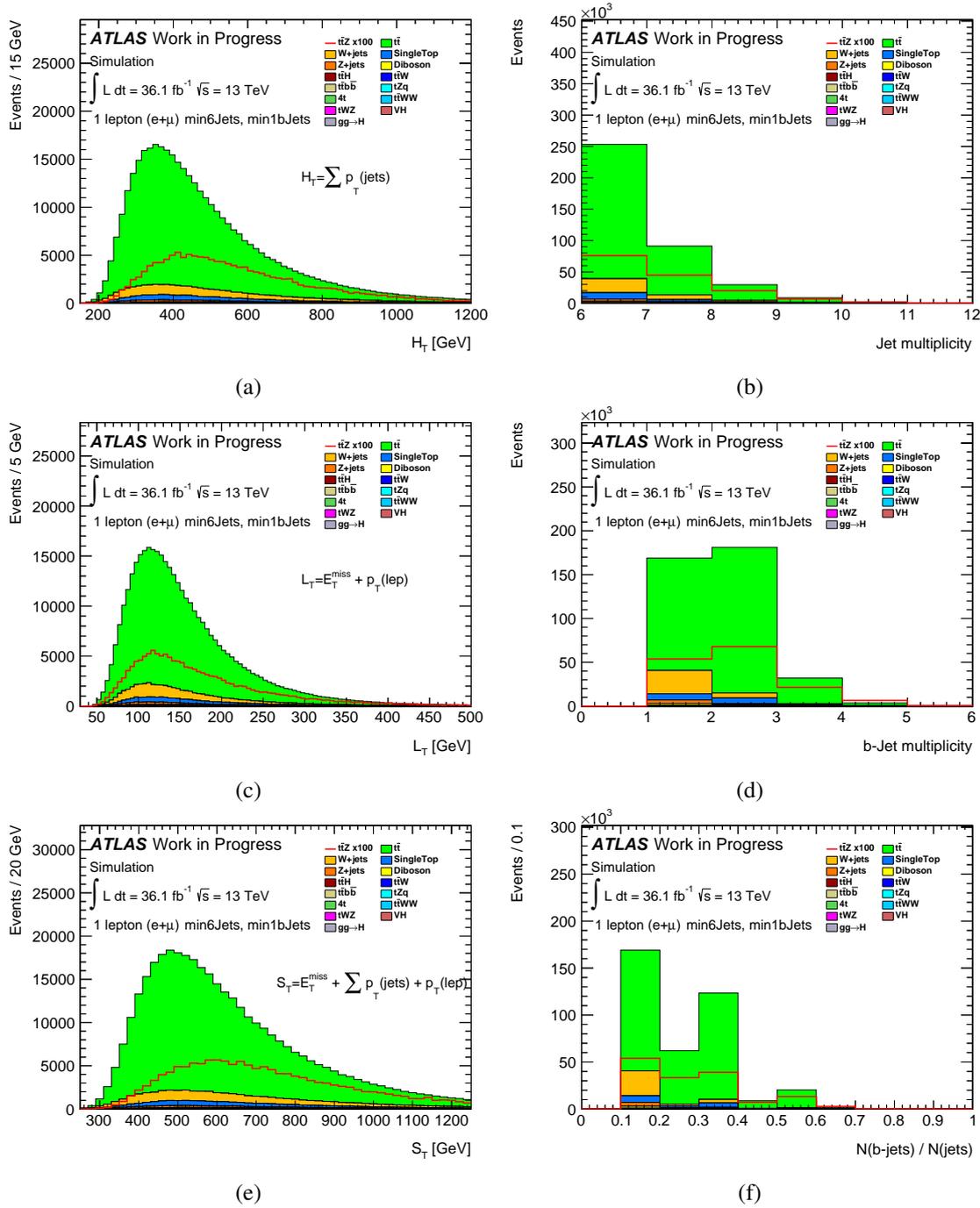


Fig. 6.3.: Presentation of distributions of composite variables for the selected events using Monte Carlo simulations. Fig. (a)–(c) show the sum of the transverse momenta of all jets in a given event, the sum of the lepton transverse momentum and the missing transverse energy, and the sum of all transverse momenta in an event, referred to as  $H_T$ ,  $L_T$  and  $S_T$ , respectively. On the contrary, Fig. (d)–(f) depict counting variables as the jet multiplicity, the  $b$ -jet multiplicity and the ratio of them.

### 6.3. Event reconstruction

As mentioned in Section 2.3.2, due to its high mass, the top quark has a lifetime much shorter than the time scale of hadronisation. It is hence possible to study its properties directly, based on the decay products and their corresponding signatures in the detector, i.e. jets, charged leptons and missing transverse momentum. A full reconstruction of the top quark four-momentum is essential for many precision measurements of top quark properties, e.g. spin, as they are smeared by hadronisation otherwise.

The quarks in the final state of the hard-scattering process are detected as jets. The number of jets only agrees with the number of quarks if no jet is lost<sup>(2)</sup> and if no additional jets<sup>(3)</sup> are present. However, even in case of a clear event jets cannot be associated uniquely to the partons of the hard-scattering process, since it is not a priori known how to assign a final state object to a parton of the physics process. Hence, reconstruction algorithms have to be used to find the best corresponding match between them. Shortcomings in these reconstruction algorithms result in incorrect jet-parton assignments, also referred to as combinatorial background, which may decrease the precision of measurement.

As stated in Sec. 6.1, the  $t\bar{t}Z$  1-lepton channel final state, depicted in Fig. 2.13, consists of a charged light lepton, a neutrino which manifests itself as missing transverse energy in the detector, and six jets<sup>(4)</sup> whereof at least one has to be tagged as a  $b$ -jet, implicating  $6! = 720$  possible jet-parton assignments in the first instance. Because of the indistinguishability of the decay products of the  $W$  and the  $Z$  boson, this huge number can indeed be lowered by a factor of two for each vector boson, which results in 180 remaining combinations. In addition,  $b$ -tagged jets are not allowed to be set on the position of a light-flavoured<sup>(5)</sup> quark. The requirement of at least one  $b$ -jet in the event selection (cf. Sec. 6.1) further reduces the count of possible permutations down to at the most 60, depending on the number of  $b$ -tagged jets. Consequently, with the help of combinatorics, the number of valid permutations in a  $t\bar{t}Z$  1-lepton channel event can be mathematically expressed as

$$N_{\text{Permutations}} = \frac{n_{\text{light}}! n_{b\text{-tag}}!}{4} \cdot \begin{cases} 2 & n_{b\text{-tag}} \text{ odd} \\ 1 & n_{b\text{-tag}} \text{ even} \end{cases} \quad (6.3)$$

where the factor of 4 in the denominator comes from the indistinguishability of the decay products of the  $W$  and the  $Z$  boson. Since the number of  $b$ -tagged and non- $b$ -tagged jets always have to sum up to the same value,

$$n_{\text{light}} + n_{b\text{-tag}} = 6 \quad , \quad (6.4)$$

Eq. (6.3) can be rewritten as a function of the number of  $b$ -tagged jets,

$$N_{\text{Permutations}}(n_{b\text{-tag}}) = \frac{(6 - n_{b\text{-tag}})! n_{b\text{-tag}}!}{4} \cdot 2^{(n_{b\text{-tag}} \bmod 2)} \quad (6.5)$$

In this analysis the number of  $b$ -quarks in the final state is dynamically varied in such a way that for events containing up to two  $b$ -tagged jets the decay particles from the  $Z$  boson are requested to be light-flavoured, whereas in the case of more than two  $b$ -jets the reconstruction explicitly aims to consider the  $Z$  boson as decaying into a pair of  $b$ -quarks. According to Eq. (6.5) the

<sup>(2)</sup> A jet could lie outside the detector acceptance or does not pass the minimum selection criteria. Besides it could be misidentified as a soft lepton and vice versa, albeit not very often.

<sup>(3)</sup> Such jets likely stem from initial or final state radiation.

<sup>(4)</sup> In this analysis the first six jets in an event had been used. This choice is based on the fact that jets are stored in the Monte Carlo datasets ordered by their transverse momentum, so softer jets from QCD or other radiation effects should have higher numbers.

<sup>(5)</sup>  $u, d, s, c$

number of permutations left the algorithm has to deal with first decreases, but increases again for sufficiently enough  $b$ -jets in an event. This is shown in Fig. 6.4 where the number of possible jet-parton assignments is plotted against the number of  $b$ -tags in an event. As a side remark, for more than 3  $b$ -jets in an event the factor of  $2^{(n_{b\text{-tag}} \bmod 2)}$  is omitted because it only makes sense if one of the  $b$ -quark positions could remain unoccupied by a  $b$ -jet which cannot be the case for  $n_{b\text{-tag}} \geq 4$ . Altogether, following this rule, the correct numbers of valid permutations in a  $t\bar{t}Z$  1-lepton channel event are those given in Tab. 6.4.

Number of $b$ -tags	Number of permutations
0	180
1	60
2	12
3	18
4	12

Tab. 6.4.: Number of valid permutations in a  $t\bar{t}Z$  1-lepton event as a function of the number of  $b$ -tags in an event. It is assumed that an interchange of the decay products of the  $W$  and  $Z$  bosons is redundant thus reducing the number of possible jet-partons assignments by a factor of 4 in the first place.

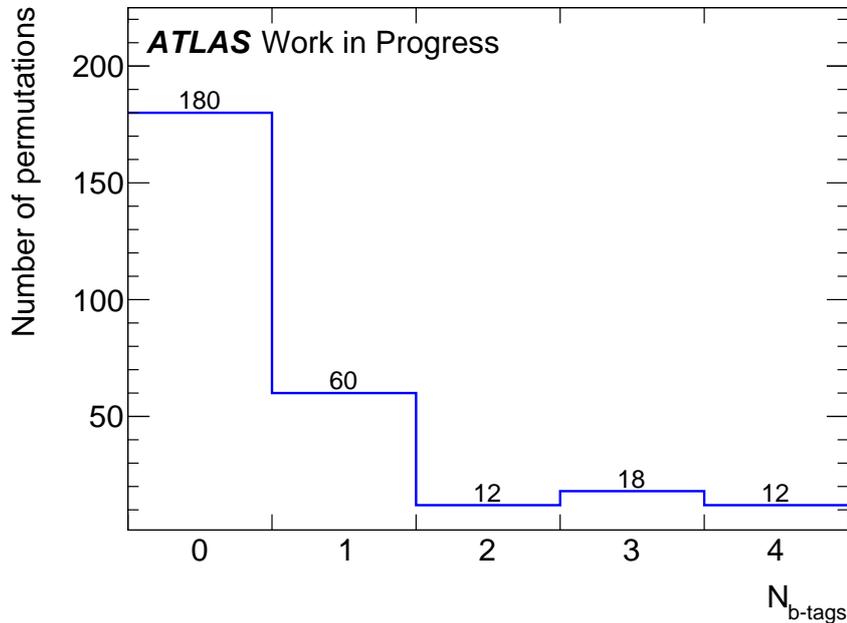


Fig. 6.4.: Number of valid permutations in a  $t\bar{t}Z$  1-lepton event as a function of the number of  $b$ -tags in an event. It is assumed that an interchange of the decay products of the  $W$  and  $Z$  bosons is redundant thus reducing the number of possible jet-partons assignments by a factor of 4 in the first place.

### 6.3.1. KLFFitter

In this analysis, the jet-parton assignment, i.e. the reconstruction of the  $t\bar{t}Z$  system has been done with a kinematic likelihood approach using the Kinematic Likelihood Fitter (KLFFitter) [200, 201]. This is a package written in C++ for the estimation of precise physical parameters using the maximum likelihood method. With its structure based on ROOT [202], for maximisation routines it deploys the Bayesian Analysis Toolkit (BAT) [203] which provides the required numerical instruments and also makes use of the numerical minimisation program MINUIT<sup>(6)</sup> [204]. The KLFFitter works stand-alone, but is also implemented in the *Athena* software framework or the *AnalysisTop* framework used in the ATLAS experiment, and can easily be adjusted to many other analysis frameworks. Originally, this tool had been developed for top quark pair reconstruction in the lepton+jets channel but it can be applied straight forward to processes other than top quark pair production.

In general, a likelihood is defined as the probability for observing a set of measured quantities given a model and a corresponding set of model parameters. The probability is then calculated according to the model, here the  $t\bar{t}Z$  1-lepton channel. The quantities from the measurements which define the dataset are the energies of the six jets  $\tilde{E}_i$ , the energy of the charged lepton  $\tilde{E}_{\text{lep}}$ <sup>(7)</sup> and the missing transverse energy stemming from the neutrino.

The energies of the quarks and the charged lepton are only measured with non-vanishing uncertainties given by the corresponding energy resolution. These uncertainties are not necessarily distributed symmetrically and can vary with energy, for which reason they are parametrised by transfer functions,  $W(\tilde{E}_i | E_i)$ , which map the measured energy of an object  $\tilde{E}_i$  to the true energies of the final state particles  $E_i$ . Although neutrinos leave no direct signal in the detector, their momentum can be obtained by applying energy and momentum conservation<sup>(8)</sup> to all measured objects. Hence, transfer functions for the  $x$ - and  $y$ -components of the neutrino momentum,  $W(\tilde{E}_{x,y}^{\text{miss}} | p_{x,\nu})$ , can be defined. It is assumed that all transfer functions are well-known and all jets are detected. More information about the KLFFitter transfer functions is given in Appendix B.

For setting up the likelihood function, a couple of additional conditions has to be fulfilled. As a first constraint, the invariant di-jet mass  $m_{jj}$  and the invariant mass of the charged lepton and the neutrino  $m_{l\nu}$  coming from the  $W$  boson decay are required to be Breit-Wigner ( $BW$ ) distributed around a pole mass of  $M_W = 80.385$  GeV within a decay width  $\Gamma_W = 2.085$  GeV [34]. Secondly, the invariant top quark masses  $m_{jjj}$  and  $m_{jl\nu}$  must also follow a Breit-Weigner distribution. In addition the invariant di-jet mass of the  $Z$  boson must be Breit-Wigner distributed as well, with a pole mass of  $M_Z = 91.1876$  GeV and a width of  $\Gamma_Z = 2.4952$  GeV [34]. In the KLFFitter package, the user can choose whether the top pole mass should be fixed to a certain value within the fit, as done in this analysis, or whether it is taken into account as an additional free parameter. It either equals the value taken from the MC generation ( $M_{\text{top}} = 172.5$  GeV) or it comes from a previous measurement. Since this analysis has its focus on the proper reconstruction of the additional  $Z$  boson, the top quark mass and width are fixed to the current measured values of  $M_{\text{top}} = 173.21$  GeV and  $\Gamma_{\text{top}} = 1.41$  GeV [34], respectively. The normalised and relativistic Breit-Wigner functions used in the fitting procedure are given by

$$BW(m|M) = \frac{2}{\pi} \frac{\Gamma M^2}{(m^2 - M^2)^2 + \Gamma^2 M^2} \quad , \quad (6.6)$$

where  $\Gamma$  depicts the decay width of a particle with mass  $m$  centred around  $M$ . Thus, the respective

<sup>(6)</sup> MINUIT was originally written in Fortran, but ROOT provides a straight forward conversion into C++.

<sup>(7)</sup> In case of a muon the transverse momentum  $p_{T,\text{lep}}$  has to be taken instead of the energy

<sup>(8)</sup> At the LHC only applicable in the transverse plane of the detector

probability densities (Eq. (6.6)) constrain the combined momenta of the final-state particles which are built according to Eq. (6.7). Here Eq. (6.7a) represents both the hadronically decaying  $W$  and the  $Z$  boson and Eq. (6.7b) the mass of the hadronically decaying  $W$  boson. The masses of the two top quarks are calculated according to Eq. (6.7c) and Eq. (6.7d), respectively.

$$m_{jj}^2 = (p_q + p_{\bar{q}'})^2 \quad (6.7a)$$

$$m_{l\nu}^2 = (p_l + p_\nu)^2 \quad (6.7b)$$

$$m_{jjj}^2 = (p_q + p_{\bar{q}'} + p_b)^2 \quad (6.7c)$$

$$m_{jl\nu}^2 = (p_l + p_\nu + p_{\bar{b}})^2 \quad (6.7d)$$

The  $z$ -component of the neutrino momentum is calculated separately from the initial values, resulting in two solutions whereof the one giving rise to a larger likelihood value is taken. In case one solution is imaginary it is set to zero and therefore not taken into account any further. The derivation of  $p_Z^\nu$  can be found in Appendix C. Taking into account the transfer functions of the measured objects and the constraints just discussed, the following 10 (11) parameters enter the fitting procedure:

- The transverse energies of the six quarks  $E_i$  and the charged lepton  $E_{\text{lep}}$  (7 parameters)
- The momentum components of the neutrino  $p_{x,y,z}^\nu$  (3 parameters)
- Optionally the top pole mass  $M_{\text{top}}$  (1 parameters)

During the kinematic fit, every parameter is varied within a certain range, which is set individually for each event. The energies of the partons and the charged lepton have to lie within an interval of

$$E \in \left[ \min \left( 0, \tilde{E} - n \cdot \sqrt{\tilde{E}} \right) < E < \tilde{E} + n \cdot \sqrt{\tilde{E}} \right]$$

around the measured values, where  $n = 7$  for partons/jets and  $n = 2$  for leptons<sup>(9)</sup>. The  $z$ -component of the neutrino momentum has to be within a range of  $\pm 1$  TeV, whereas the  $x$ - and  $y$ -components must not exceed a deviation of  $\pm 100$  GeV from the measured components of the missing transverse energy. In case of a free top quark mass, it is restricted into a range of  $M_{\text{top}} \in [100, 1000]$  GeV.

Finally, the heart of this method is the kinematic likelihood function which contains all the information of a particular event topology and, in the 1-lepton channel case, reads

$$\begin{aligned} L = & BW \{ m(q_1 q_2) | m_W, \Gamma_W \} \cdot BW \{ m(l\nu) | m_W, \Gamma_W \} \cdot \\ & BW \{ m(q_1 q_2 b_{\text{had}}) | m_{\text{top}}, \Gamma_{\text{top}} \} \cdot BW \{ m(l\nu b_{\text{lep}}) | m_{\text{top}}, \Gamma_{\text{top}} \} \cdot \\ & BW \{ m(q_{Z_1} q_{Z_2}) | m_Z, \Gamma_Z \} \cdot \\ & W \left( \tilde{E}_{jet_1} | E_{b_{\text{had}}} \right) \cdot W \left( \tilde{E}_{jet_2} | E_{b_{\text{lep}}} \right) \cdot W \left( \tilde{E}_{jet_3} | E_{q_1} \right) \cdot W \left( \tilde{E}_{jet_4} | E_{q_2} \right) \cdot \\ & W \left( \tilde{E}_{jet_5} | E_{q_{Z-1}} \right) \cdot W \left( \tilde{E}_{jet_6} | E_{q_{Z-2}} \right) \cdot \\ & W \left( \tilde{E}_x^{\text{miss}} | p_{x,\nu} \right) \cdot W \left( \tilde{E}_y^{\text{miss}} | p_{y,\nu} \right) \cdot \begin{cases} W \left( \tilde{E}_{\text{lep}} | E_{\text{lep}} \right) & \text{single electron channel} \\ W \left( \tilde{p}_{T,\text{lep}} | p_{T,\text{lep}} \right) & \text{single muon channel} \end{cases} \end{aligned} \quad (6.8)$$

As already mentioned in Section 6.3, an a priori association of jet with quarks is not possible, so all possibilities have to be permuted in order to find the most likely one. Again, the symmetry of the likelihood function (6.8) under the exchange of the decay particle of the hadronically decaying  $W$

<sup>(9)</sup> All ranges given in GeV

boson and the  $Z$  boson reduces the number of permutations, as well as weighting each permutation with  $b$ -tagging weights. For each possible jet-parton assignment, KLFitter tries to maximise the given likelihood function with respect to the assumed kinematics and particle properties given the particular event topology. In doing so, the algorithm starts from the measured values and assumes to have found the correct jet-parton assignment for the permutation with the highest likelihood value in an event. However, this does not imply that this permutation is indeed the true one. For computational reasons, instead of maximising the likelihood function  $L$ , the negative logarithm  $-\ln(L)$  is minimised, using the minimisation technique MINUIT.

In the end, apart from the fit parameters for each permutation and the corresponding value likelihood value, KLFitter also returns (i) the best fit parameters, (ii) the corresponding value of the likelihood, referred to as  $-\ln(L^*)$  or  $L^*$ , (iii) and a relative weight for each jet permutation, which gives the ordering of the permutations in an event. This weight is built from the likelihood value and the  $b$ -tagging weights for each permutation such that the sum of relative weights is normalised to one. Fig. 6.5 shows the distribution of the logarithmic likelihood  $\ln(L)$  obtained for the reconstruction of the  $t\bar{t}Z$  1-lepton channel applied in this analysis.

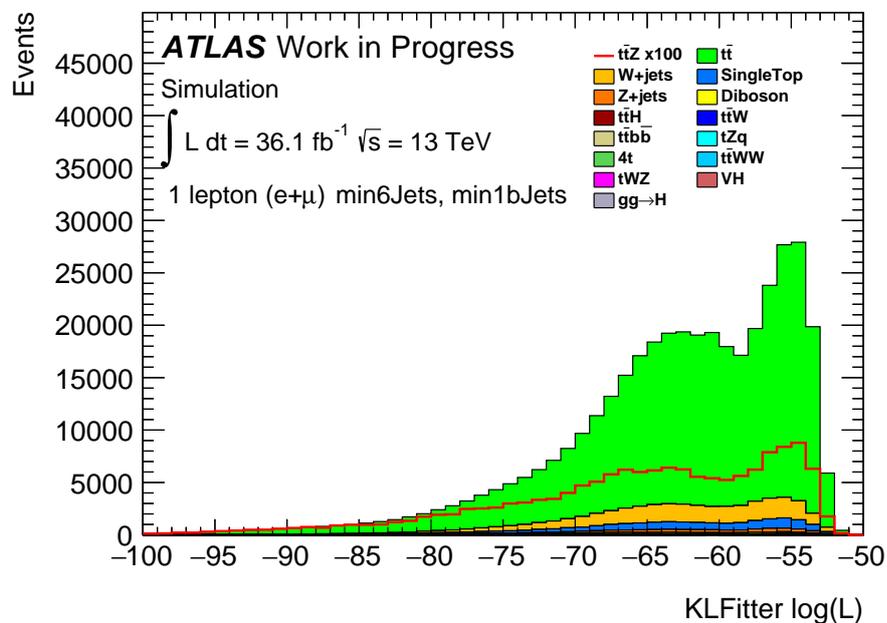


Fig. 6.5.: Distribution of the KLFitter logarithmic likelihood resulting from the fit. The signal process is scaled by a factor of 100 in order to make the signal visible.

As a technical remark, KLFitter stores not the measured, but the fitted and therefore varied variables. In addition, an index labelling the jet-parton assignment of the best permutation is returned. These labels indicate which jet of the input dataset is placed on which parton position of the final state signature exploiting the specific ordering of jets within the Monte Carlo samples. To be physically correct, the  $t\bar{t}Z$  system are reconstructed from the four-momenta not affected by the fit of their assumed decay particles, according to Eq. (6.7).

## 6.4. Signal region definition

The number of permutation the reconstruction algorithm has to run through per event clearly depends on the  $b$ -jet multiplicity (cf. Tab 6.4). As explained later on in Sec. 6.6 in detail, this

has a significant impact on the performance of the KLFilter. Thus, a grid of different jet and  $b$ -jet multiplicity, respectively, has been investigated in order to tag the regions of phase space in which the reconstruction performs best. Starting by constructing so-called exclusive regions in which an exact ( $b$ -)jet multiplicity is required, the scan over the phase space has been extended to inclusive regions which cover a larger part of the available phase space and therefore might improve the sensitivity. For each of those bins the significance corresponding to the amount of signal and background events left after applying the respective selection has been calculated, using Eq. (6.2) and assuming neither statistical nor systematic errors on the event numbers. The results of this procedure are shown in Fig. 6.6 depicting the calculated significance within each region for both exclusive and inclusive selection criteria, with no uncertainties taken into account. The corresponding plots for the significance calculations with uncertainties are depicted in Fig. E.1 and Fig. E.2.

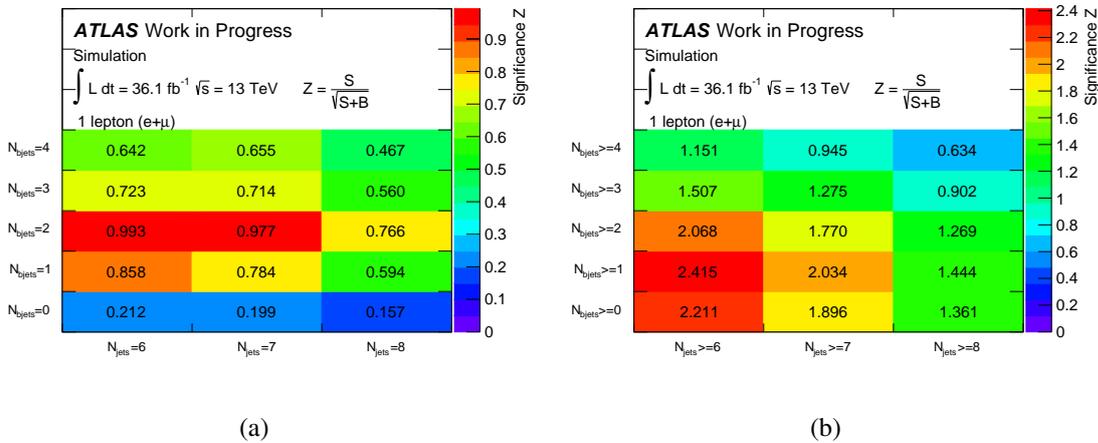


Fig. 6.6.: Illustration of the scan through the phase space of ( $b$ -)jet multiplicity, divided into exclusive (cf. Fig. (a)) and inclusive (cf. Fig. (b)) selection criteria. Both images exhibit two promising regions each, in the following referred to as signal regions. The calculation of the respective significance values has been done according to the formula depicted within the images assuming no uncertainties.

According to the respective values, the optimisation with respect to the ( $b$ -)jet multiplicity results in four promising regions, two of which originate from the exclusive and inclusive phase space grid, respectively. Thus, the further analysis is split into four parallel threads each focussing on one particular phase space region, in the following also referred to as signal region (SR). The four signal region considered are:

- **SR1** – the first inclusive region, with at least six jets and allowing for any  $b$ -jet multiplicity
- **SR2** – the second inclusive region, requiring at least six jets and at least one  $b$ -jet
- **SR3** – the first exclusive region, characterised by exactly six jets and two  $b$ -jets, respectively
- **SR4** – the second exclusive region, defined by exactly seven jets and exactly two  $b$ -jets

The selection criteria regarding lepton multiplicity, missing transverse energy and transverse  $W$  boson mass remain unchanged, whereas each signal region has other selection criteria on the ( $b$ -)jet multiplicity. For SR1 this means even relaxing the criteria made before (cf. Sec. 6.1). The event yields for each signal region are given in Tab. 6.5 for the combined  $e + \mu$  channel only. It is remarkable that although SR1 is missing any restriction on the  $b$ -jet multiplicity thus exhibiting a large amount of  $W$ +jets background, it still delivers the second best result for all

Process	SR1	SR2	SR3	SR4
$t\bar{t}Z$ signal	1629 ± 8	1503 ± 8	343 ± 3	203 ± 3
$t\bar{t}$ +jets	353477 ± 371	327469 ± 358	108902 ± 208	39646 ± 125
Single top	16610 ± 74	14823 ± 70	4552 ± 40	1424 ± 21
$W$ +jets	137716 ± 421	32914 ± 182	3639 ± 48	1333 ± 23
$Z$ +jets	20770 ± 115	5354 ± 56	719 ± 14	260 ± 5
Diboson	9836 ± 86	2635 ± 49	284 ± 14	126 ± 13
$t\bar{t}H$	1213 ± 3	1158 ± 3	205 ± 1	131 ± 1
$t\bar{t}W$	1181 ± 3	1079 ± 3	274 ± 2	145 ± 1
$t\bar{t}+b\bar{b}$	336 ± 4	326 ± 4	82 ± 2	34 ± 1
$tZ$ +jets	91 ± 1	84 ± 1	25 ± 1	9 ± 0
$4t$	62 ± 0	61 ± 0	2 ± 0	4 ± 0
$t\bar{t}WW$	49 ± 1	44 ± 1	6 ± 0	6 ± 0
$tWZ$	32 ± 1	27 ± 1	5 ± 0	3 ± 0
$VH$	44 ± 6	22 ± 4	7 ± 2	1 ± 1
$gg \rightarrow H$	114 ± 49	13 ± 13	13 ± 13	0 ± 0
Total background	541532 ± 585	386010 ± 415	118716 ± 219	43122 ± 130
Total expected (MC)	543161 ± 585	387513 ± 415	119059 ± 219	43326 ± 130

Tab. 6.5.: Final event yields for all signal regions after all selection criteria have been applied. The expected contributions from Monte Carlo simulations have been scaled to an integrated luminosity of  $36.1 \text{ fb}^{-1}$  in order to match the data luminosity. The numbers given in the table are rounded to next whole number. The associated uncertainties correspond to the statistical uncertainties only.

considered inclusive regions. To not great surprise, the two most sensitive exclusive signal regions are those fulfilling the twofold  $b$ -jet requirement of the top quark pair production signature. Here, the most sensitive signal region corresponds to the mere  $t\bar{t}Z$  jet multiplicity without any additional off-radiations, while for the second most sensitive one additional gluon off-radiation must have happened which is very likely at the LHC. The  $W$ +jets background, which is relatively strongly suppressed in SR3 and SR4, outperforms the  $t\bar{t}Z$  signal by a factor of almost 100 in SR1 as no  $b$ -jet veto has been applied there. This factor is significantly reduced within SR2 to roughly 20. The ratio between the  $W$ +jets and  $Z$ +jets backgrounds is nearly constant, amounting 6.3 and 6.2 in SR1 and SR2, respectively, whereas in SR3 and SR4 it is even the same within some rounding errors, namely 5.1. From this it can be concluded that the applied selection cuts have a rather similar impact on the  $Z$ +jets background as on  $W$ +jets. The diboson contribution is steadily further decreased when going from SR3 to SR4. The ratio of irreducible backgrounds  $t\bar{t}W$  and  $t\bar{t}H$  to the  $t\bar{t}Z$  signal remains roughly the same for each, oscillating between 1.3 and 1.5.

Alike in Sec. 6.1, the ratio of  $t\bar{t}$  to  $t\bar{t}Z$  events has been calculated for each signal region and therefrom the corresponding factor of increase in sensitivity with respect to the theoretical value of 991.66 (cf. (2.48)) has been derived. The results are listed in Tab. 6.6, together with the significance determined imposing either no uncertainties, only statistical uncertainties or both statistical and systematic uncertainties on the input values. Again, a systematic uncertainty of  $\sim 30\%$  is assumed. While statistical uncertainties do almost not affect the resulting significance, the impact of a systematic uncertainty of 30% manifests itself in all signal regions in a decrease of significance of roughly 12% which is the same as for the inclusive regions. It turns out that SR2,

	SR1	SR2	SR3	SR4
$t\bar{t}/t\bar{t}Z$	216.97	217.87	317.89	194.92
decrease w.r.t. theory [%]	78.12	78.01	67.94	80.34
significance (no err.)	2.211	2.415	0.993	0.977
significance (stat. err.)	2.209	2.413	0.992	0.976
significance (stat. & syst. err.)	1.939	2.119	0.871	0.858

Tab. 6.6.: Summary of the gain in sensitivity achieved by applying the selection criteria given above, listed for all signal regions. The calculation of the statistical significance according to Eq. (6.2) has been done with either no uncertainties, only the statistical uncertainties or both the statistical and systematic uncertainties, assuming a systematic uncertainty of 30 % at this stage of the analysis. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

which is identical to the phase space region selected with the criteria applied in Sec. 6.1 in the first place, provides the best result in terms of the significance independent of the imposed uncertainties. However, due to the overwhelming  $W$ +jets contribution the systematic uncertainty might be slightly underestimated by the assumed 30 % in SR2 and especially in SR1. Generally, the differences between SR1 and SR2 are very small in terms of both significance and  $t\bar{t}$ -to- $t\bar{t}Z$  ratio, where the latter amounts to approximately 78 %. In contrast, the differences between the two exclusive signal regions are much more apparent. Whereas SR3 delivers a worse reduction of the  $t\bar{t}$ -to- $t\bar{t}Z$  ratio with respect to the theoretical value, SR4 outperforms not only its exclusive counterpart but also the two inclusive signal regions with a value of around 80 %.

## 6.5. Reconstruction results

The two top quarks and the  $Z$  boson have been reconstructed following the instruction given at the bottom of Sec. 6.3.1, resulting in the distributions depicted in Fig. 6.7–6.10. As can be seen in Fig. 6.7 from the overlap of the orange area depicting  $W$ +jets and the red signal line, the  $W$ +jets background exceeds the  $t\bar{t}Z$  signal much more than in the other signal regions, as already mentioned above. The shapes of the reconstructed top quark masses are fairly broad compared to those of the  $Z$  boson and exhibit a long tail to the right, reflecting the large number of possible jet-parton assignments within the reconstruction procedure. The same holds for the combined mass distributions of both top quarks, which naturally receive the properties from each of the top quark mass distributions. Since on the  $W$  and  $Z$  boson masses quite tight constraints are set (cf. Sec. 6.3.1), the invariant mass distributions are therefore more narrow, except for the leptonically decaying  $W$  boson. There, a considerable uncertainty from the a priori unknown  $z$ -component of the neutrino momentum enters the reconstruction, leading to smearing of the invariant mass distribution. All  $Z$  mass distributions exhibit sort of a flat basis starting at about 20 GeV, referred to as combinatorial background, on which the actual  $Z$  mass peak is placed ranging from roughly 50 GeV to 150 GeV. This feature will be discussed in more detail later on in Sec. 6.6.

As it is obvious that even after the kinematic reconstruction of the  $t\bar{t}Z$  system still no cut-based methods are suited to isolate the signal from backgrounds, a more sophisticated technique has to be used, as explained in Chap. 7.

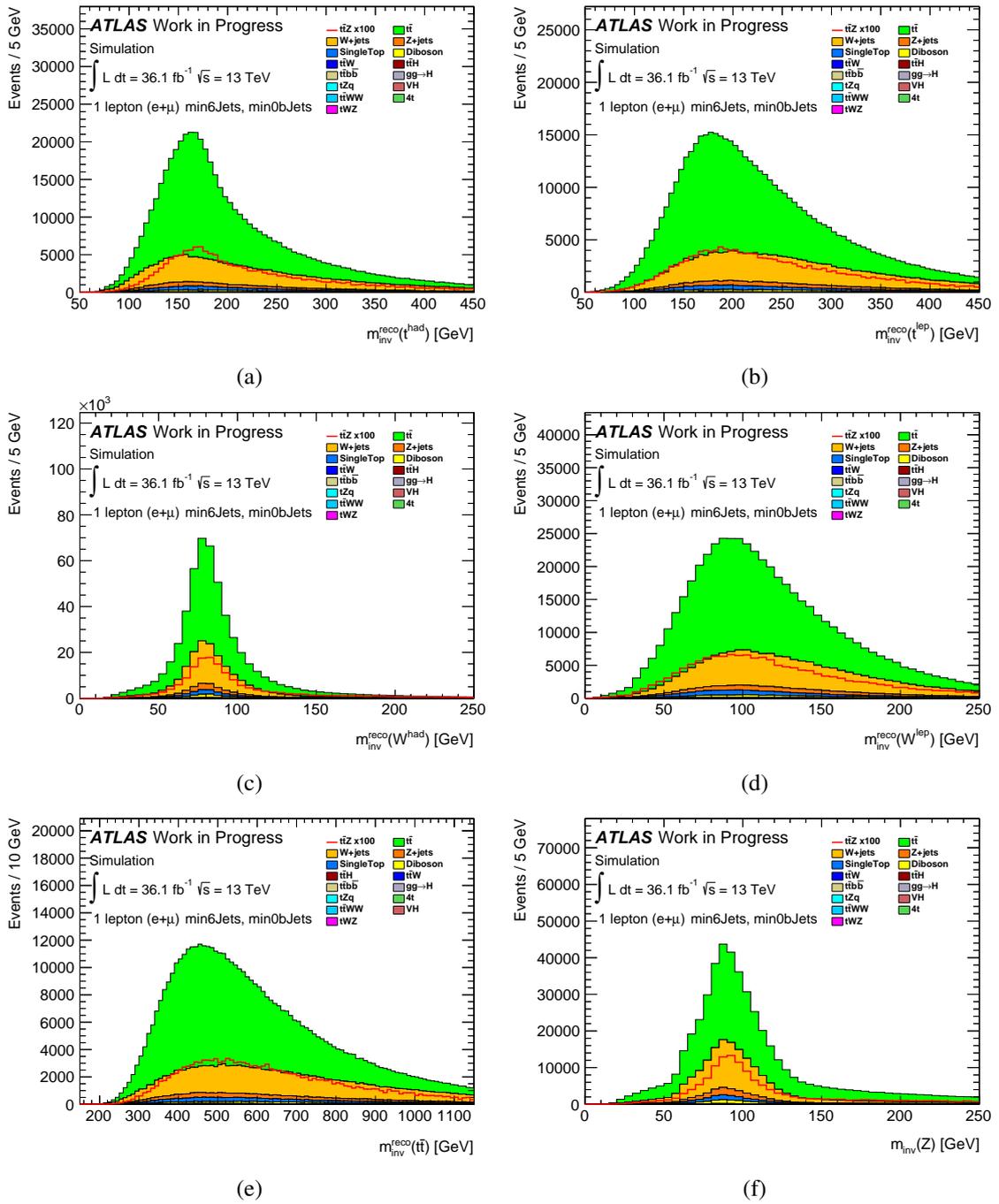


Fig. 6.7.: Reconstructed mass distributions shown for the first signal region (SR1). In Fig. (a) and Fig. (b) the reconstructed mass of the top quark from the hadronic and the leptonic decay chain is depicted, whereas Fig. (c) and Fig. (d) show the reconstructed mass distributions of the hadronically and leptonically  $W$  boson, respectively. Fig. (e) presents the combined mass distribution of both top quarks, and finally Fig. (f) illustrate the invariant mass of the reconstructed  $Z$  boson.

## 6. Event selection and reconstruction

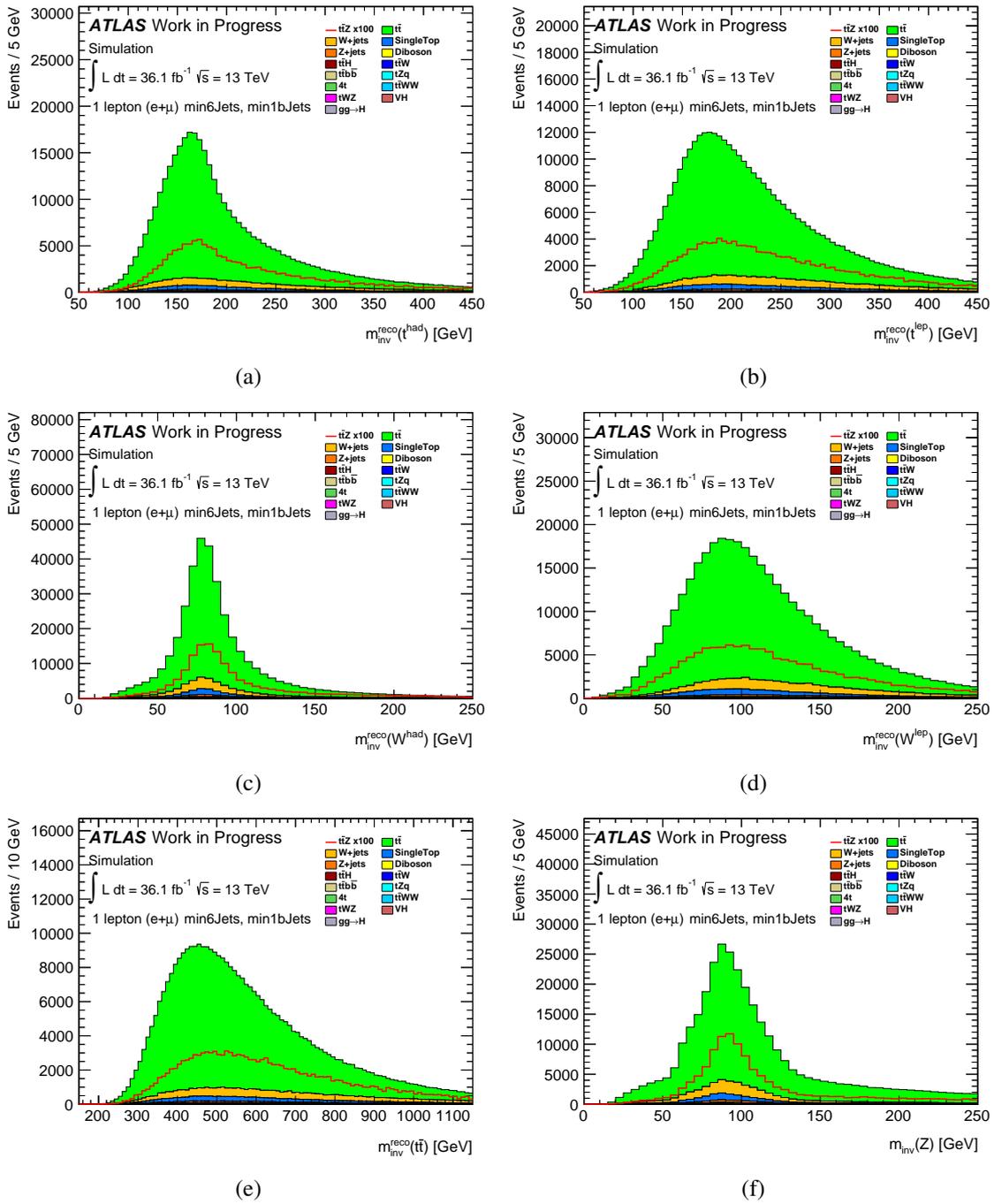


Fig. 6.8.: Reconstructed mass distributions shown for the second signal region (SR2). In Fig. (a) and Fig. (b) the reconstructed mass of the top quark from the hadronic and the leptonic decay chain is depicted, whereas Fig. (c) and Fig. (d) show the reconstructed mass distributions of the hadronically and leptonically  $W$  boson, respectively. Fig. (e) presents the combined mass distribution of both top quarks, and finally Fig. (f) illustrate the invariant mass of the reconstructed  $Z$  boson.

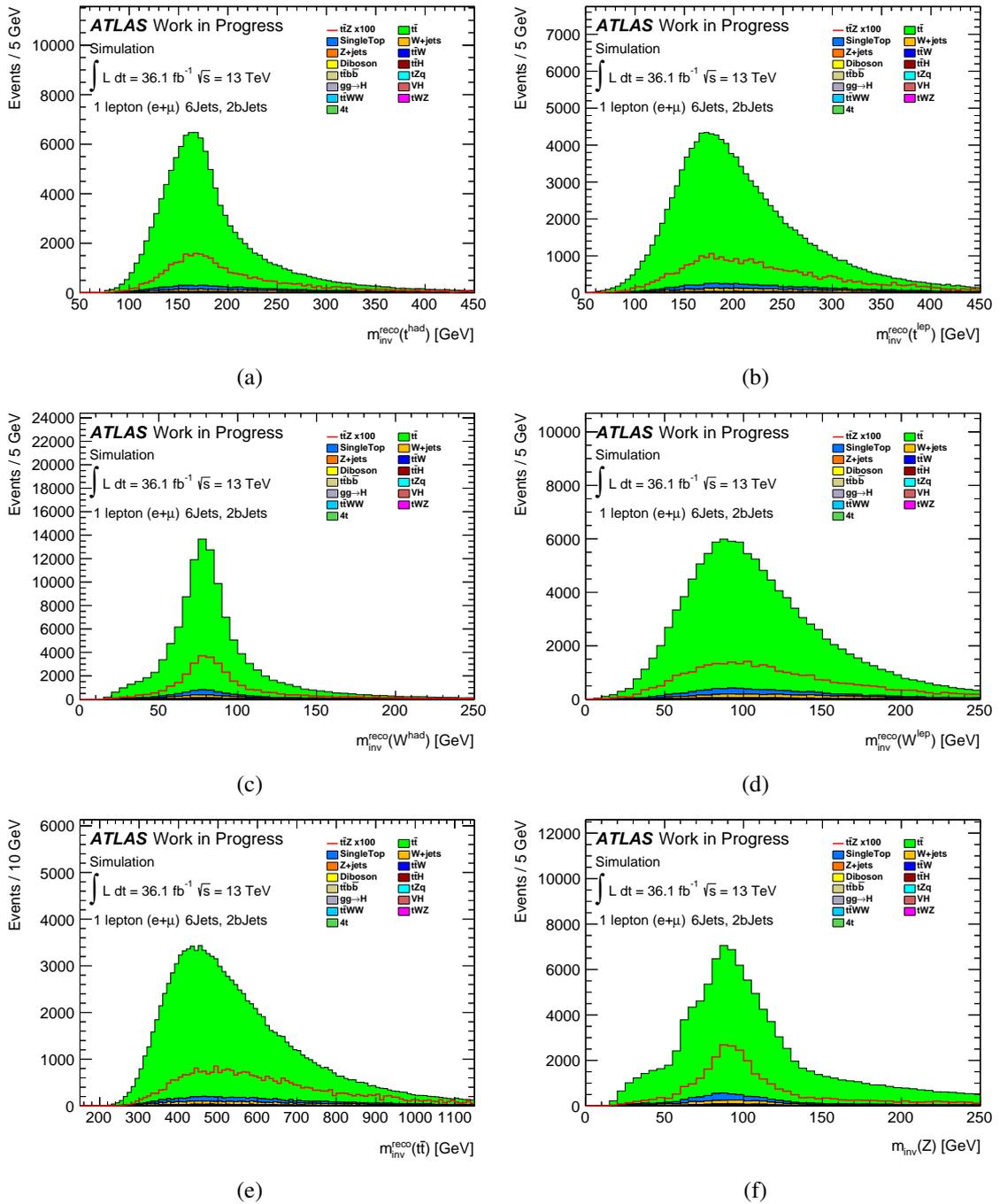


Fig. 6.9.: Reconstructed mass distributions shown for the third signal region (SR3). In Fig. (a) and Fig. (b) the reconstructed mass of the top quark from the hadronic and the leptonic decay chain is depicted, whereas Fig. (c) and Fig. (d) show the reconstructed mass distributions of the hadronically and leptonically  $W$  boson, respectively. Fig. (e) presents the combined mass distribution of both top quarks, and finally Fig. (f) illustrate the invariant mass of the reconstructed  $Z$  boson.

## 6. Event selection and reconstruction

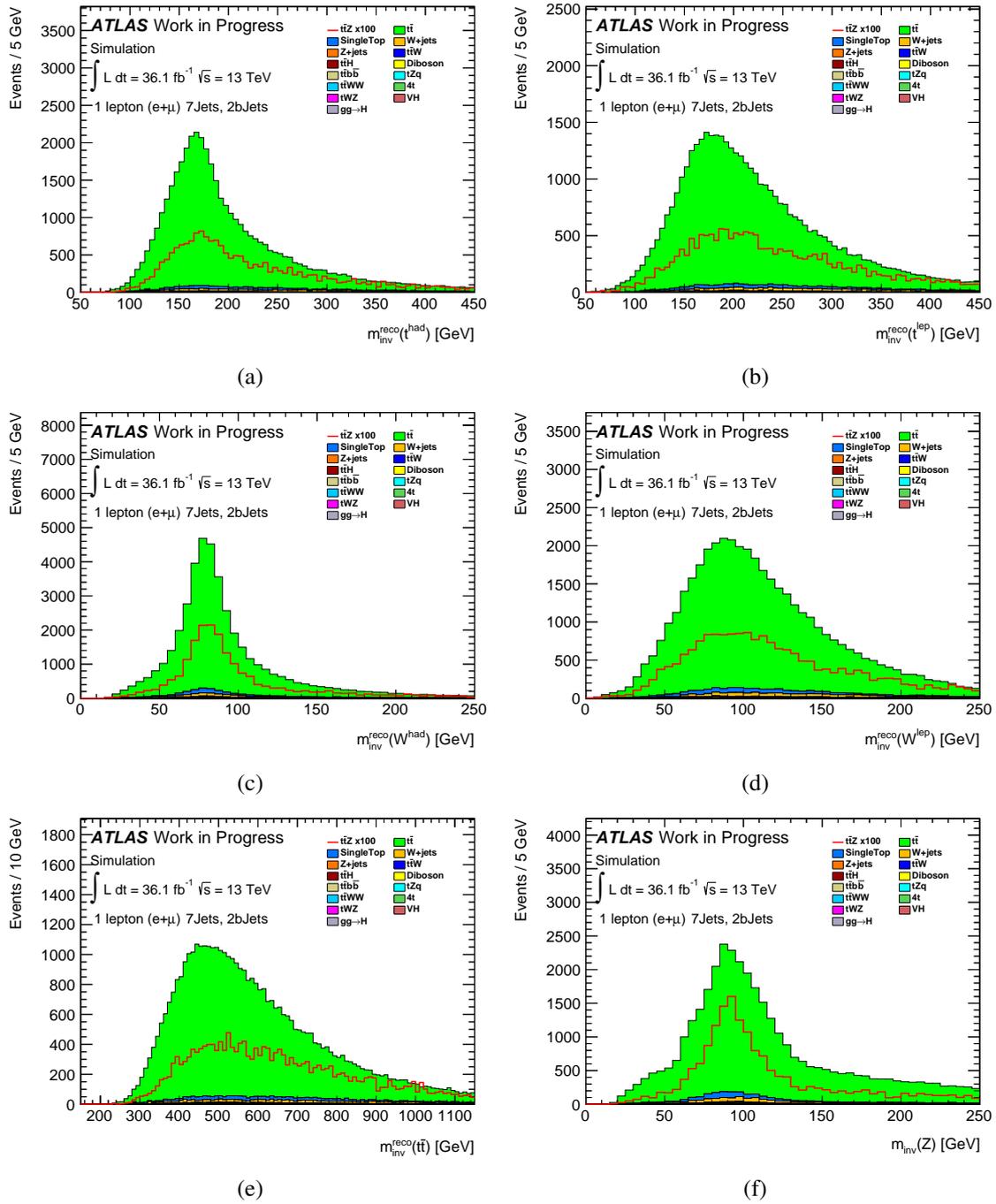


Fig. 6.10.: Reconstructed mass distributions shown for the fourth signal region (SR4). In Fig. (a) and Fig. (b) the reconstructed mass of the top quark from the hadronic and the leptonic decay chain is depicted, whereas Fig. (c) and Fig. (d) show the reconstructed mass distributions of the hadronically and leptonically  $W$  boson, respectively. Fig. (e) presents the combined mass distribution of both top quarks, and finally Fig. (f) illustrate the invariant mass of the reconstructed  $Z$  boson.

## 6.6. Reconstruction performance

It has been shown in Sec. 6.3 that the number of possible jet-parton assignments in the reconstruction of the  $t\bar{t}Z$  system depends on the  $b$ -jet multiplicity of a given event. With an increasing number of possible permutations in an event an incorrect assignment becomes more likely. In order to confirm that the reconstruction works sufficiently well, additional studies on its results have been performed.

### 6.6.1. Investigation of the $Z$ boson mass

As stated in Sec. 6.5, the distributions of the reconstructed  $Z$  boson mass exhibit a quite broad plateau on which the actual mass peak is put on. This plateau, so far referred to as combinatorial background, has been further investigated. As the signal region for this analysis had been found by scanning through a grid of different ( $b$ -)jet multiplicities, the same has been done for the distribution of the  $Z$  boson mass. The results are shown in Fig. 6.11. It can be clearly seen in Fig. 6.11a

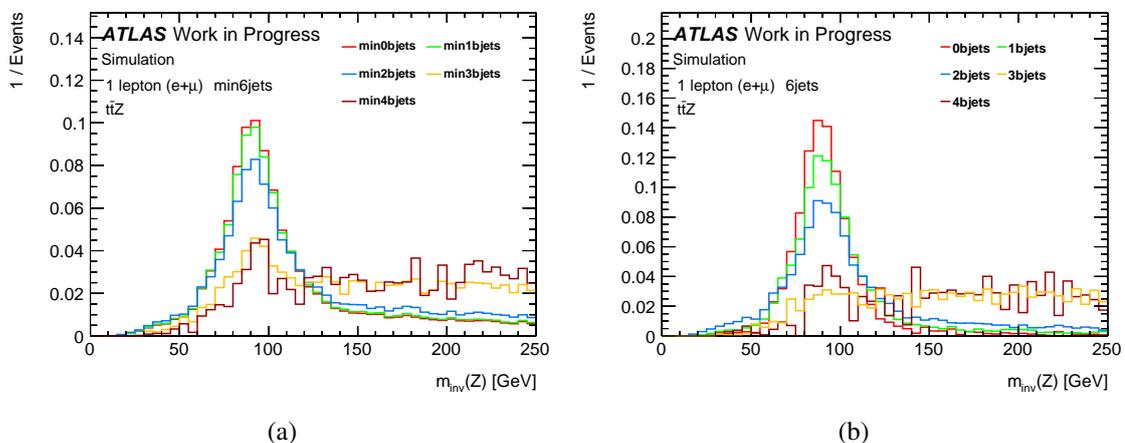


Fig. 6.11.: Illustration of the  $Z$  boson mass distribution depicted for different ( $b$ -)jet multiplicities. Fig. (a) depicts the  $Z$  boson mass as a function of the number of  $b$ -jets for at least jets in an event, whereas Fig. (b) shows the situation for exactly 6 jets. All distributions are normalised to total number of events in order to compare their shapes.

and 6.11b that the high flat tail of the mass distribution predominantly develops for events with a high  $b$ -multiplicity, whereas the actual mass peak rather arises in phase space regions with a small number of  $b$ -jets for which both exclusive and inclusive selection criteria have been applied. Since most of the events have rather small  $b$ -jet multiplicity (cf. Fig 6.3d), the peak is still well visible within the regions including (almost) all events.

However, the reconstruction seems to fail for events containing many  $b$ -tagged jets, albeit the increase of the  $b$ -jet multiplicity should actually lower the number of possible jet-parton assignments in an event, thus reducing the probability of mis-assignments.

### 6.6.2. Investigation of the KLFilter likelihood

During the reconstruction procedure a likelihood (cf. Eq. (6.8)) is maximised probing every allowed jet-parton assignment in an event. Hence, the distribution of the KLFilter likelihood offers information about the performance of the reconstruction. Besides the peak on the far right, the distribution of the logarithmic likelihood from KLFilter (cf. Fig. 6.5) shows a second broader peak,

which has not been expected. Therefore, as for the  $Z$  boson mass, the KL Fitter likelihood has been logarithmically plotted for different ( $b$ -)jet multiplicities as shown in Fig. 6.12. Splitting the

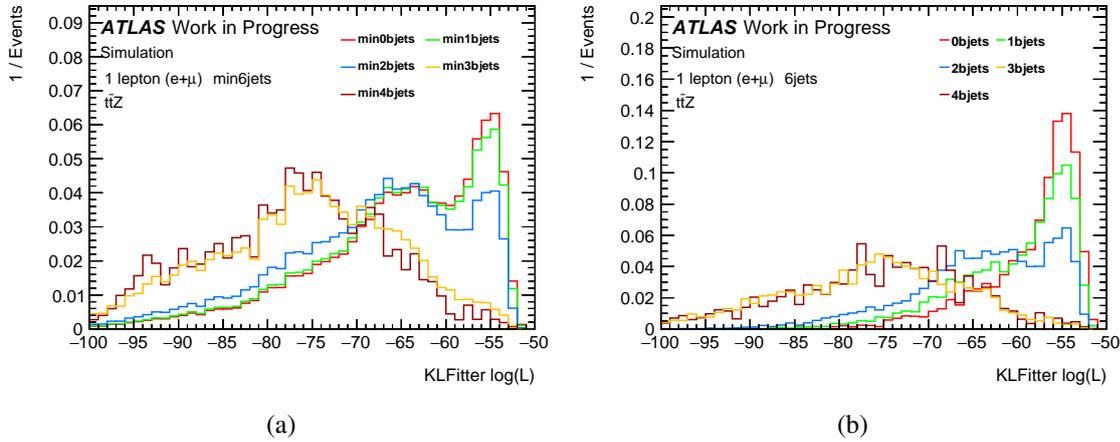


Fig. 6.12.: Illustration of the logarithmic KL Fitter likelihood depicted for different ( $b$ -)jet multiplicities. Fig. (a) depicts the KL Fitter likelihood as a function of the number of  $b$ -jets for at least jets in an event, whereas Fig. (b) shows the situation for exactly 6 jets. All distributions are normalised to total number of events in order to compare their shapes.

phase space in inclusive regions (cf. Fig 6.12a), a third peak arises on the left for events containing more than 2  $b$ -jets, whereas the second peak is visible for events containing at least one  $b$ -jet. As depicted in Fig. 6.12b, the second peak is mainly caused by events with a  $b$ -jet multiplicity of 2 and the third peak only develops for more than 2  $b$ -jets in an event, meaning that the performance of the reconstruction algorithm significantly decreases.

The KL Fitter performance seems thus to be highest for events containing at most 1  $b$ -jet and becomes smaller for an increasing  $b$ -jet multiplicity. This is a remarkable behaviour because for a  $t\bar{t}Z$  event at least 2  $b$ -jets have to be present. Furthermore, the attempt of explicitly reconstructing a  $Z$  boson decaying to  $b\bar{b}$  in the case of more than 2  $b$ -jets in an event further lowers the reconstruction performance.

### 6.6.3. Investigation of the KL Fitter jet indices

Besides the likelihood distribution, KL Fitter also returns an index labelling the jet-parton assignment of the best permutation for each event. According to this labels, the initial state particles are reconstructed. Within the (Monte Carlo) datasets the jets are stored in a descending order in transverse momentum. In this analysis, the first six jets, i.e. the six jets with the highest transverse momentum, are put into the reconstruction algorithm. An uniform labelling has been chosen where the labels correspond to the storage positions of the jets within the Monte Carlo datasets. Each jet which is finally assigned to a final state parton receives a particular label according to the assignment the reconstruction algorithm indicates to be the best for a given event. Thus, the jet index provided by KL Fitter indicates which jet from the Monte Carlo sample is finally assigned to which final state parton in a given event.

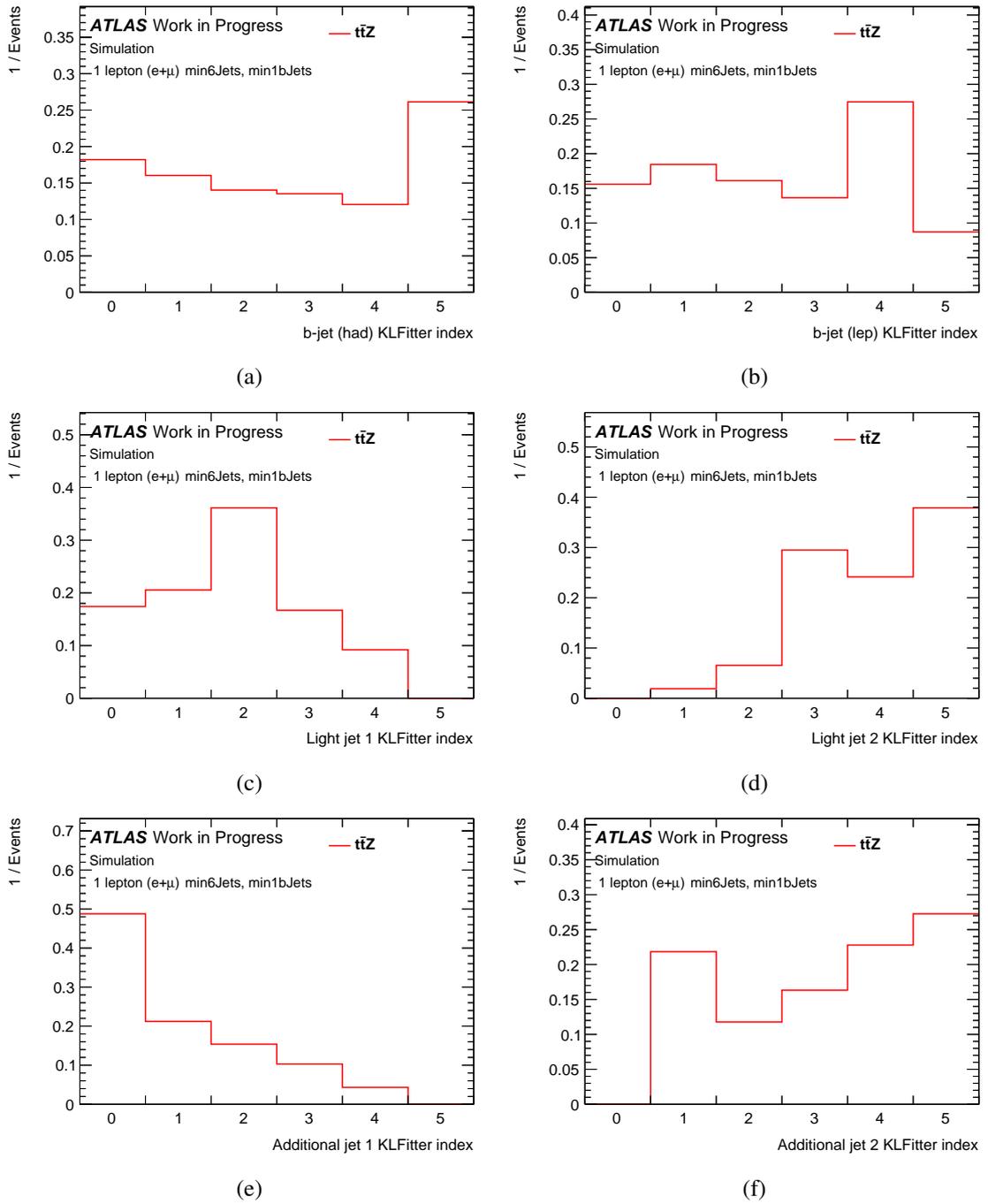


Fig. 6.13.: Illustration of the KL Fitter jet indices for all final state particles. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the indices of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

The assignment results from the reconstruction for SR2 are shown in Fig. 6.13 depicted for each parton position. The corresponding images for the other signal regions are given in Fig. E.3 – E.5

It can be seen that for inclusive selection criteria the two jets identified as originating from the  $b$ -quarks associated to top quark decay are in many cases the jets with the lowest provided momenta. In contrast, one of the jets associated to the  $Z$  boson is for about 50 % of the events the one with the highest momentum. The other jet supposed to stem from  $Z$  boson decay is labelled to either carry also high momentum for roughly 25 % of all events, or tends towards lower momenta quite often. The two jets associated to the  $W$  boson are mainly assumed either to have a quite small momentum or are mainly set on position three in the middle. The two  $b$ -jets from top quark decay have a slight tendency towards the high-momentum regime, while the jets from the  $W$  and the  $Z$  boson decay depicted on the right side of the images only lean to smaller momenta anymore. The picture for the two remaining jets does not heavily change.

Parton	Label
Up quark	1
Down quark	2
Strange quark	3
Charm quark	4
Bottom quark	5
Top quark	6
Gluon	21
Jets from pile-up	-1

Tab. 6.7.: Depiction of the jet flavour labels. For the labelling the flavour of the hardest parton within the jet is taken as in jet reconstruction jets are most likely formed around the particles with the highest transverse momentum in a certain region of space. The choice of numbers is based on the particle labels as stated by the Review of Particle Physics [34].

In order to examine the performance of the reconstruction it has been checked whether the flavour of the parton the jet originates from is in agreement with the flavour of the quark on which position it has been set by the algorithm. The particular quarks are labelled as given in Tab. 6.7. Within this comparison it does not matter at all if the quark is actually a particle or an antiparticle. Of course, this works only for Monte Carlo simulation as in real data the flavour of the initiating parton of jet is hardly accessible. As for the  $Z$  boson mass and the KL Fitter likelihood, it has been done for different phase space regions defined by the ( $b$ -)jet multiplicity. The results are shown in Fig. 6.14 depicting the number of jets which the KL Fitter algorithm correctly set on final state parton positions according to their flavour. Since for inclusive selection criteria (cf. Fig. 6.14a) not only one dedicated phase space region but more than one are simultaneously taken together, the differences between the different requirements of the ( $b$ -)multiplicity are quite small, although a slight tendency is already visible. A correct assignment in terms of parton flavours can be achieved most frequently for at least four  $b$ -jets in an event followed by a  $b$ -jet multiplicity of two, whereas the worst performance occurs imposing no selection criterion on the number of  $b$ -jets. In Fig. 6.14b the situation for exclusive requirements is depicted. There the assignment performance according to quark flavours differs quite intensively, depending on the region in phase space. However, the performance is best for four  $b$ -jets and second best for two  $b$ -jets. Similarly, the lowest performance is provided for one or even no  $b$ -jets. Thus, the algorithm works best for events theoretically predicted as expected. However, events containing three  $b$ -jets are kept since in this analysis a fixed cut on the  $b$ -tagging BDT output serves as a decision criterion whether a jet originates from a  $b$ -quark or not (cf. Sec. 5.5), thus a small percentage of real  $b$ -jets are not tagged to be such and vice versa. Thus, Fig. 6.14a and 6.14b indicate that for almost no event all jets are

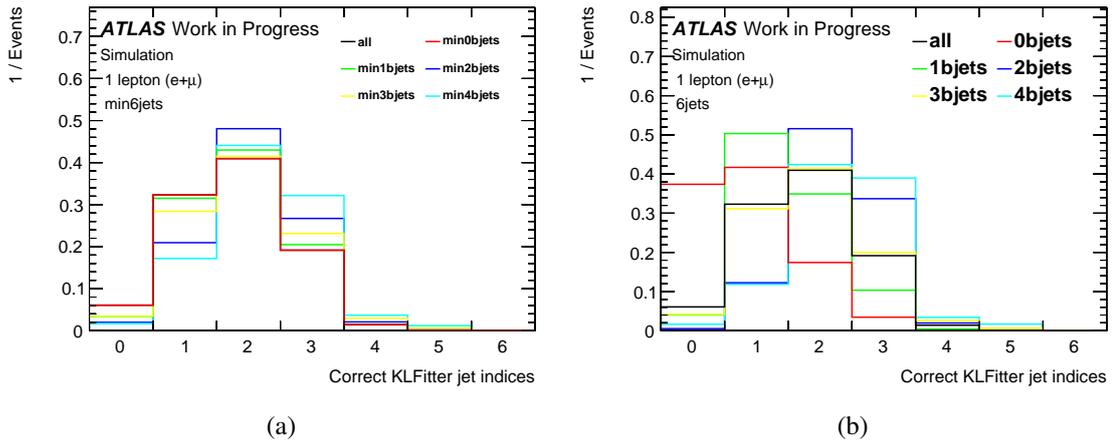


Fig. 6.14.: Illustration of the KLFitter jet indices depicted for different ( $b$ -)jet multiplicities. Fig. (a) depicts the jet indices as a function of the number of  $b$ -jets for at least jets in an event, whereas Fig. (b) shows the situation for exactly 6 jets.

correctly assigned with respect to their flavour.

Looking at the jet flavour labels of SR2 depicted in Fig. 6.15, the numerous mis-assignments become quite clear. Besides the two  $b$ -quarks originating from top quark decay on which positions mainly  $b$ -tagged jets have been set, all other positions are occupied to roughly a quarter by jets stemming from gluons which is for the  $b$ -quarks to only 10% the case. As more than two  $b$ -jets are allowed for inclusive selection criteria, the positions of the two quarks from  $W$  and  $Z$  boson decay, respectively, are more often manned by  $b$ -jets than for exclusive criteria, naturally. These four positions are furthermore occupied by jets labelled to arise from pile-up events by 2–5%. Differently to the quark positions of the  $Z$  boson,  $b$ -tagged jets should not have been assigned to quarks from the hadronically decaying  $W$  boson at all as actually a veto for putting  $b$ -jets on light quarks positions is applied in the reconstruction. Thus, summing up all the various contributions shows that each quark position is only to approximately 30% correctly occupied by a jet with respect to the quark flavour. For completion, in Fig. E.6 – Fig. E.8 the corresponding distributions of the jet flavour label are shown for the other signal regions.

In Fig. 6.16 and 6.17 the dependence of the jet flavour labels is depicted for inclusive and exclusive selection criteria in jet phase space, respectively. As it can be observed, the more  $b$ -jets in are in an event, the more often they are set on positions of quarks from vector boson decay, whereas the probability of finding a  $b$ -tagged jet on a  $b$ -quark position simultaneously decreases. However, as in the reconstruction for more than two  $b$ -jets the  $Z$  boson is supposed to explicitly decay into a pair of bottom quarks, this observation seems not to be surprising at least for the two jets originating from the  $Z$  boson. In case of more than two  $b$ -tagged jets in an event the reconstruction algorithm tends to assign them rather to positions of quark from vector boson decay than to the  $b$ -quarks from the decay of the top quarks. Especially events containing three  $b$ -jets provide a rather bad agreement of jet positions and their corresponding parton flavours. Generally, the comparison of the positions jets have been set on with their flavour label reflects the distribution of the number of correct assignments.

#### 6.6.4. Summary

In order to investigate the performance of the KLFitter algorithm several properties related to the reconstruction have been tested. For the distribution of the invariant mass of the  $Z$  boson as well

as the logarithmic likelihood, dependent on the particular phase space region, rather strongly varying shapes have been observed. Furthermore, the jet-to-parton assignment as performed by the reconstruction has been investigated, revealing a mostly insufficient ability of the reconstruction algorithm to correctly relate the jets to the final state quark they actually belong to. In particular, although the veto on placing  $b$ -jets to light quark positions, a considerably high amount of  $b$ -tagged jets is assigned to the two quarks from  $W$  boson decay. Therefore, further investigation is inevitable in order to tackle these chinks. However, when making statements about the performance of the reconstruction algorithm it has to be taken into account that the examination of a correct jet-to-parton assignment has been done only in terms of quark flavours. No information about the actual relation of the jets to the initial state particle is given wherefore it cannot be stated what jet exactly corresponds to which final state quark. A possible improvement of the reconstruction is to not only consider kinematic quantities as the transverse momentum or the energy of a particle, but also to take angular variables into account.

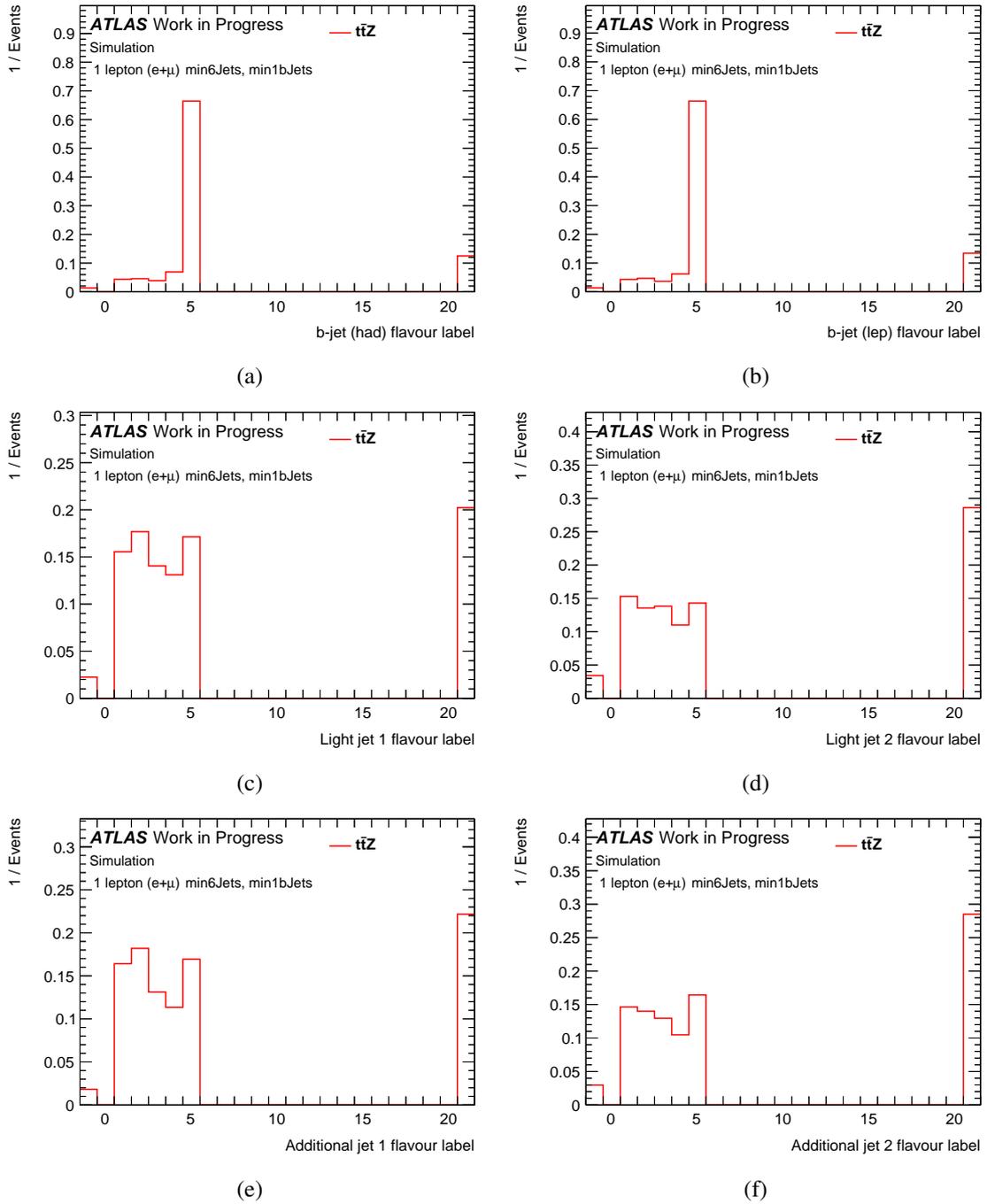


Fig. 6.15.: Illustration of the jet flavour labels for all final state particles. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

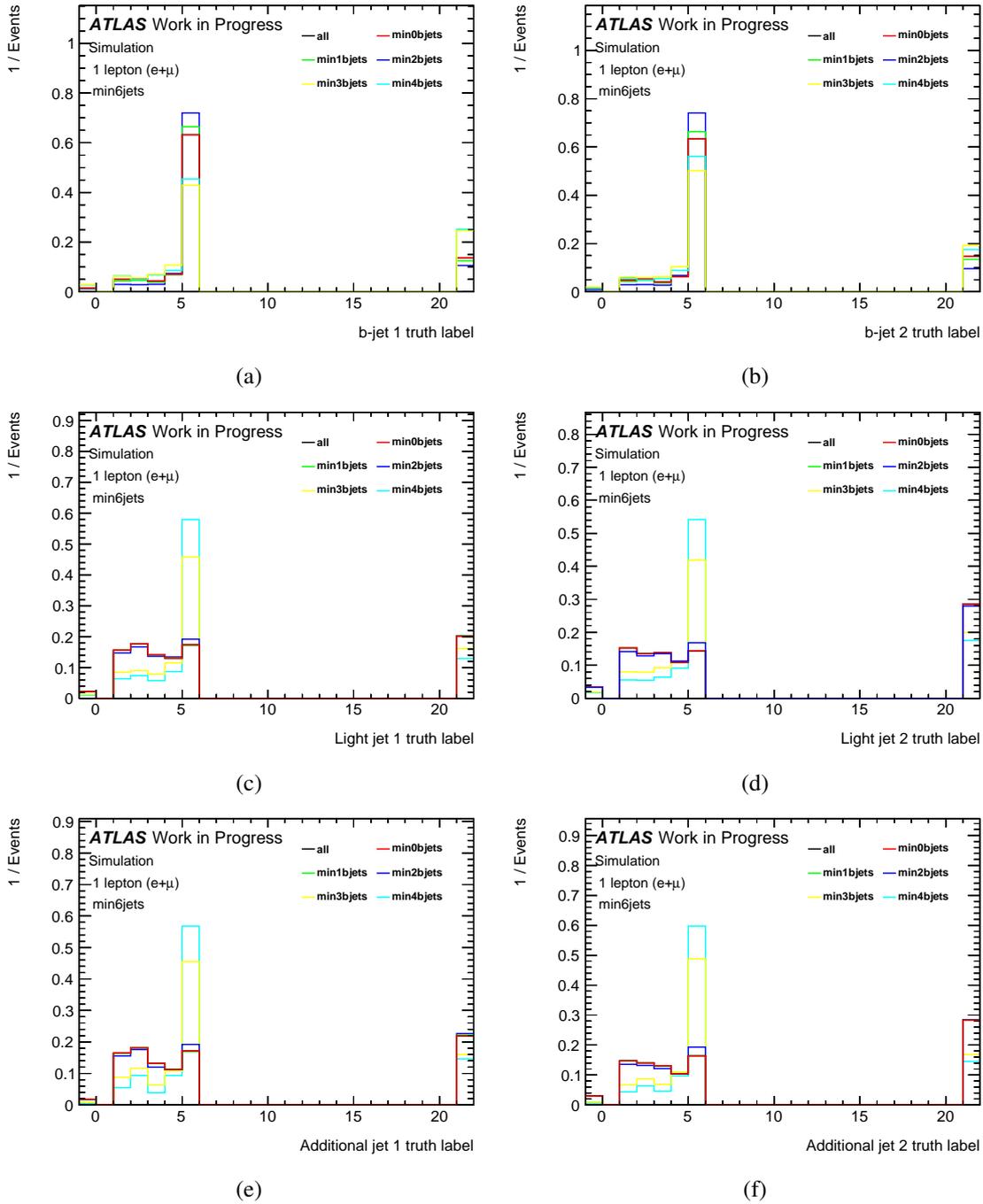


Fig. 6.16.: Illustration of the jet flavour labels for all final state particles in the region of phase space with at least six jets as a function of various requirements on the  $b$ -jet multiplicity. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

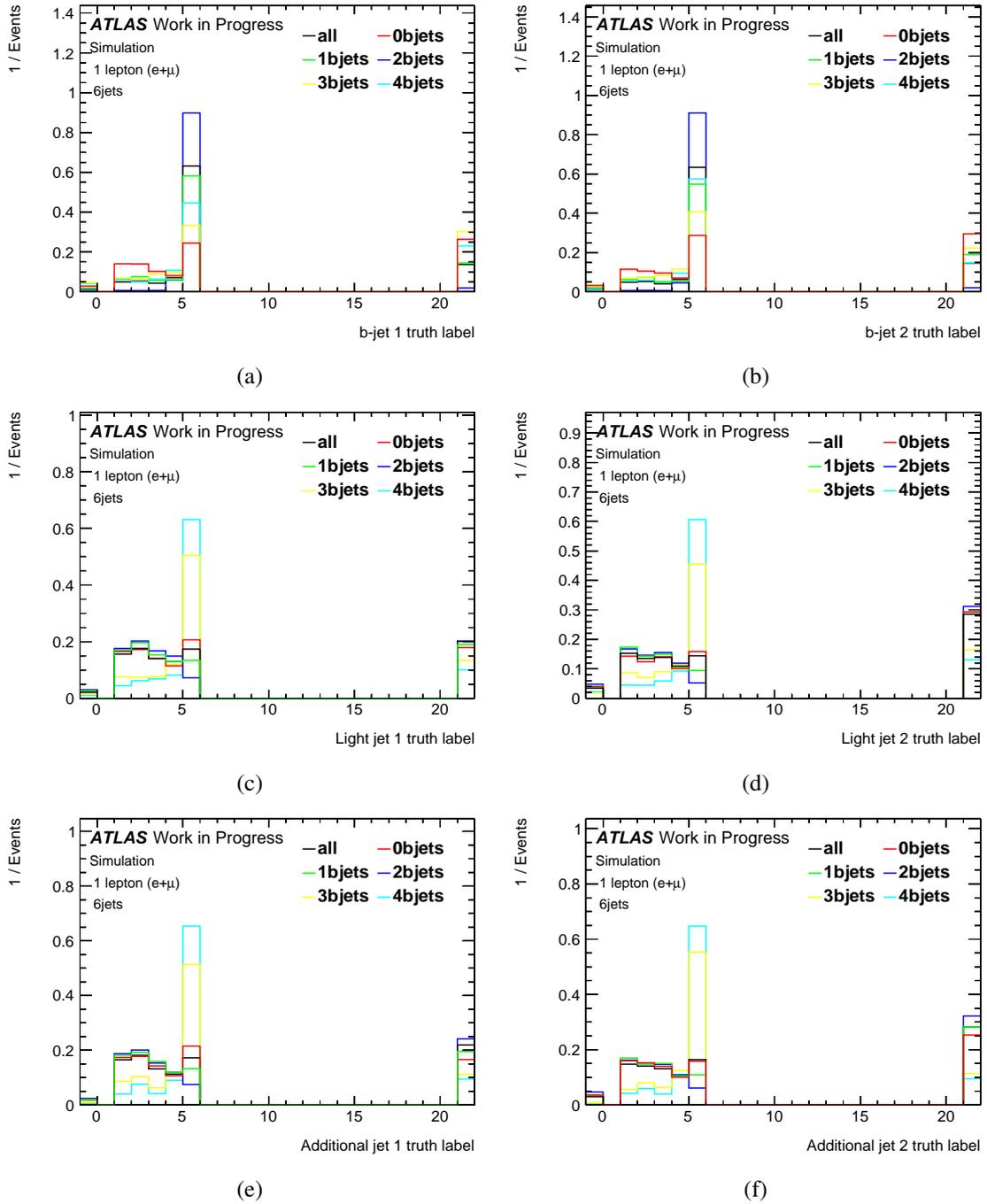


Fig. 6.17.: Illustration of the jet flavour labels for all final state particles in the region of phase space with exactly six jets as a function of various requirements on the  $b$ -jet multiplicity. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.



## 7. Multivariate Analysis

In order to separate between signal and background events, simple cuts on different variables can be used if the shape of these distributions differ reasonably. If this is not the case, simple cuts might not be the best solution anymore. Instead, it is more favourable to use multivariate techniques to separate between signal and background, because it considers both the shapes of the distributions and correlations among them.

### 7.1. Boosted decision tree

Using a multivariate technique means to train a set of discriminating variables with a specific method (algorithm) – as, for example, a boosted decision tree [205]. A classification technique using such a method allows to combine several discriminating input variables into one final output variable (here: BDT response), which is sensitive to all the differences between signal and background incorporated in the input variables. This equals a mapping of the space of  $d$  input variables  $\mathbb{R}^d$  onto a one-dimensional space  $\mathbb{R}$  of the single discriminator. In case of boosted decision trees the output variable is defined in the range of  $[-1, +1]$  in which  $-1$  represents background-like events and  $+1$  signal-like events. In this analysis the Toolkit for Multivariate Analyses (TMVA) [205], which provides a ROOT-integrated [202] environment for processing, evaluation and application of multivariate classification and regression techniques, is used.

The main reason for choosing a boosted decision tree for this analysis is its robustness against the training on statistical fluctuations. This is important for both cases dealing with a rather high or small amount of input data and comes along with its construction, as explained more detailed in the following.

#### 7.1.1. General concepts

A boosted decision tree (BDT) [205] is a binary, tree structured classifier using yes-/no-decisions on one single variable until some specified stop criterion is fulfilled [205]. As illustrated in Fig. 7.1 the training starts at one single root node, where an initial splitting criterion on the whole dataset is determined by cutting on the variable with the highest discrimination power at this stage. Thus, the dataset is split up into two (smaller) subsets of training events and two nodes, which serve as root nodes for the following steps. This procedure is repeated until the events are sufficiently separated (“classified”) into final nodes, also referred to as leaf nodes, which are eventually tagged as signal or background leafs depending on the the majority of events per category.

TMVA provides several methods which can be used to determine the best discriminating variable and its optimum cut value at each node, i.e. the variable and its corresponding cut value, that optimises the separation between the parent node and the two prospective daughter nodes by maximising the method’s associated mathematical expression. The same variable might thus be used at several nodes, while others might not be used at all. For this analysis the most frequently used Gini Index has been chosen, which depends on the purity  $p^{(1)}$ . Tab. 7.1 shows a set of different methods for the calculation of the highest separation power implemented in TMVA. Since a cut

---

<sup>(1)</sup> The purity of a node is given by the ratio of signal events to all events in that node, i.e. pure background nodes have zero purity.

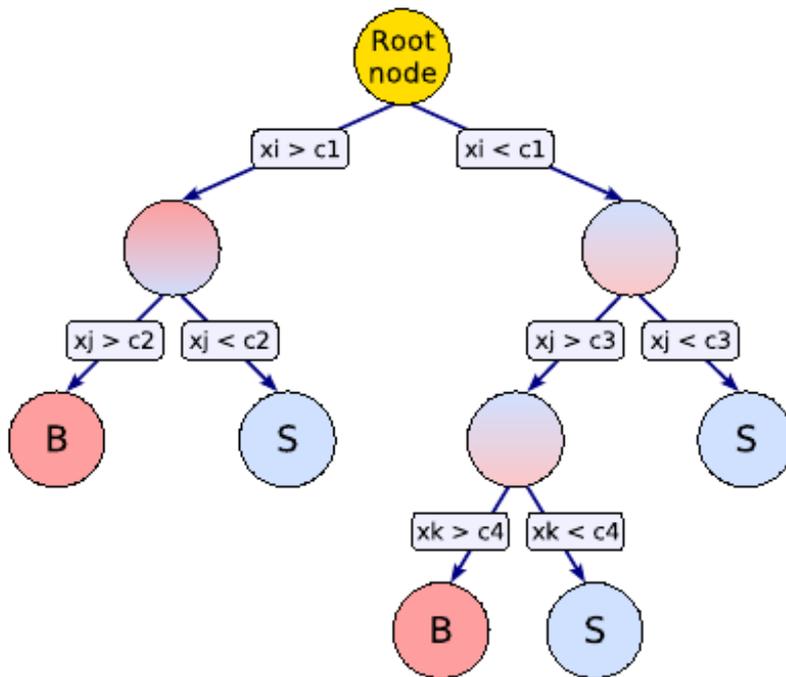


Fig. 7.1.: Illustration of a decision tree. Starting from the root node, a sequence of binary splits is applied to a dataset using the variable  $x_i$  with the highest discrimination power at the particular node. The leaf nodes at the bottom end of the tree are labeled “S” for signal and “B” for background, depending on the majority of events that end up in the respective nodes. [205]

that predominantly selects background is as valuable as a cut that selects signal, the criteria are symmetric with respect to the event classes as depicted is Fig. 7.2.

<i>Gini Index</i>		$p \cdot (1 - p)$
<i>Cross Entropy</i>	$-p \cdot \ln(p) - (1 - p) \cdot \ln(1 - p)$	
<i>Misclassification error</i>		$1 - \max(p, 1 - p)$
<i>Statistical significance</i>		$S/\sqrt{S + B}$

Tab. 7.1.: Selection of separation techniques to calculate the separation power at the split nodes in order to determine the optimum variable on which it has to be cut on. All methods have their maximum at purity  $p = 0.5$  and fall off to zero when the sample consists of one event class only.

The importance or ranking of a variable within a training procedure is proportional to the number of splitting decisions the variable has been used for. In doing so, each split is weighted by the separation gain squared and by the number of events in the node.

A shortcoming of decision trees is their instability with respect to statistical fluctuations within the training events. In order to prevent training on statistical fluctuations due to some cuts, many decision trees, all constructed from the whole training sample, are combined into a forest instead of growing one single decision tree. Furthermore, overtraining, i.e. training on statistical fluctuations of the datasets, is counteracted by limiting the tree depth confining either the maximal number of

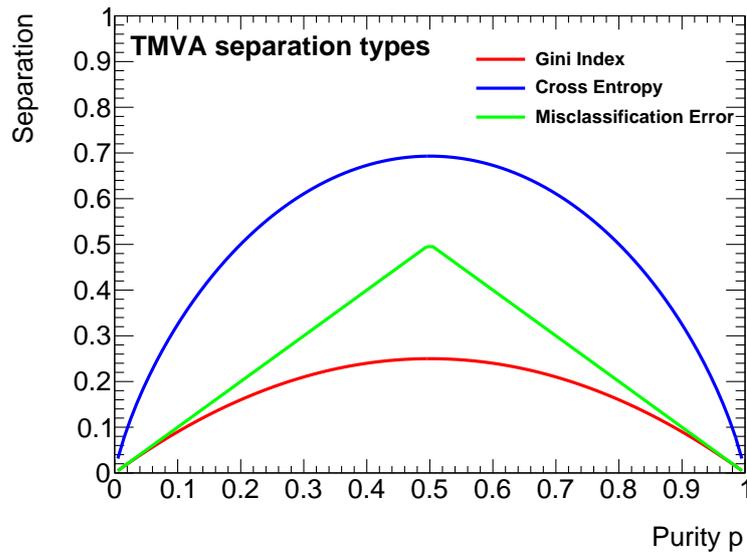


Fig. 7.2.: Comparison of three different methods for the calculation of the highest separation in TMVA, the Gini Index, the Cross Entropy and the Misclassification error. They all only depend on the purity  $p$  of the node at which a splitting decision has to be made. All methods are symmetric with respect to the event classes with their maximum at  $p = 0.5$ . The corresponding formulas are given in Tab. 7.1.

nodes in a tree or the minimal number of events in a leaf node. In addition, the number of grown trees itself can be adjusted.

### 7.1.2. Boosting

A way of enhancing the classification performance and to increase the stability with respect to statistical fluctuations in the training sample is to reweight (“boost”) incorrectly assigned events in the training sample after each training, to end up in the correct category in the newly grown tree. Typically, growing a forest of boosted trees leads to a large improvement of separation performance and stability regarding statistical fluctuations. However, the downside is the loss of clear decision tracing since the number of individual trees is in general too large. A more detailed explanation of the most popular boosting algorithm, the so-called AdaBoost, as well as of Gradient Boost, the boosting method applied in this analysis, can be found in Appendix D.

### 7.1.3. Data preprocessing

Before putting discriminating variables or training events into multivariate methods, it is possible to subject them to a preprocessing. This can be useful in order to reduce correlations among some variables or to bring their shapes in a more appropriate form.

In TMVA currently five preprocessing methods are implemented [205]:

1. Variable normalisation
2. Decorrelation via the square root of the covariance matrix
3. Decorrelation via a principal component decomposition
4. Transformation of the variables into Uniform distributions (“Uniformisation”)
5. Transformation of the variables into Gaussian distributions (“Gaussianisation”)

With the first of them, the minimum and maximum range for the input variables are determined from the training events and used afterwards to linearly scale these variables to lie within a range of  $[-1, +1]$ . Such a transformation turns out to be useful in order to directly compare the classification weights assigned to the variables, where large absolute values may indicate a strong separation power.

Several classifiers, for instance rectangular cuts, underperform in presence of variable correlations, while other classifiers as the projective likelihood classifier even ignore correlations at all, leading to noticeable performance losses in most real scenarios [205]. Linear correlations can be easily taken into account by computing the square root of the covariance matrix, i.e. by finding the matrix  $C'$  that multiplied with itself yields  $C$ :  $C = (C')^2$  [206]. This can be achieved by means of diagonalising the (orthogonal) covariance matrix  $D$ ,

$$D = S^T C S \quad \stackrel{(C')^2 \equiv C}{\implies} \quad C' = S \sqrt{D} S^T \quad , \quad (7.1)$$

where  $D$  is a diagonal matrix and  $S$  orthogonal. With  $\sqrt{D}$  the square root of all entries is meant, i.e.  $\sqrt{D} = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_n}) \in \mathbb{R}^{n \times n}$ . The linear decorrelation of the selected variables is then obtained by multiplying the initial variable tuple  $\mathbf{x}$  by the inverse of  $C'$ ,

$$\mathbf{x} \mapsto (C')^{-1} \mathbf{x} \quad . \quad (7.2)$$

It is important to note that the decorrelation as sketched above only works properly for linearly correlated and Gaussian distributed variables [205].

A rather similar preprocessing method is the principal component decomposition or principal component analysis (PCA) [207, 208]. It is a linear transformation that rotates a sample of data points such that the maximum variability is visible, thus identifying the most important gradients. Within the PCA-transformed coordinate system, the largest variance irrespective of a particular projection of the data is identified with the first coordinate, hence denoted by the first principal component. Going on, the second largest variance comes to lie on the second coordinate, et cetera. Consequently, with PCA it is possible to reduce the dimensionality of a problem initially given by the number of input variables by simply removing dimensions with an insignificant variance. The tuples

$$\mathbf{x}_U^{\text{PC}}(i) = (x_{U,1}^{\text{PC}}(i), \dots, x_{U,N_{\text{var}}}^{\text{PC}}(i)) \quad (7.3)$$

of principal components of a tuple of input variables  $\mathbf{x}(i) = (x_1(i), \dots, x_{N_{\text{var}}}(i))$ , measured for the event  $i$  belonging to either signal ( $U = S$ ) or background ( $U = B$ ), are obtained by the transformation

$$x_{U,k}^{\text{PC}}(i) = \sum_{l=1}^{N_{\text{var}}} (x_{U,l}(i) - \bar{x}_{U,l}) v_{U,l}^{(k)}, \quad \forall k = 1, \dots, N_{\text{var}} \quad , \quad (7.4)$$

with the tuples  $\bar{x}_U$  and  $\mathbf{v}_U^{(k)}$  being the sample means and the eigenvectors, respectively. The matrix of eigenvectors  $V_U = (\mathbf{v}_U^{(1)}, \dots, \mathbf{v}_U^{(N_{\text{var}})})$  then obeys the relation

$$C_U \cdot V_U = D_U \cdot V_U \quad , \quad (7.5)$$

where  $C$  again is the covariance matrix of the sample  $U$  and  $D_U$  is the diagonal matrix of eigenvalues. With this transformation linear correlations for Gaussian distributed variables can be eliminated [205].

Since PCA is indeed not very different from the linear decorrelation discussed above, both

methods have been introduced. However, in the analysis only PCA has been used for handling correlated variables. At the moment, in fact only two methods can be applied for both the training and the evaluation of a fully trained and tested BDT in TMVA at all, namely the normalisation method and the principal component decomposition. The first one does not fully meet the requirements as in this stage of the analysis special attention is not placed on a dedicated investigation of the single variables's performance but rather on probing the BDT for the  $t\bar{t}Z$  1-lepton channel on the whole. Therefore merely the PCA preprocessing has been used in order to improve the BDT's performance.

#### 7.1.4. Performance

BDTs are widely used in high energy physics, for example for the reconstruction of physical objects at the ATLAS detector (cf. Chap. 5) or as analysis tool for various analyses. Since not much tuning is required in order to obtain reasonable results, they are considered as the best “out of the box” classifiers. This gets favoured by the simplicity of the methods where each training step, i.e. node splitting involves only a one-dimensional cut optimisation. Another advantage of decision trees is their insensitivity to poorly discriminating variables. For each node splitting only the best discriminating variables is used so that non-discriminating variables are basically ignored by the decision tree training algorithm. In addition, a BDT does not need to have as many training events as other multivariate techniques such as artificial neural networks in order to provide good results, especially in case of quite complex correlations among the input variables [205].

To sum up, compared to a cut-based analysis, multivariate algorithms like the BDT have the advantage of selecting hyperplanes in variable space which can be identified as signal-like events. The subsequent signal and background event classification is based on a majority vote, taking into account all individual decision trees which have been trained. The final BDT can be visualised as a forest of them with each having different node properties, one exemplary shown in Fig. 7.1. The events's output weight, called BDT response, ranging from  $-1$  to  $+1$ , can then be interpreted as a new variable that can be subjected to a final analysis step.

## 7.2. BDT analysis

As stated at the end of Sec. 6.5, the previously presented results indicate that only poor separation of signal and background can be achieved by conventional means, for example cut&count methods. Therefore, more sophisticated analysis techniques are explored. In this analysis a boosted decision tree is trained for each signal region separately to provide an optimally separating variable. A first attempt has been done trying to separate the signal from the dominant  $t\bar{t}$  background.

### 7.2.1. BDT configuration

In order to improve the performance of the BDT classifier for a given problem, multiple configuration options have to be tuned. Due to its robustness, in this analysis the Gradient Boost method has been applied on bagged sample fractions, therefore using stochastic gradient boosting to enhance the stabilisation of the BDT response by smearing over statistical fluctuations in the training samples (cf. Sec. D.2). To further enhance the algorithm's robustness a shrinkage parameter different from one has been introduced raising the classifier's accuracy by slightly modifying the weights of the individual trees (cf. Sec. D.2). As separation type the mostly used Gini Index has been chosen. To avoid overtraining on statistical fluctuations in the input datasets, the depth of the decision trees is limited by restricting both the maximal number of nodes in a tree as well as the minimal allowed number of events left in a leaf node. In doing so it is important to walk the

fine line between not fully tapping the full potential of the BDT classifier and overtraining, each leading to significant lack of performance. Next, the number of trees to be trained applying the boosting procedure had to be set, where a reasonable tradeoff had to be found between a sufficiently high mixing rate of the input events and an overshoot in computing time. The input events are randomly mixed before being put into the BDT classifier in order to prevent from being biased by the order in which the events are stored in the training samples. The input events are split into two fractions, one for the training of the multivariate classifier and the other one for testing its outcome, resulting in four different event fractions in total as this procedure is done for both signal and background events, naturally. In order to provide a maximally unbiased training, the events are again randomly picked for testing and training. If the number of signal and background events are not equal, the event numbers are adjusted to each other since the BDT classifier needs to have the same amount of signal and background events, respectively. The cut values at the particular nodes are optimised by scanning over the variable range with a granularity that has to be given to the algorithm, a compromise between computing time and step size has to be made. Furthermore, during the training negative event weights, which can occur in simulated datasets generated by some specific Monte Carlo generators, e.g. MADGRAPH5\_AMC@NLO [123], are ignored because otherwise the training might reveal peculiar results due to reweighting issues that might occur during the boosting procedure. The configuration parameters which turned out to suit best, are listed in Tab. 7.2.

BDT parameter	Optimal value
Boosting algorithm	Gradient Boost
Shrinkage parameter	0.1
Bagged sample fraction	0.5
Separation type	Gini Index
Maximal tree depth	3
Minimal node size	2.5 %
Number of trees	400
Cut granularity	20 bins

Tab. 7.2.: Summary of the BDT configuration parameters which turned out to suit best. Only those parameters are listed that have not yet been mentioned in the text. The minimal node size is given as the percentage of the original sample size.

### 7.2.2. Discriminating variables

For each signal region a separate classifier has been trained whose configuration parameters are adjusted according to those listed in Tab. 7.2. Therefore, a set of variables, which more or less well discriminate between the  $t\bar{t}Z$  signal and the  $t\bar{t}$  background, has been defined and serves as an input for the construction of the BDT output variable. A list of all variables used for the exclusive and inclusive signal regions for the discrimination between  $t\bar{t}$  and  $t\bar{t}Z$  is given in Tab. 7.3. As some variables behave differently depending on whether the selection criteria for an exclusive or an inclusive signal region are applied, two different sets of discriminating variables have been investigated. A short physical motivation for each variable will be given in the following.

As the probability of gluon irradiation goes with  $\alpha_s^n$ , where  $n$  is the number of radiated gluons, it decreases with the amount of gluon emission provided that  $\alpha_s < 1$ . Hence, the jet multiplicity in  $t\bar{t}Z$  events is typically higher than in  $t\bar{t}$  events. In order to be selected, in  $t\bar{t}$  events at least one

inclusive SRs	exclusive SRs
$p_T(1^{\text{st}} \text{ jet})$	$p_T(1^{\text{st}} \text{ jet})$
$p_T(5^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$
$p_T(6^{\text{th}} \text{ jet})$	$p_T(6^{\text{th}} \text{ jet})$
$H_T$	$H_T$
$H_T^{\text{lep}}$	$H_T^{\text{lep}}$
$S_T$	$S_T$
$m_T^W$	$m_T^W$
$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$
$m_T^{\text{reco}}(Z)$	$m_T^{\text{reco}}(Z)$
$p_T^{\text{reco}}(Z)$	$p_{\text{reco}T}(Z)$
$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$
$m_T^{\text{reco}}(t\bar{t})$	$m_T^{\text{reco}}(t\bar{t})$
$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$
$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$
$\Delta R(\text{add. jets})$	$\Delta R(\text{add. jets})$
$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$
$\log L_{\text{KLFitter}}$	$\log L_{\text{KLFitter}}$
$p_T(7^{\text{th}} \text{ jet})$	–
$N(\text{jets})$	–
$N(b\text{-jets})$	–
$N(b\text{-jets})/N(\text{jets})$	–

Tab. 7.3.: List of all input variables for the BDT training in order to separate the signal from the dominant  $t\bar{t}$  background. The last four variables from above are used only within the inclusive signal regions as the  $(b\text{-})$ jet multiplicities are constant quantities in the exclusive signal region and they thus cannot be trained on (cf. Fig. 7.3). As can be observed in Fig. 7.5c, the transverse momentum of the seventh jet in an event has a too small discrimination in the exclusive signal regions for which reason it is just used within the inclusive ones. The reconstructed invariant and transverse quantities are indicated with “inv” and “T” as subscript, respectively.

gluon must have already been emitted, whereas in  $t\bar{t}Z$  events the first two additional jets stem from the  $Z$  boson decay. Thus, the chance to observe another gluon emission for signal events is much more likely than for  $t\bar{t}$ , as indicated by the different jet multiplicities in Fig. 7.3. As these variables are constant within exclusive signal regions, the depicted images show the distributions of SR1 and SR2, respectively. Though, this criterion can only be applied in the two inclusive signal regions, naturally. Since the  $Z$  boson decays to about 15 % into a pair of  $b$ -quarks, the  $b$ -jet multiplicity is also higher than for the  $t\bar{t}$  background as there only two  $b$ -quarks are expected. Consequently, the ratio of the numbers of  $b$ -jets and jets in an event differs for signal and background, respectively. Fig. 7.6 shows the  $H_T$  variable which is equal to the scalar sum of all transverse jet momenta in an event. As signal events have on average a higher jet multiplicity, also the sum of momenta will be slightly higher. The  $Z$  boson is either radiated from one of the top quarks or from the whole  $t\bar{t}$  system, hence the  $t\bar{t}$  system is supposed to carry more momentum as it would have gained in case of a gluon radiation due to the stronger recoil. Therefore, the jets of the final state have higher momenta, thus enhancing the discrimination power of  $H_T$ . Another two variables,  $H_T^{\text{lep}}$  and  $S_T$  which are the scalar sum of all transverse momenta of all jets and leptons and the scalar sum of

all transverse momenta in an event, respectively, exhibit the same behaviour as shown in Fig. 7.4b and 7.4c.

In events containing a real  $Z$  boson the sum of momenta of their decay products is higher than for jets originating from gluons, since the  $Z$  boson transfers its whole energy to the two selected additional jets when decaying (cf. Fig. 7.4d). Of course, these discrepancies can also be observed for the transverse momenta of the two jets separately (cf. Fig. 7.4e and 7.4f). For the transverse mass of the  $Z$  boson, defined as

$$m_{\text{T}}^2 = E^2 - p_{\text{T}}^2 = m^2 + p_z^2 \quad , \quad (7.6)$$

depicted in Fig. 7.5a, the discrimination originates from the difference in the invariant masses of a  $Z$  boson and a gluon which amount to about 91 GeV and 0 GeV, respectively. As for the radiation of a real  $Z$  boson much more energy is needed than for gluon emission, the transverse mass of the whole  $t\bar{t}$  system is on average higher for  $t\bar{t}Z$  than for  $t\bar{t}$  events, which is illustrated in Fig. 7.5b. A seventh jet is investigated regarding its momentum. Every gluon emission takes a certain amount of energy and momentum away, consequently further gluon radiations will be less energetic. However, this feature can only be made use of in the inclusive signal regions, as for the exclusive ones it either delivers not enough discrimination power or even cannot be applied in SR3 due to the selection criteria on the jet multiplicity. (cf. Fig. 7.5c). Because of the big mass difference between the top quark and its decay products, the jet with the highest transverse momentum in a given event is assumed to belong to one of the top quark decay chains. Due to stronger recoil from a real  $Z$  boson this jet should have an even larger momentum in  $t\bar{t}Z$  events than for the  $t\bar{t}$  background, which is illustrated by Fig. 7.5d.

Besides kinematic variables, separation between  $t\bar{t}Z$  and  $t\bar{t}$  can be observed when looking at the difference in angular space between the decay products of the  $Z$  boson,  $\Delta R(\text{add. jets})$ , depicted in Fig. 7.5e, as the angle between the decay products of particles that propagate with high momentum is smaller than for less boosted particles. A similar behaviour can be seen in Fig. 7.5f, which shows again the angular difference between the decay products of the  $Z$  boson but in the rest frame of the  $t\bar{t}Z$  system.

Furthermore, discrimination power has been observed for the invariant masses of the  $Z$  boson, both the top quark from the hadronic and leptonic decay chain, as well as for the invariant mass of the combined  $t\bar{t}$  system and the transverse mass of the  $W$  boson. The corresponding illustrations are presented in Fig. 7.6a–7.6e. Also the logarithm of the KL Fitter likelihood shows different slopes for the  $t\bar{t}Z$  signal and  $t\bar{t}$  background, as can be seen in Fig. 7.6f, which is further investigated in Sec. 6.6.

For the other signal regions, the corresponding images of all variables are depicted in Fig. E.9–E.17. All the presented variables do not discriminate much on their own, though in combination they provide better separation power.

### 7.2.3. BDT training results

When training a BDT, a test for overtraining is immediately done using a randomly mixed subset of the training events as mentioned in Sec. 7.2. The results of training and testing is shown in Fig. 7.7 for SR2, whereas the corresponding pictures for the other signal regions are depicted in Fig. E.18. On the left side of the images, the BDT response is shown for signal in blue and for background in red, where the coloured dots represent the training outcome and the coloured areas depict the results after the BDT training had been tested. On the right side the so-called receiver operating characteristic (ROC) curve is plotted for both the training in dark cyan and for its test in black. In the presented form the ROC curve shows the relation between the efficiencies of

background rejection and signal acceptance reflecting the performance of a binary classifier. For a classifier yielding no discrimination at all, the corresponding ROC curve would be a straight line with a ROC integral<sup>(2)</sup> of 0.5, whereas a perfect classifier would have a rectangular ROC curve with a ROC integral of 1 [205]. As the shapes of the BDT response composed of the dots and the areas, respectively, are in good agreement, overtraining could be prevented for all signal regions. This observation is enforced by the almost perfect superposition of the training and testing ROC curves.

During the training at each node a cut is set on the respective variable with the highest discrimination power, thus the importance of the input variables throughout the whole training procedure can be determined. This ranking might differ from the order in discrimination power, which has been generated before the training only by looking on the overlap of the shapes of the input variables given in Sec. 7.2.2. The listings of the variable rankings after training are presented in Tab. F.1 for all signal regions, whereas the corresponding ranking before the training is given in Tab. F.2.

#### 7.2.4. BDT evaluation

After the training has been successfully tested, the trained classifiers are evaluated for each background and the signal sample, respectively. The results of the BDT evaluation for SR2 are shown in Fig. 7.8, while the corresponding pictures of the other signal regions are depicted in Fig. E.20. On the left side of these images, all considered backgrounds are stacked on top of each other while the signal is plotted before with a red line. In order to make the impact of the BDT evaluation visible, signal and backgrounds are normalised to their respective number of entries. In contrast, the right side shows the significance when scanning from the left to right through the distributions. On both sides, the dashed arrow marks the BDT response value on which a cut must be set in order to gain the maximum significance. Furthermore, the actual cut value is given within the significance plots together with the corresponding significance and the formula with which it has been calculated. In order to quantify the enhancement of separation power achieved with the BDT, in Tab. 7.4 the corresponding significance values for all signal regions are listed. The significance has been calculated three times for each signal region according to Eq. (6.2) with either no error, only a statistical error or with both a statistical and systematic error on the input values. Assuming no uncertainties, the significance has been increased by a factor of roughly 18 % for the inclusive signal regions and 12 % for the exclusive signal regions, respectively, with respect to the cut-based analysis. Statistical uncertainties have almost no influence, lowering the results by only 0.002. The assumption of 30 % systematic uncertainties again reduces the significance within all signal regions by a factor of approximately 12 % as for the cut-based results in Tab. 6.6. However, the increase in significance with respect to cut-based results remains the same as for significances calculated without any uncertainty imposed, namely 18 % for inclusive signal regions and 12 % for exclusive signal regions.

#### 7.2.5. Variable decorrelation

It likely happens that some variables are more or less correlated to each other, as it can be seen in Fig. E.23 – E.26, depicting the situation for each signal region, respectively. In particular, this occurs for variables which either share a rather similar definition, for example  $H_T$ ,  $H_T^{\text{lep}}$  and  $S_T$ , or are kinematically related to each other, as  $p_T(5^{\text{th}} \text{ jet})$ ,  $p_T(6^{\text{th}} \text{ jet})$  and  $p_T(Z)$ . As such correlations possibly reduce the performance of a classifier, thus, in order to counteract some preprocessing can be applied on the input samples (cf. Sec. 7.1.3).

<sup>(2)</sup> The ROC integral is calculated from the area beneath the ROC curve.

	SR1	SR2	SR3	SR4
significance	2.627	2.842	1.118	1.090
significance (stat. err.)	2.625	2.840	1.116	1.088
significance (stat. & syst. err.)	2.305	2.495	0.981	0.957

Tab. 7.4.: Summary of the increase in sensitivity achieved by evaluating the trained BDT classifiers, listed for all signal regions. The calculation of the statistical significance according to Eq. (6.2) has been done with either no uncertainties, only the statistical uncertainties or both the statistical and systematic uncertainties, assuming a systematic uncertainty of 30 %. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

In this analysis the PCA preprocessing method has been applied on the most correlated input variables before the training of the classifiers for all signal regions, leading to the training results depicted in Fig. 7.9 for SR2 and in Fig. E.19 for the other signal regions. With the decorrelation procedure a very small increase in the ROC integral and thus in the performance of the BDT training could be achieved, though the distributions for training and testing show a similar behaviour to those without variable decorrelation. As for the training without a decorrelation preprocessing presented in Sec. 7.2.3, the training with decorrelated input variables has not run into overtraining. The corresponding variables rankings are presented in Tab. F.3 – F.4.

The newly trained BDT classifiers have been evaluated on all Monte Carlo simulation samples, resulting in the distributions and the corresponding significance plots shown in Fig. 7.10 for SR2 and in Fig. E.21 for the other signal regions. As before, for each signal region the corresponding increase in sensitivity and the significance value which has been calculated according to Eq. (6.2). In Tab. 7.5 the results are shown. With the variable decorrelation a further increase in significance

	SR1	SR2	SR3	SR4
significance	2.606	2.877	1.114	1.096
significance (stat. err.)	2.604	2.874	1.112	1.094
significance (stat. & syst. err.)	2.287	2.525	0.977	0.962

Tab. 7.5.: Summary of the increase in sensitivity achieved by evaluating the trained BDT classifiers, listed for all signal regions. Before the BDT training a variable decorrelation has been performed. The calculation of the statistical significance according to Eq. (6.2) has been done with either no uncertainties, only the statistical uncertainties or both the statistical and systematic uncertainties, assuming a systematic uncertainty of 30 %. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

could be achieved for SR2 by 1.2 % and for SR4 by 0.6 % with respect to the undecorrelated BDT training. On the contrary, for SR1 and SR3 the significance even decreased by 0.8 % and 0.04 %, respectively. While for only statistical uncertainties only a very low decrease in significance can be noticed, systematic uncertainties of 30 % lowers the calculated significance in each signal region by a factor of around 12 % as for the cut-based results in Tab. 6.6 or the results from BDT training and evaluation without variable decorrelation (cf. Tab 7.4). Due to the poor increase in separation power between signal and background contributions the preprocessing of the input variables has not been proven to be too beneficial, for which reason it will not be applied anymore in further studies.

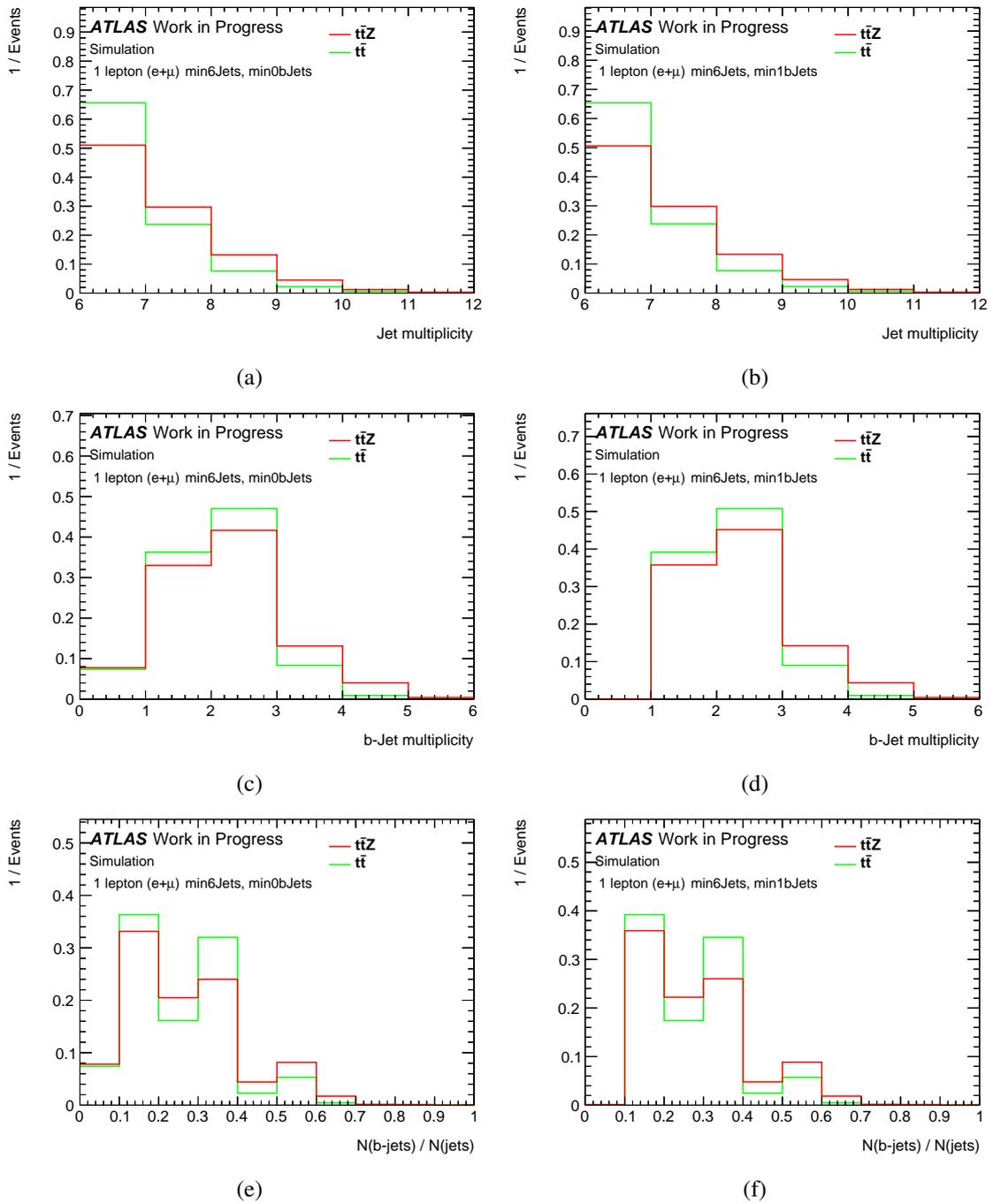


Fig. 7.3.: Illustration of jet multiplicity (cf. Fig. (a), (b)), the  $b$ -jet multiplicity (cf. Fig. (c), (d)) and the ratio of both (cf. Fig. (e), (f)) for the  $t\bar{t}Z$  signal and the dominant  $t\bar{t}$  background. On the left side the situation for SR1 is shown, whereas on the right side SR2 is shown. All distributions are normalised to total number of events in order to compare their shapes.

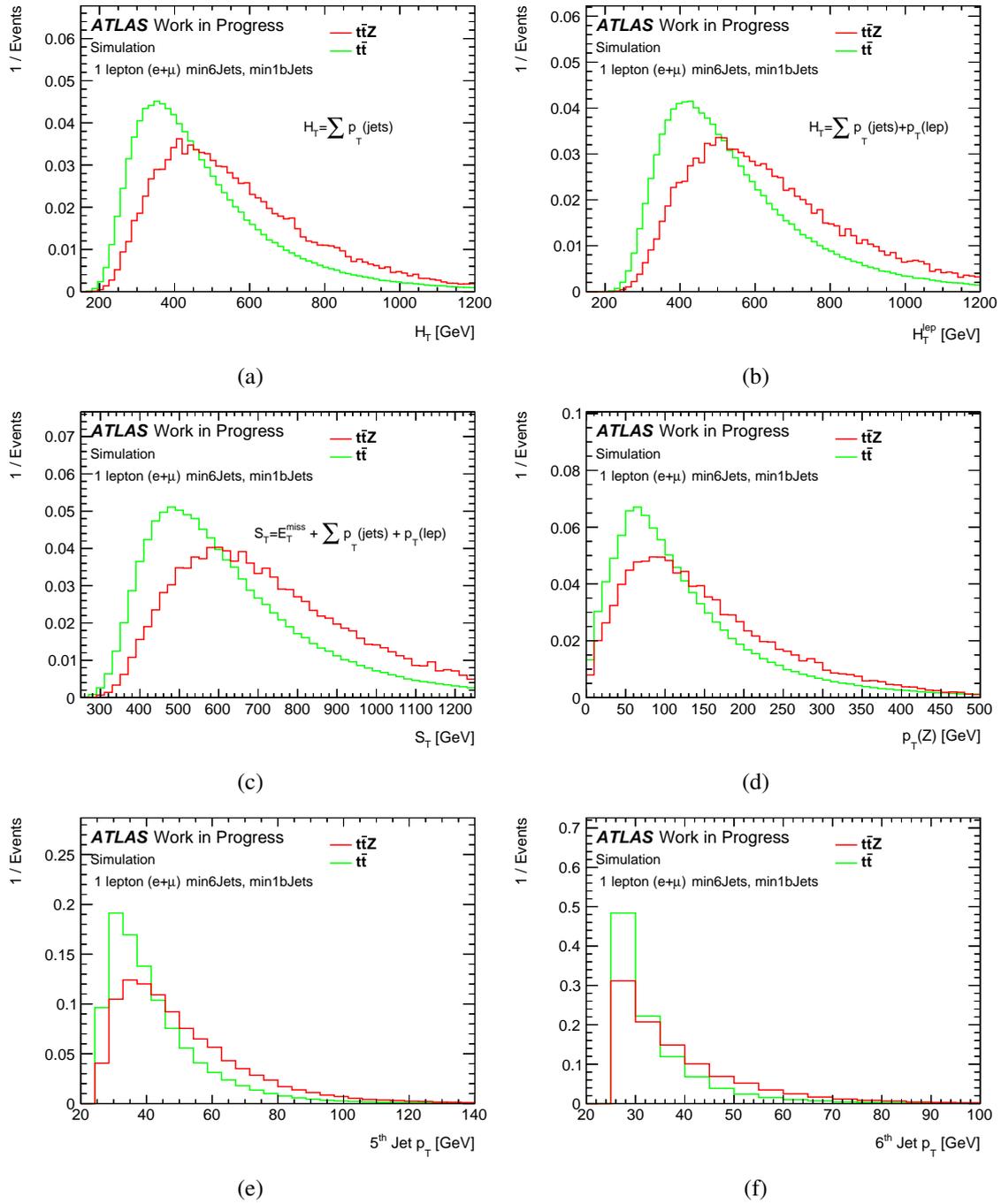


Fig. 7.4.: Illustration of  $H_T$  in the upper left figure (Fig. (a)) and of  $H_T^{\text{lep}}$  (Fig. (b)),  $S_T$  (Fig. (c)),  $p_T^{\text{reco}}(Z)$  (Fig. (d)),  $p_T(5^{\text{th}} \text{ jet})$  (Fig. (e)) and of  $p_T(5^{\text{th}} \text{ jet})$  (Fig. (e)), respectively. All distributions are shown for SR2 and are normalised to total number of events in order to compare their shapes.

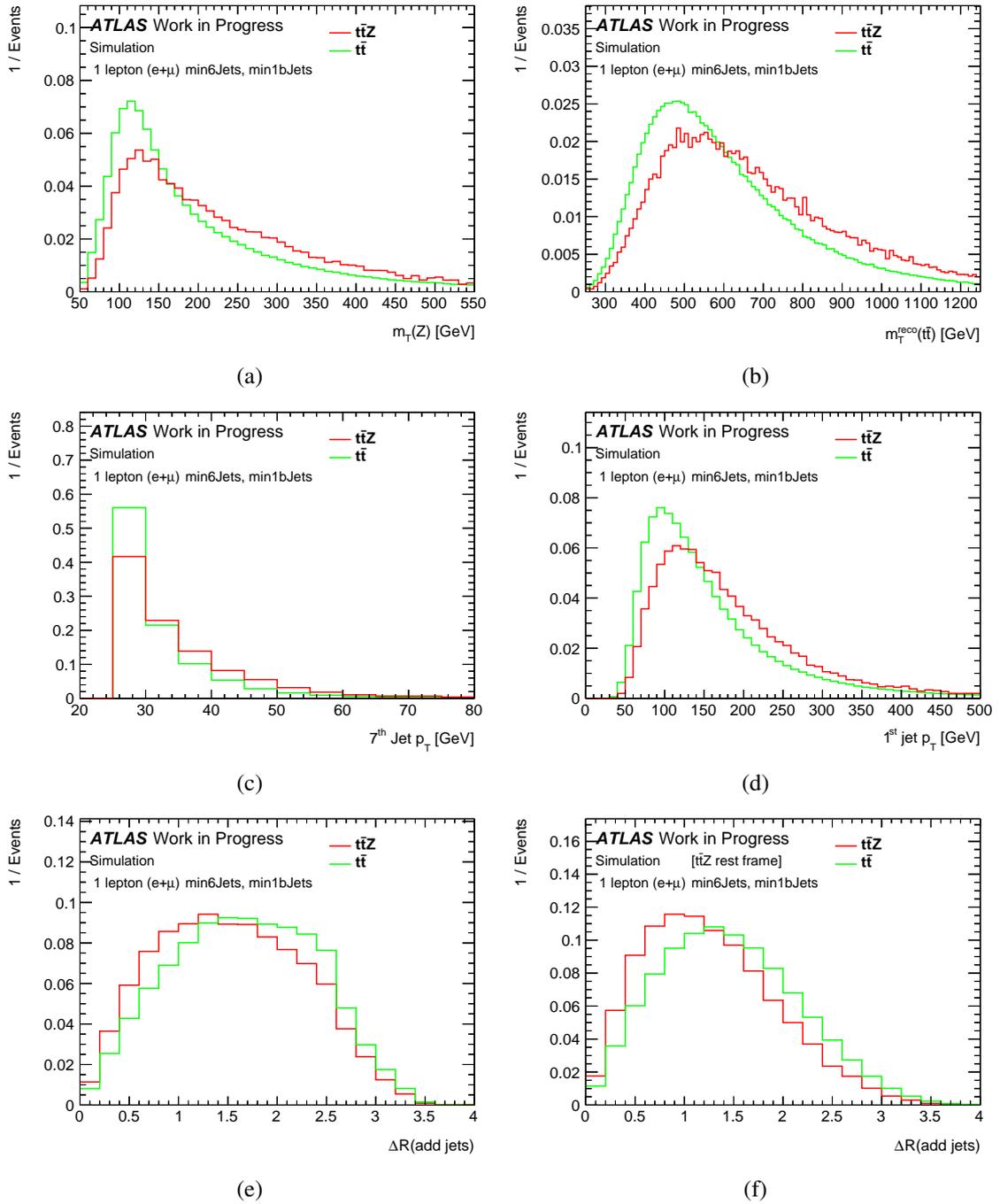


Fig. 7.5.: Illustration of  $m_T^{\text{reco}}(Z)$  in the upper left figure (Fig. (a)) and of  $m_T^{\text{reco}}(t\bar{t})$  (Fig. (b)),  $p_T(7^{\text{th}} \text{ jet})$  (Fig. (c)),  $p_T(1^{\text{st}} \text{ jet})$  (Fig. (d)),  $\Delta R(\text{add. jets})$  (Fig. (e)) and of  $\Delta R^{t\bar{t}Z}(\text{add. jets})$  (Fig. (f)), respectively. All distributions are shown for SR2 and are normalised to total number of events in order to compare their shapes.

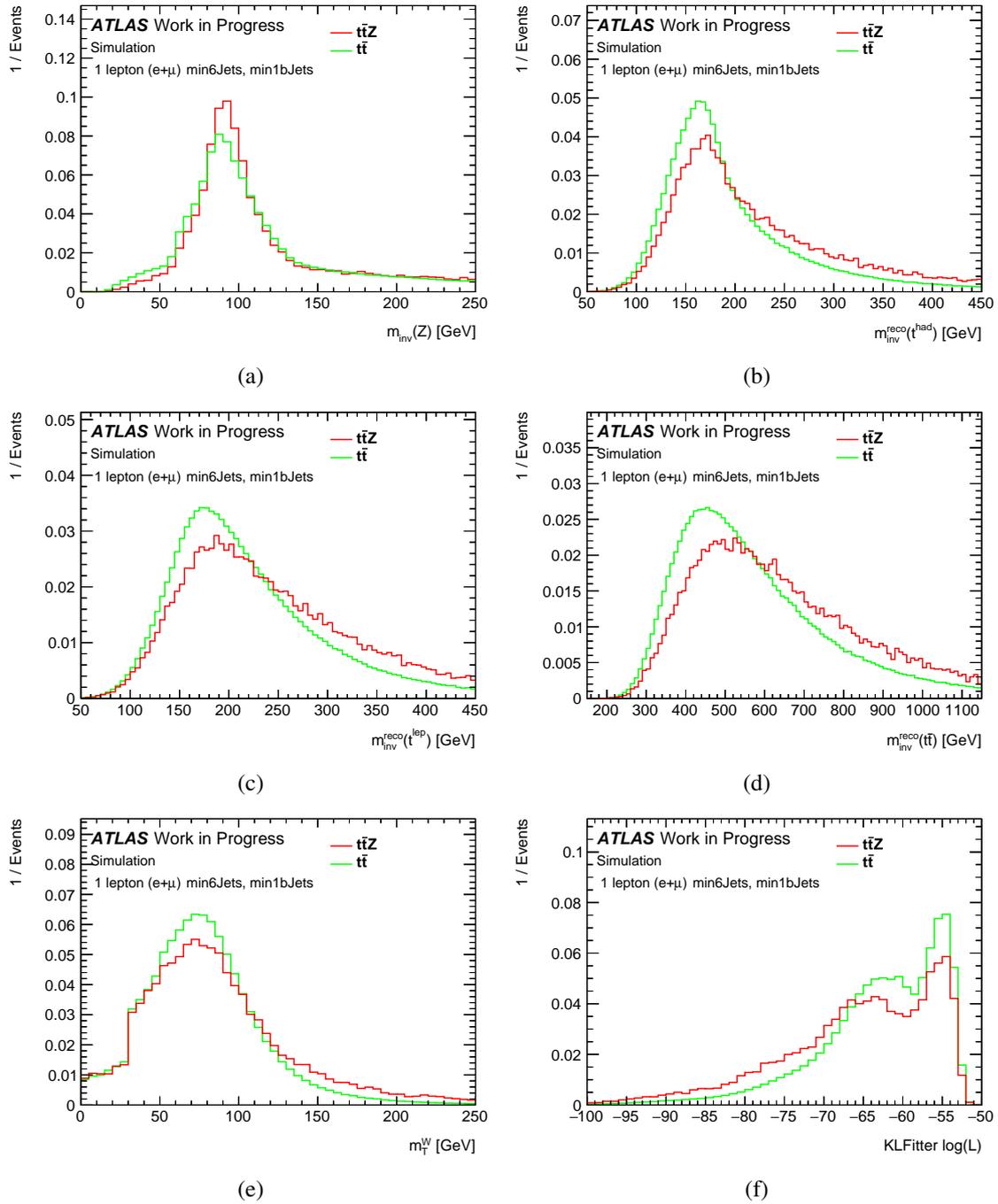


Fig. 7.6.: Illustration of  $m_{inv}^{reco}(Z)$  in the upper left figure (Fig. (a) and, of  $m_{inv}^{reco}(t^{had})$  (Fig. (b)),  $m_{inv}^{reco}(t^{lep})$  (Fig. (c)),  $m_{inv}^{reco}(t\bar{t})$  (Fig. (d)),  $m_T^W$  (Fig. (e)) and of  $\log L_{KL\text{Fitter}}$  (Fig. (f)), respectively. All distributions are shown for SR2 and are normalised to total number of events in order to compare their shapes.

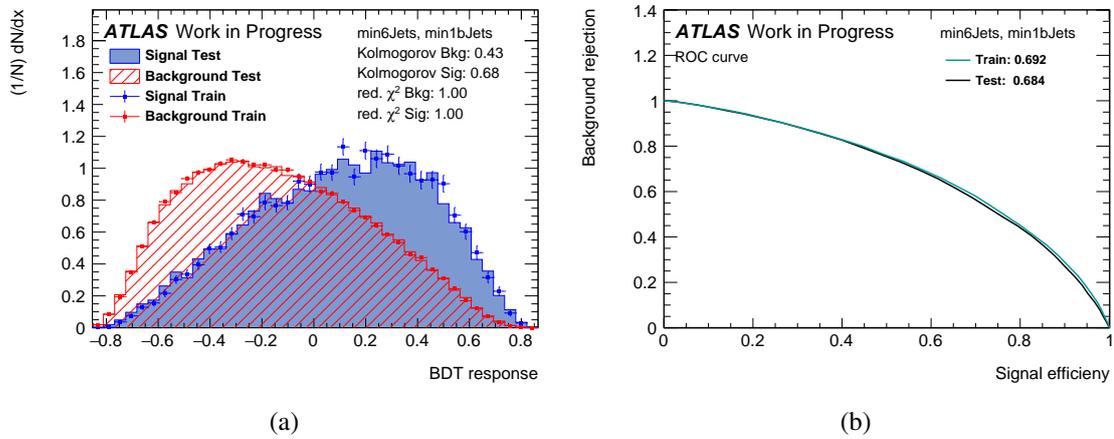


Fig. 7.7.: Presentation of the BDT training results for the second signal region, depicting the BDT response (Fig. (a)) for signal in blue and for background in red, respectively, and the corresponding ROC curves for training and testing (Fig. (b)). Both training and testing seems to be in good agreement, thus showing no indication of overtraining.

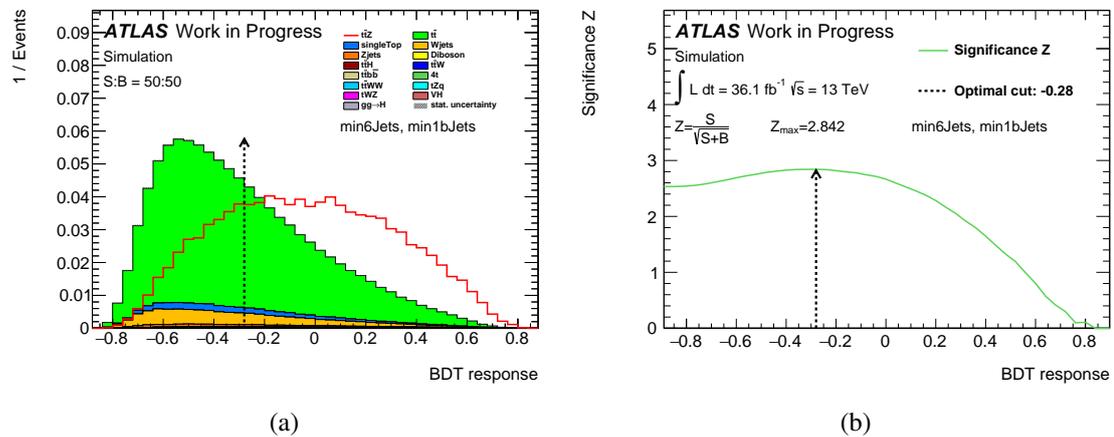


Fig. 7.8.: Presentation of the BDT evaluation for the second signal region, depicting the evaluated BDT response (Fig. (a)) and the corresponding significance curve (Fig. (b)) for an integrated luminosity of  $36.1 \text{ fb}^{-1}$ . Both signal and backgrounds are normalised to their respective number of events, thus illustrating the achieved separation. The dashed arrow indicates the optimal cut on the BDT response.

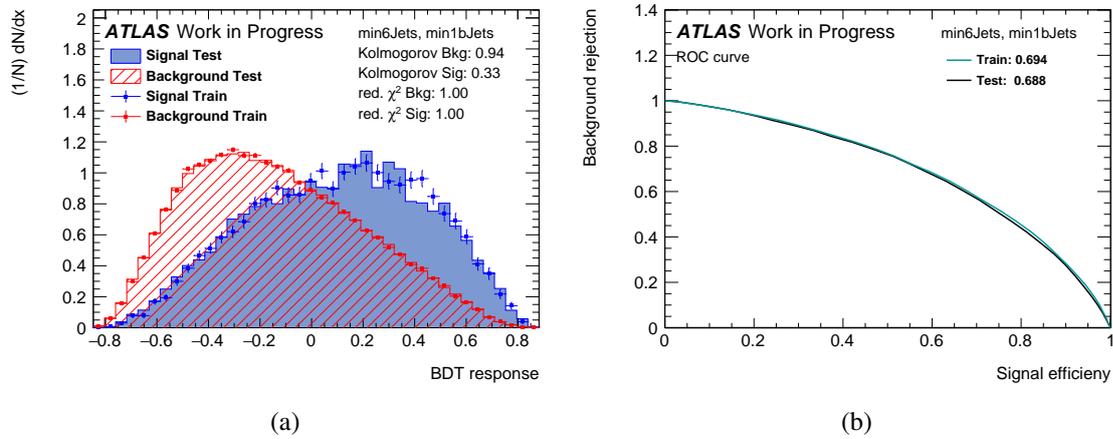


Fig. 7.9.: Presentation of the BDT training results for the second signal region, depicting the BDT response (Fig. (a)) for signal in blue and for background in red, respectively, and the corresponding ROC curves for training and testing (Fig. (b)). Both training and testing seems to be in good agreement, thus showing no indication of overtraining. Before the BDT training a variable decorrelation has been performed.

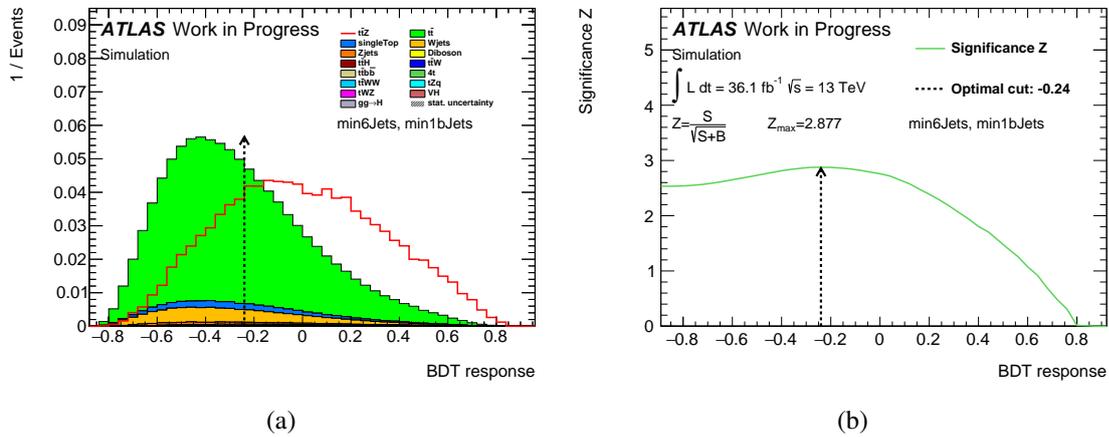


Fig. 7.10.: Presentation of the BDT evaluation for the second signal region, depicting the evaluated BDT response (Fig. (a)) and the corresponding significance curve (Fig. (b)) for an integrated luminosity of  $36.1 \text{ fb}^{-1}$ . Both signal and backgrounds are normalised to their respective number of events, thus illustrating the achieved separation. The dashed arrow indicates the optimal cut on the BDT response. Before the BDT training a variable decorrelation has been performed.

## 8. Sensitivity studies towards $100 \text{ fb}^{-1}$

This analysis is contributing to the ATLAS sensitivity studies focussing on new analyses and techniques for top quark studies which likely lead to new or more precise results regarding measurements of the top quark cross section, mass or properties, using the whole Run 2 dataset which corresponds to  $100 \text{ fb}^{-1}$ .

### 8.1. Cut-based analysis

After the event yields in Tab. 6.5 are scaled to the increased luminosity, the increase in sensitivity and the corresponding significances have been calculated. As all event yields are scaled up to  $100 \text{ fb}^{-1}$  with a constant factor, the  $t\bar{t}$ -to- $t\bar{t}Z$  ratio stays the same as for  $36.1 \text{ fb}^{-1}$ , naturally. In contrast, the significance considerably increases due to the structure of Eq. (6.2). The results are summarised in Tab. 8.1, including the significance determined imposing either no uncertainties, only statistical uncertainties or both statistical and systematic uncertainties on the input values. The significance values calculated with only statistical uncertainties are slightly smaller than those

	SR1	SR2	SR3	SR4
$t\bar{t}/t\bar{t}Z$	216.97	217.87	317.89	194.92
decrease w.r.t. theory [%]	78.12	78.01	67.94	80.34
significance (no err.)	3.680	4.018	1.652	1.626
significance (stat. err.)	3.678	4.016	1.651	1.624
significance (stat. & syst. err.)	3.229	3.526	1.450	1.427

Tab. 8.1.: Summary of the increase in sensitivity achieved by applying the selection criteria given above, listed for all signal regions for an integrated luminosity of  $100 \text{ fb}^{-1}$ . The calculation of the statistical significance according to Eq. (6.2) has been done with either no uncertainties, only the statistical uncertainties or both the statistical and systematic uncertainties, assuming a systematic uncertainty of 30%. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

without any uncertainty, as expected. The inclusion of a systematic uncertainty of 30% lowers the results more noticeably, namely by a factor of 12%. It could be noticed that the ratio between the significances of the particular signal regions remains the same for  $100 \text{ fb}^{-1}$  as it is for  $36.1 \text{ fb}^{-1}$ .

### 8.2. BDT analysis

The evaluation of the BDT classifiers trained on each signal region is repeated for the lifted luminosity, resulting in the BDT response distributions and significance plots depicted in Fig. 8.1 for SR2 and in Fig. E.22 for the other signal regions. As for the training of the BDT classifier the luminosity, which is a constant multiplicative factor, is not of importance but much more the shapes and correlations of the input variables, the training procedure has not been repeated. The significance values have been determined using Eq. (6.2) imposing either no error, only a statistical error or with both a statistical and systematic error on the input values. The results for all



## 9. Systematic uncertainties

In Sec. 6.4 and 8.1, various results for the expected significance have been stated in Tab. 6.6 and Tab. 8.1, respectively. These values have been calculated imposing either no uncertainties, only statistical uncertainties or both statistical and systematic uncertainties on the respective event yields used for the calculations. The statistical uncertainties have been directly taken from the event yields assuming Poisson statistics with a simultaneous consideration of the Monte Carlo weights, whereas a constant value of 30 % is presumed for the estimation of systematic uncertainties. Although not a full evaluation involving all uncertainties is done within this analysis, an up- and down-variation for the most important systematic uncertainties is performed separately. The goal is to determine this source of uncertainty which has the largest impact on the result of the BDT evaluation.

### 9.1. Experimental uncertainties

The term systematic uncertainties comprises a multitude of uncertainties on the experimental methods for data acquisition, calibration and correction as well as theoretical uncertainties on the modelling of signal and background processes. Technically, to each systematic uncertainty one or more parameters are assigned to quantify and model its impact on the number of selected events and their kinematics.

The integrated luminosity used for normalisation of simulated events is assigned with an uncertainty of 2.1 %, derived from a calibration of the luminosity scale using  $x$ - $y$  beam-separation scans [209]. As stated in Sec. 3.4, a lot of supplementary scattering processes happen alongside the hard interactions of interest, referred to as pile-up. In order to reflect the observations in data with the Monte Carlo simulations, a pile-up reweighting (PRW) is applied to which one uncertainty parameter is assigned. Experimental uncertainties related to muons are considered for the muon momentum calibration and the efficiency corrections on muon identification, isolation, and trigger selection [163], each represented by one parameter. Similarly, systematic uncertainties associated to electrons are taken into account for the electron calibration [166] and electron efficiency corrections on reconstruction, identification, isolation, and trigger selection [167]. For the jet vertex tagging, one systematic parameter represents uncertainties related to the statistical precision and choice of the Monte Carlo generator in the measurements and residual contamination from pile-up jets [210]. Furthermore, several systematic parameters are considered for the  $b$ -tagging calibration, representing uncertainties on the  $b$ -jet and  $c$ -jet efficiency and light-flavour jet mistag efficiency measurements [192, 193], respectively. Further systematic uncertainties related to jets not being investigated here are systematic components from the jet energy scale (JES) and the jet energy resolution (JER) [180]. Uncertainties on the measurement and reconstruction of missing transverse energy [195] are also not taken into account. However, as those systematic uncertainties are expected to give major contributions, they shall be investigated soon.

### 9.2. Theoretical uncertainties

In order to estimate the impact of uncertainties related to the top quark background, the nominal Monte Carlo samples are evaluated against datasets with varied generator settings or samples

produced with alternative generators [211]. To account for uncertainties due to the choice of different Monte Carlo generators for the dominant  $t\bar{t}$  background process, the nominal POWHEG event generator is compared to AMC@NLO, where the modelling of the parton shower, hadronisation and underlying event is performed by HERWIG++ using the UEEE5 tune. For the description of hadronisation in  $t\bar{t}$  two different models are compared, namely on the one hand the Lund string model [114, 150] implemented in PYTHIA 6, and on the other hand the cluster fragmentation model [115] used in HERWIG++. The hard scatter interaction is simulated with POWHEG both times (cf. Sec. 4.3). In the hadronic environment of the LHC, additional radiation in the initial and final state occurs very likely. In order to estimate the effects of a change in the amount of initial and final state radiation on the BDT output, different parameters have been varied within the nominal  $t\bar{t}$  Monte Carlo simulation to allow for an increase or decrease in QCD radiation [212]. The uncertainty on the treatment of interference between  $t\bar{t}$  and the  $Wt$ -channel of single top quark production is estimated by comparing the diagram removal (DR) and diagram subtraction (DS) schemes which are used to subtract interference contributions from simulated  $tW$  events. Furthermore, the uncertainties on the theoretical predictions of the production cross sections of  $t\bar{t}$  and  $t\bar{t}Z$  are evaluated by varying the normalisation of the respective processes within the theoretical boundaries.

### 9.3. Impact of systematic uncertainties on analysis

In the following the impact of the systematics which tend to most distort the result of the BDT evaluation is to be found. Therefore, in a first step the behaviour of several systematics on the input variables of the BDT classifier is investigated. Due to the excellent resolution of the electromagnetic calorimeter and the muon spectrometer, the experimental uncertainties on both the electron and muon measurement are negligibly small, inducing deviations of not even a few percent. Thus, they are not further investigated. However, systematic uncertainties related to jets have a greater impact on the distributions of the input variables, so they are taken into account furthermore. In Fig. 9.1 the impact of the uncertainties on the  $b$ -tagging procedure is depicted for the  $t\bar{t}Z$  signal process as well as for the dominant  $t\bar{t}$  background in SR2. The corresponding pictures for SR1 can be found in Fig. E.27. Exemplary shown is one parameter with the largest impact for the uncertainty on the  $b$ -jet and  $c$ -jet tagging efficiency as well as on the light jet mistag efficiency for the ratio of  $b$ -jet multiplicity and the number of all jets in an event, respectively, exhibiting considerably deviations of higher than 20%. It turns out that all other parameters from each category can be neglected since they do not significantly contribute. Rather small changes are caused by the uncertainties related to the jet vertex tagging and to the pile-up reweighting procedure, respectively, resulting in a variation of hardly a few percent which is smaller than the corresponding statistical uncertainties. Thus, they are not considered further. The uncertainty on the integrated luminosity of 2.1%, resulting in a constant shift of the whole distributions, is very small compared to theoretical uncertainties and therefore not further investigated.

The impact of theoretical systematics in SR2 is shown in Fig. 9.2 on the basis of the inverse mass of the reconstructed top quark belonging to the leptonically decaying  $W$  boson. The corresponding illustrations for the other signal regions are depicted in Fig. E.28 and E.29, respectively. The largest difference between the curves can be noticed comparing different Monte Carlo generators for parton shower modelling (cf. Fig. 9.2c) which amounts up to 40% or even more. Deviations of the same order of magnitude arise for the comparison of the diagram removal and diagram subtraction techniques applied for the single top  $Wt$ -channel. However, although the discrepancy between the distributions for diagram removal and diagram subtraction are fairly high, its actual impact should be quite low because single top production is by far not the most dominant back-

ground process. Additional QCD radiation as well as the comparison of different Monte Carlo generators for the hard scatter interactions has an impact of about 10–20 % (cf. Fig. 9.2a and 9.2b) on the distribution shape. Despite the uncertainties of the different Monte Carlo generators for the hard interaction, all theoretical uncertainties turn out to not only vary in the absolute amount of events, but also in the shape of the distributions. It is remarkable that in SR3 the systematic variations for additional QCD radiation (cf. Fig. E.28c) and the hard scatter simulation (cf. Fig. E.28d) are in very good agreement with the nominal shape despite from statistical fluctuations in regions with a low event multiplicity. A possible reason could be the limitation to six jets in SR3 and therefore a more accurate simulation of the hard interaction due to the reduced amount of additional partons with respect to the other signal regions. The uncertainties on the theoretical calculations of the  $t\bar{t}$  and  $t\bar{t}Z$  cross sections have been taken together in such a way that the maximal deviation from the nominal value is obtained by summing up all deviations in each direction. Thus, with the numbers stated in Sec. 2.3.1.2, the overall uncertainty on the  $t\bar{t}$  cross section amounts to  $-7.7\%/ + 6.6\%$ . With the same approach the overall uncertainty on the  $t\bar{t}Z$  cross section is  $-16.9\%/ + 15.5\%$  using the values from Sec. 2.4. Both uncertainties result in a horizontal shift of the distributions. From the systematic parameters which are left for further studies the theoretical systematics seem to affect the nominal input variables more profoundly than the experimental systematics which have been so far investigated.

After the evaluation of the BDT on the respective Monte Carlo samples, the systematics with a non-negligible impact on the input variables have been investigated regarding their influence on the BDT. It appears that the three experimental uncertainties from the  $b$ -tagging calibration procedure seem to be reduced by the BDT training as their fluctuations around the nominal value lie now within the statistical uncertainties, as depicted in Fig. 9.3 for SR2 and in Fig. E.30 for SR1. On the contrast, the theoretical uncertainties significantly influence the distribution of the BDT response as depicted in Fig. 9.4 for SR2. The corresponding illustrations for the other signal regions are depicted in Fig. E.31 and E.32, respectively. However, the deviation from the nominal shape for additional QCD radiation (cf. Fig. 9.4a) and for the comparison of diagram removal and subtraction (cf. Fig. 9.4d) increased a bit, whereas the impact of the comparison of different generators for the hard interaction seems to stay the same (cf. Fig. 9.4b). In contrast, uncertainties on the parton shower modelling decreased as shown in Fig. 9.4c to around 30 %. As for the input variables, in SR3 the shapes of the nominal and the systematic distribution are in very good agreement for additional QCD radiation (cf. Fig. E.31c) and the hard scatter interaction (cf. Fig. E.31d), albeit for the latter a slight tendency to the left is noticed for the green shape. Non-linear variations of the shapes can still clearly be seen for additional QCD radiation and the comparison of diagram removal and subtraction. On the contrary, for the parton shower modelling the ratio between nominal and systematic variation has a more horizontal slope after the evaluation of the BDT. In addition, the uncertainties on the cross sections of  $t\bar{t}Z$  and  $t\bar{t}$  give rise to a horizontal shift of the distributions, as explained above. The uncertainties on the  $t\bar{t}$  cross section is considerably smaller than other theoretical uncertainties wherefore its impact is reasonably low. Due to the small amount of signal events the impact of the uncertainties on the signal cross section should be also rather small.

## 9. Systematic uncertainties

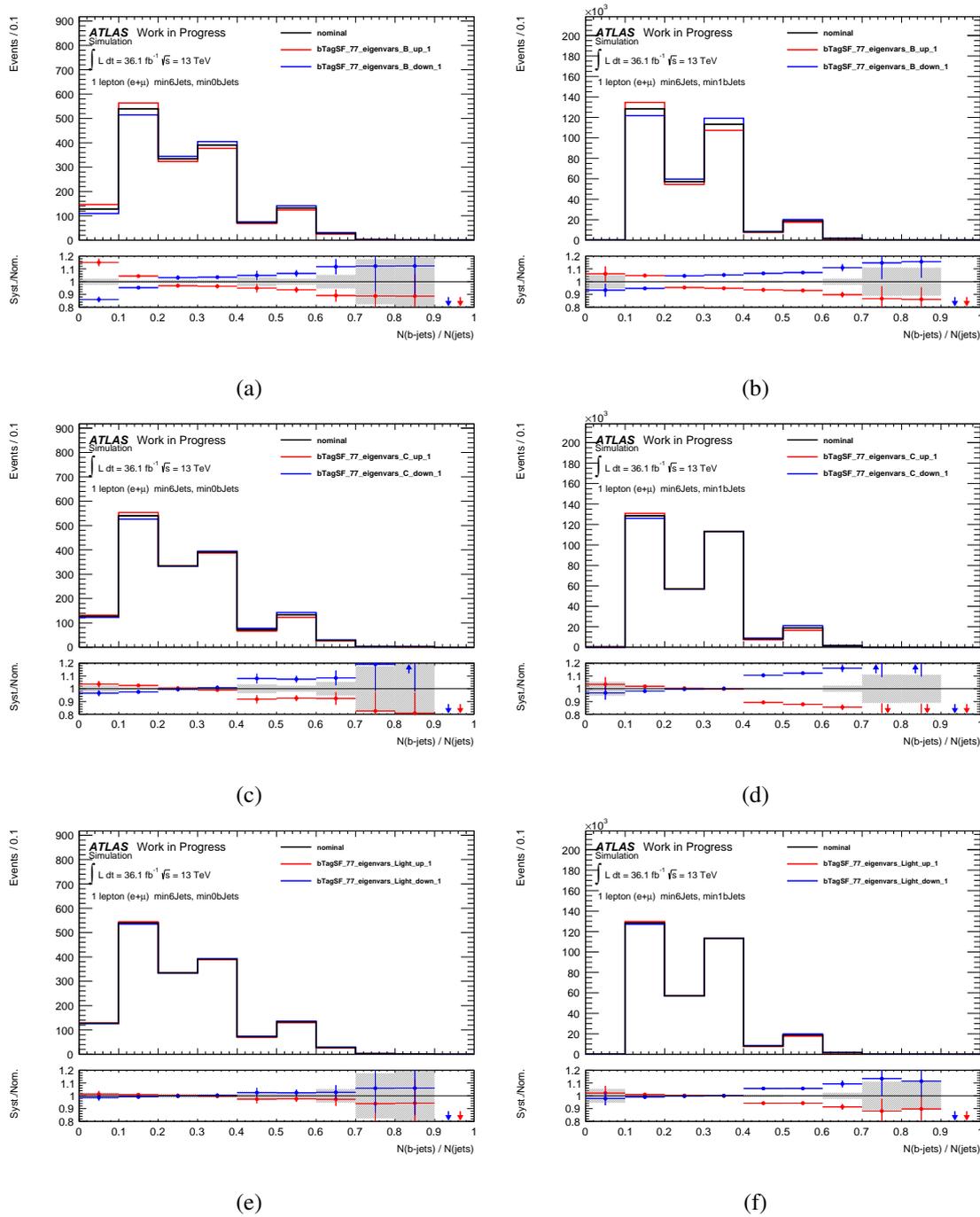


Fig. 9.1.: Illustration of the impact of the largest systematic uncertainties on the  $b$ -tagging calibration, shown for the ratio of the  $b$ -jet multiplicity to the number of all jets in an event in SR2. The two upper images show the distributions for the most striking systematic parameter for the  $b$ -jet efficiency (cf. Fig. (a), (b)). In the middle, the situation for the relevant parameter for the  $c$ -jet efficiency is depicted (cf. Fig. (c), (d)) and on the bottom the significant parameter for the light jet mistag rate is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. On the left side the situation for the  $t\bar{t}Z$  signal is illustrated, whereas on the right side the dominant  $t\bar{t}$  background is depicted.

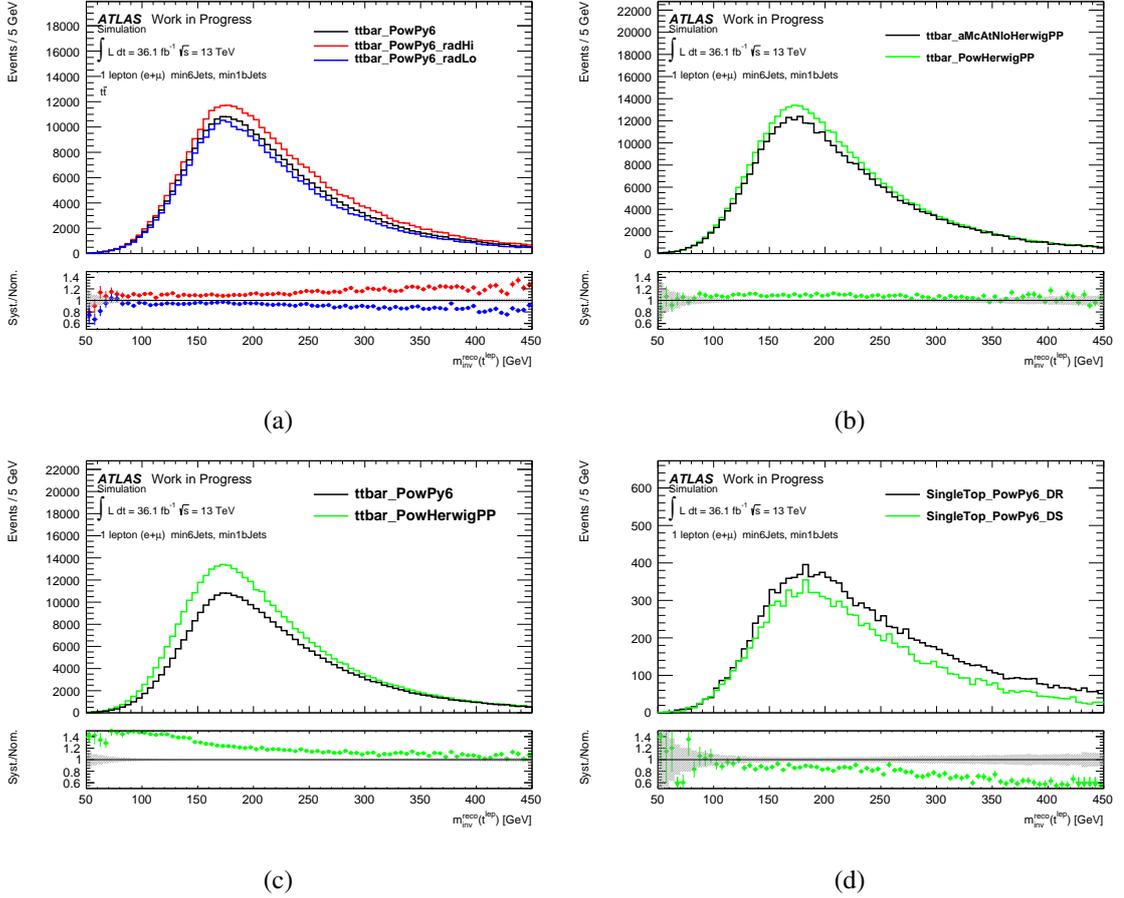


Fig. 9.2.: Illustration of the impact of theoretical uncertainties for the dominant  $t\bar{t}$  background and the  $Wt$ -channel of single top quark production in SR2, shown for  $m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$ . In the upper part, in Fig. (a) the impact of additional QCD radiation is shown and in Fig. (b) the comparison of different Monte Carlo generators for the simulation of the hard scatter event. Below, Fig. (c) depicts the comparison of different Monte Carlo generators for the parton shower modelling while Fig. (d) illustrates the comparison between diagram removal and diagram subtraction. The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.

## 9. Systematic uncertainties

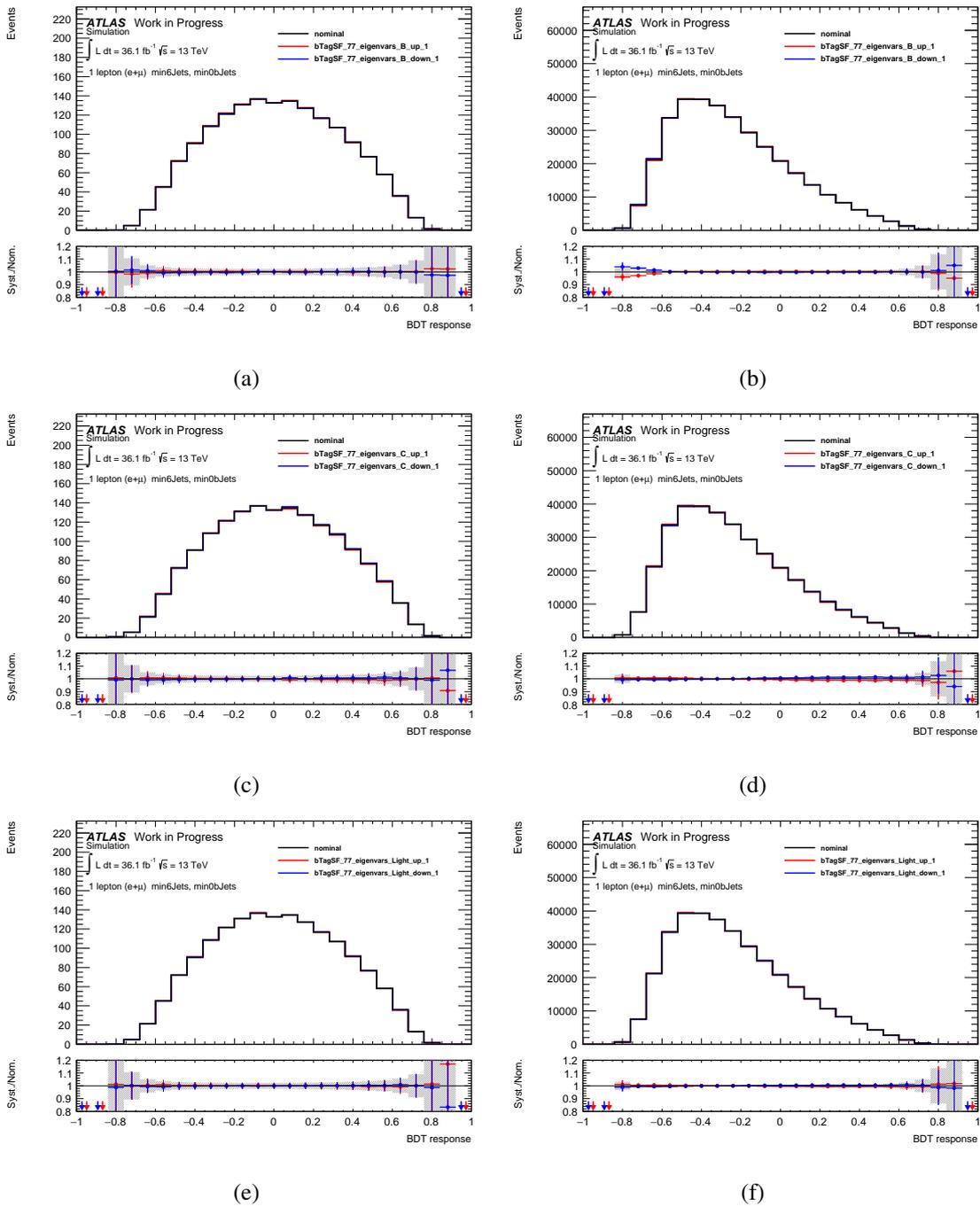


Fig. 9.3.: Illustration of the impact of the largest systematic uncertainties on the  $b$ -tagging calibration, shown for the BDT response in SR2. The two upper images show the distributions for the most striking systematic parameter for the  $b$ -jet efficiency (cf. Fig. (a), (b)). In the middle, the situation for the relevant parameter for the  $c$ -jet efficiency is depicted (cf. Fig. (c), (d)) and on the bottom the significant parameter for the light jet mistag rate is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. On the left side the situation for the  $t\bar{t}Z$  signal is illustrated, whereas on the right side the dominant  $t\bar{t}$  background is depicted.

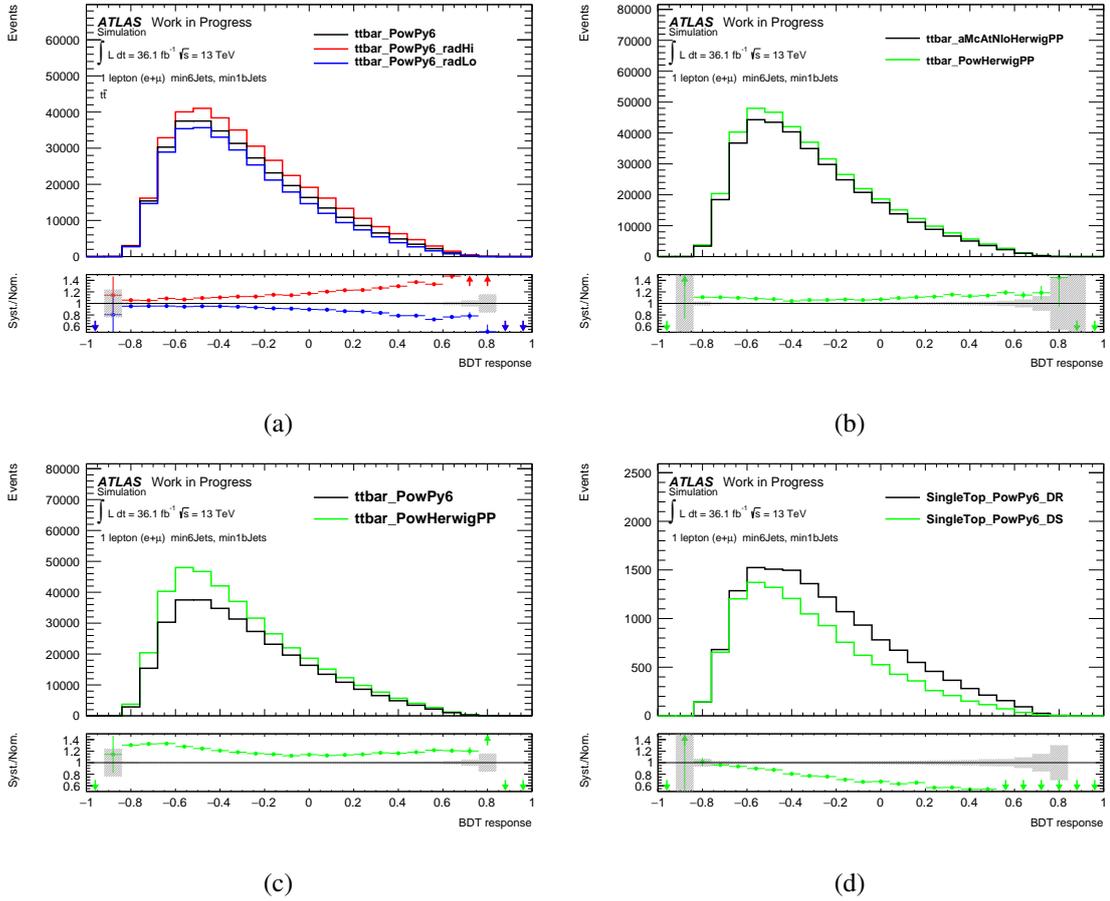


Fig. 9.4.: Illustration of the impact of theoretical uncertainties on the BDT response for the dominant  $t\bar{t}$  background and the  $Wt$ -channel of single top quark production in SR2. In the upper part, in Fig. (a) the impact of additional QCD radiation is shown and in Fig. (b) the comparison of different Monte Carlo generators for the simulation of the hard scatter event. Below, Fig. (c) depicts the comparison of different Monte Carlo generators for the parton shower modelling while Fig. (d) illustrates the comparison between diagram removal and diagram subtraction. The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.



## 10. Conclusion

In this analysis, first studies on the measurement of the production cross section of top quark pairs in association with a  $Z$  boson in the 1-lepton channel have been presented, using proton-proton collision data at a centre-of-mass energy of 13 TeV recorded by the ATLAS experiment at the LHC in the years 2015 and 2016. In contrast to multi-lepton channels, the 1-lepton channel suffers from the huge amount of  $t\bar{t}$ +jets background events which exceeds the  $t\bar{t}Z$  signal by a factor of roughly 1000. However, a major advantage is the enhanced amount of signal events due to the high branching ratio of the hadronic decay of the  $Z$  boson. Both signal and background processes are modelled with Monte Carlo simulations, which have been scaled to an integrated luminosity of  $36.1 \text{ fb}^{-1}$  corresponding to the total proton-proton collision data taken by ATLAS in the years 2015 and 2016.

First, a dedicated event selection has been designed, mainly based on jet multiplicities, in order to enrich signal events. Afterwards, the  $t\bar{t}Z$  system has been kinematically reconstructed using the KLFFitter tool [200, 201], which permutes all possible assignments of the jets in an event to the partons of the signal signature. For each permutation a likelihood function, which depends on the transverse energies and momenta of the permuted jets, is calculated. In the end, the permutation which yields the maximum likelihood value for a given event is assumed to be the physical correct one. For this purpose, the KLFFitter likelihood function for the  $t\bar{t}$  lepton+jets channel has been modified by adding a further mass term for the  $Z$  boson. The KLFFitter algorithm has been tested on its performance revealing a decreasing capability of correctly relating the jets to the final state quarks they actually belong to with an increasing number of  $b$ -tagged jets in an event. A possible improvement of the reconstruction is to not only consider kinematic quantities as the transverse momentum or the energy of a jet, but also to take angular variables into account.

A total of four signal regions has been defined, which depend on the number of jets and  $b$ -tagged jets in an event, respectively. In the first signal region, denoted by SR1, events must have at least six jets irrespective of the  $b$ -jet multiplicity, whereas the second signal region, labelled SR2, additionally requires at least one  $b$ -jet. Furthermore, two exclusive regions have been defined requiring exactly two  $b$ -jets, referred to as SR3 and SR4, respectively. Events of SR3 have exactly six jets and events of SR4 are required to have exactly seven jets.

Within the signal regions variables which discriminate between the signal and the huge  $t\bar{t}$ +jets background have been studied. These contain simple event-based variables like the sum of the transverse momenta of all jets in an event,  $H_T$ , as well as quantities from the reconstruction, for example the invariant masses of the reconstructed top quarks and the  $Z$  boson. In addition, more sophisticated variables as angular relations between certain objects in specific rest frames have been explored. As none of the variables which have been found to give a reasonable separation has a sufficient discrimination power for a single-variable fit or to set an appropriate cut, all variables are combined into one output variable of a multivariate classifier. For this purpose, the Toolkit for Multivariate Analyses (TMVA) [205] has been used. In this analysis a so-called boosted decision tree (BDT) has been trained on a set of 21 and 17 input variables for the two inclusive and exclusive signal regions, respectively. When growing the decision tree, the algorithm performs yes-/no-decisions by setting cuts on the most discriminating variable at each splitting point. Based on the respective set of input variables, the BDT classifiers have been optimised with respect to their ROC values by the tuning of several setting parameters. The ROC value is a measure for the discrimination power of a BDT and defined as the integral of the area beneath the

receiver operator characteristic (ROC) curve, which shows the relation between signal efficiency and background rejection. No overtraining, i.e. training on statistical fluctuations of the input datasets, could be observed. After the evaluation of the trained classifiers a cut on the BDT output has been defined yielding the maximum significance value. The statistical significance  $Z$ , i.e. the probability that an observed excess is caused by the presence of signal and is not originating from statistical fluctuations of the background, has been calculated according to

$$Z = \frac{S}{\sqrt{S + B + \Delta B}} \quad , \quad (10.1)$$

where  $S$  and  $B$  denote the number of signal and background events and  $\Delta S$  and  $\Delta B$  represent their statistical errors, respectively. With the help of boosted decision trees, the significance could be increased by at most 18 % with respect to the cut-based analysis for the inclusive signal regions SR1 and SR2, whereas the increase is limited to 12 % for the exclusive signal regions SR3 and SR4. In case of an assumed systematic uncertainty of 30 % the significance of the respective signal regions is lowered by roughly 12 %.

Furthermore, in order to handle strongly correlated input variables given to the BDT, a decorrelation procedure has been explored. The here applied PCA decorrelation transforms the input datasets into the eigenspace of variances. As this has not enhanced the sensitivity for the  $t\bar{t}Z$  signal at all, variable decorrelation via PCA is not further pursued. However, another decorrelation algorithm might deliver better results, but due to the limited application spectrum of TMVA this cannot be realised in the moment.

In Tab. 10.1 a summary of the results for all signal regions before and after the BDT evaluation with an assumed systematic uncertainty of 30 % is given. It also includes the outcome for the BDT evaluation at  $100 \text{ fb}^{-1}$ , which corresponds to the expected amount of data which the ATLAS detector will have recorded by the end of the LHC Run 2. According to the results in Tab. 10.1, SR2 has throughout turned out to deliver the best results. Applying only the selection criteria of the respective signal regions results in a maximum significance of somewhat above  $2\sigma$  in SR2 for  $36.1 \text{ fb}^{-1}$ . This could be increased by the training and evaluation of a BDT to almost  $2.5\sigma$  which is still below evidence. However, a more promising result is expected when using the whole Run 2 dataset resulting in a significance above  $4\sigma$ . Though, assuming 30 % of systematic uncertainties might be too optimistic, the amount of uncertainty to still have evidence ( $3\sigma$ ) has been estimated yielding an allowed uncertainty of 140 % on the background which seems to be achievable.

		SR1	SR2	SR3	SR4
$36.1 \text{ fb}^{-1}$	cut-based	1.939	2.119	0.871	0.858
	BDT	2.305	2.495	0.981	0.957
$100 \text{ fb}^{-1}$	cut-based	3.229	3.526	1.450	1.427
	BDT	3.836	4.152	1.633	1.593

Tab. 10.1.: Summary of the significances achieved by the cut-based analysis and by evaluating the trained BDT classifiers at an optimal cut on the BDT response for both  $36.1 \text{ fb}^{-1}$  and  $100 \text{ fb}^{-1}$ . The calculation of the statistical significance according to Eq. (10.1) has been done assuming a systematic uncertainty of 30 % at this stage of the analysis. The significance values are commonly given in units of the gaussian standard deviation  $\sigma$ .

In a final step a qualitative investigation of the systematic uncertainties with the most significant impact on the analysis has been performed. Experimental uncertainties on the electron and

---

muon measurement as well as uncertainties on the jet vertex tagging procedure and the pile-up reweighting have been found to be negligibly small. The only experimental uncertainties which have a noticeable impact on the input variables of the BDT are three parameters from  $b$ -tagging calibration. However, they can be absorbed into statistical fluctuations after the evaluation of the multivariate classifier. Uncertainties on the jet energy scale and the jet energy resolution which are assumed to significantly influence the final results are not yet taken into account. On the contrary, theoretical uncertainties originating from the comparison of different Monte Carlo generators for the modelling of specific physical processes like parton showers and QCD radiation in the top quark backgrounds are not reduced in BDT output compared to the input variables. The shape and rate uncertainties on the BDT output due to different amounts of QCD radiation and due to the use of either diagram removal and diagram subtraction in the single top  $Wt$ -channel, amounts roughly 20 %, respectively. However, since single top quark production is not the dominant background, its uncertainties should not have a too big influence. Uncertainties of 40 % from the comparison of different parton shower modelling algorithms are slightly reduced from around 40 % for the input variables down to approximately 30 % in the BDT output, exhibiting almost no impact on the shape. In contrast, the uncertainties on the BDT output due to the use of different Monte Carlo generators for the hard scatter interaction yields quite smaller deviations of roughly 10 %. A special behaviour could be noticed for the third signal region, where the impact of additional QCD radiation and of different Monte Carlo generators for the simulation of the hard scatter interaction is very small compared to the other signal regions. Finally, uncertainties on the theoretical calculations of the production cross sections of the  $t\bar{t}Z$  signal and the dominant  $t\bar{t}Z$ +jets background are found to play a minor part due to the quite precise calculations. Altogether, the systematic uncertainties originating from theoretical issues have been turned out to have the largest impact on the distributions of both the input variables and the BDT response itself.

In order to reach  $5\sigma$ , improvements have to be done, of course. Most urgently an improvement of the reconstruction of the  $t\bar{t}Z$  system is necessary. Instead of the bare cut on the BDT output, a fit of background and signal to the BDT distribution has to be performed including all relevant systematic uncertainties. An important task will be the reduction of experimental and theoretical systematic uncertainties exploiting instruments, for example particle-flow jets [213]. In particular, an improved simulation of additional partons by the Monte Carlo generators for the dominant backgrounds should lower the uncertainties on QCD radiation and possibly also parton shower modelling. Additionally, a (data-driven) estimation of the backgrounds arising from QCD multijet production as well as from the associated production of  $W$  bosons with jets has to be done.

However, with this analysis the foundations for the measurement of the production cross section of top quark pairs in association with a  $Z$  boson in the 1-lepton channel have been laid. The application of multivariate analysis techniques has been proven to be beneficial, revealing a considerable increase in significance. Although the amount of data yet collected seems not to be sufficient, for the whole Run 2 dataset at least evidence can be expected allowing for a huge systematic uncertainty of 140 %.



## A. Top quark pair production at the LHC

As explained in Sec. 2.3.1, there are two different mechanisms to produce top quarks: The single top quark production via the electroweak interaction and the top quark pair production via the strong interaction. At hadron colliders as the LHC, the strong force is the main source for top quarks at the LHC. This happens either through the annihilation of a quark and an antiquark or the fusion of two gluons. In order to determine which one of the two alternatives is the dominant process at the LHC at  $\sqrt{s} = 13$  TeV, the momentum fractions the incoming partons have to carry to produce a top quark pair system proves to be beneficial. The square of the centre-of-mass energy of two colliding partons is equal to the squared sum of their four-momenta,

$$s = (p_a + p_b)^2 \quad , \quad (\text{A.1})$$

where the momentum of a parton  $i$  can be expressed in terms of the fraction  $x_i$  of the proton's momentum. In the relativistic limit the momentum of the proton is equal to the beam energy, so the momenta of the two partons can be rewritten as

$$p_a = \begin{pmatrix} x_a \\ 0 \\ 0 \\ x_a \end{pmatrix} \cdot E_{\text{beam}} \quad p_b = \begin{pmatrix} x_b \\ 0 \\ 0 \\ x_b \end{pmatrix} \cdot E_{\text{beam}} \quad . \quad (\text{A.2})$$

In order to produce a top quark pair,  $\sqrt{s}$  has to be at least equal to the rest mass of the top quark pair system,

$$s = (p_a + p_b)^2 \geq (2m_{\text{top}})^2 \quad . \quad (\text{A.3})$$

For simplicity it is assumed that both partons carry a similar momentum fraction ( $x_a \approx x_b$ ). The momentum fraction required to create a top quark pair is thus given by

$$x \geq \frac{2m_{\text{top}}}{\sqrt{s}} = \frac{m_{\text{top}}}{E_{\text{beam}}} \quad (\text{A.4})$$

with  $\sqrt{s} = 2E_{\text{beam}}$  denoting the centre-of-mass energy of the LHC collider. For a given centre-of-mass energy of 13 TeV and an assumed top quark mass of 172.5 GeV the momentum fraction a parton has to carry is about 0.027. Since there are much more gluons than (sea) quarks at such low momentum fractions in a proton (cf. Fig. 2.4), the main top quark pair production mechanism at the LHC is the gluon-gluon fusion, contributing with approximately 90%. For the associated top quark pair production with a  $Z$  boson the momentum fraction a parton has to carry is about 0.0035 since the mass of a  $Z$  boson has in addition to be considered. However, the gluon-gluon fusion is still by far the dominant production mechanism.

The total number of top quark pairs at the LHC can be calculated from the respective production cross section and the total integrated luminosity, using (3.4). With a luminosity of  $36.1 \text{ fb}^{-1}$  and the theoretical top quark pair production cross section of  $\sigma_{t\bar{t}} = 832 \text{ pb}$  (cf. Sec. 2.3.1.2), this corresponds to roughly 30 million top quark pair events produced in the 2015+2016 data taking period. Similarly, approximately 30 000 top quark pairs in association with a  $Z$  boson have been produced, reflecting the thousandfold suppression of this process with respect to the bare top quark pair production.



## B. The KL Fitter transfer functions

The detector response caused by a traversing particle is in general a non-symmetric distribution around the true value. In particular, the calorimeter tends to rather measure less energy than actually deposited, e.g. due to dead material or within specific regions with a poor coverage. This depends of course on the detector geometry, but also on the energy of the incoming particle. Hence, the detector response depends on both the energy and an angular component, e.g.  $\eta$ . Therefore the detector resolution is generally parametrised by so-called transfer functions  $W(\tilde{E}_i|E_i)$ , which give the probability density of a measured value given the true value in different  $\eta$  and energy bins. In other words, a transfer function is defined as the conditional probability  $W(x_{\text{reco}}|x_{\text{truth}})$  to obtain a certain response after reconstruction,  $x_{\text{reco}}$ , given the true value,  $x_{\text{truth}}$ . They are normalised to unity,

$$\int dx_{\text{reco}} W(x_{\text{reco}}|x_{\text{truth}}) = 1 \quad , \quad (\text{B.1})$$

meaning that given a true value there will always be a reconstructed one [214].

The parametrisation of the transfer functions for the energies of the measured particles is done separately for different types of objects distinguished in electrons, muons, light jets,  $b$ -jets and neutrinos (respectively missing transverse energy). They are modelled by double Gaussian functions,

$$W(\Delta E) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \left[ \exp\left(-\frac{(\Delta E - p_1)^2}{2p_2^2}\right) + p_3 \exp\left(-\frac{(\Delta E - p_4)^2}{2p_5^2}\right) \right] \quad (\text{B.2})$$

with

$$\Delta E = \frac{E_{\text{truth}} - E_{\text{reco}}}{E_{\text{truth}}} \quad , \quad (\text{B.3})$$

where the parameters  $p_i$  are functions of the particle's true energy  $E_{\text{truth}}$ <sup>(1)</sup> and parametrised according to the detector resolution. Depending on the parameter, this dependence can be either proportional to the inverse of the square-root of the true energy as for parameter  $p_2$  with

$$p_2 = \frac{a_2}{E_{\text{truth}}} + b_2 \quad , \quad (\text{B.4})$$

or linear in the true energy of the particle as for all other parameters, i.e.

$$p_i = a_i + b_i \cdot E_{\text{truth}} \quad . \quad (\text{B.5})$$

With these assumptions, ten parameters  $a_i$  and  $b_i$  have to be derived separately for each object. For the neutrino simple Gaussian functions are used.

The transfer functions are derived from reconstructed objects which are matched to the corresponding truth particles requiring a unique match with  $\Delta R(x_{\text{reco}}, x_{\text{truth}}) = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$ . For the derivation of the transfer functions a two-dimensional binned likelihood fit within three different  $\eta$ -regions is used. One dimension corresponds to the energy of the reconstructed objects and the other dimension is the relative difference of the energies  $\Delta E$ , where the latter is calculated using events simulated by Monte Carlo. The  $\eta$ -regions considered are  $(0.0 < |\eta| < 1.0)$ ,

<sup>(1)</sup> The unit for  $E_{\text{truth}}$  used in the parametrisation is GeV.

( $1.0 < |\eta| < 1.7$ ) and ( $1.7 < |\eta| < 2.5$ ), and are motivated by the detector geometry. The energy of the reconstructed objects is subdivided into ten bins, depending on the statistics in a single bin in order to minimise statistical uncertainties. For each bin of energy, the relative energy deviation  $\Delta E$  is then fitted with a double Gaussian using a first loose range of fit parameters. Afterwards, the parameters  $p_i$  of the double Gaussians are further fitted in global fit depending on the truth energy. These globally fitted parameters are used to define the transfer functions (cf. equation (B.2)). As an example, Fig. B.1 shows the relative energy deviations obtained from Monte Carlo simulations as well as the local double Gaussian fit, its two components and the result of the global fit for light quarks in B.1a and  $b$ -quarks in B.1b.

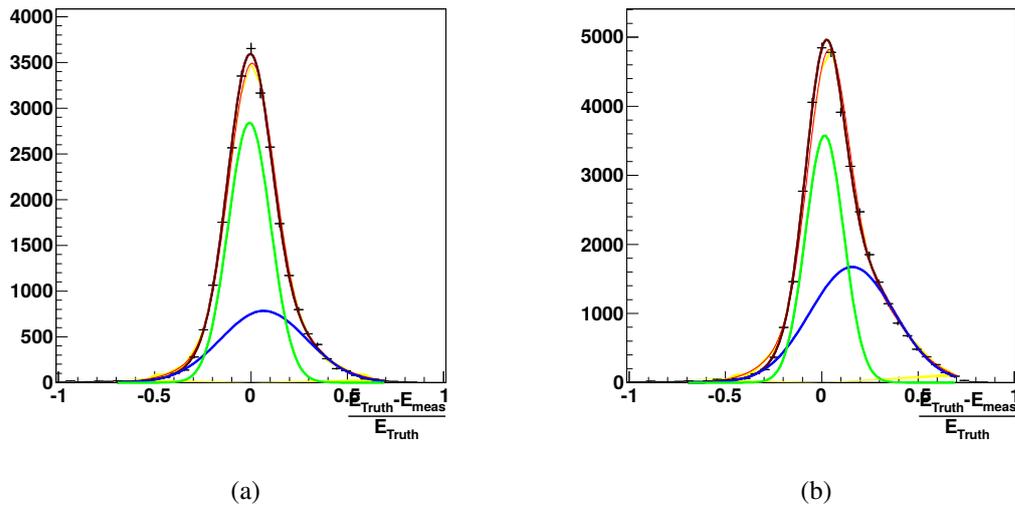


Fig. B.1.: In both plots the parametrisation of the transfer functions with a double Gaussian (brown), its subcomponents divided into the two Gaussians (green, blue) and the global fit (red) with the error band (yellow) can be seen. The Monte Carlo derived relative energy deviation is depicted by the black crosses, implicating statistical errors. In Fig. (a) the transfer function of light quarks in  $\eta$ -range ( $0.0 < |\eta| < 1.0$ ) and energy range ( $130 \text{ GeV} < E < 160 \text{ GeV}$ ) is shown, whereas Fig. (b) depicts the transfer function of bottom quarks in  $\eta$ -range ( $0.0 < |\eta| < 1.0$ ) and energy range ( $130 \text{ GeV} < E < 160 \text{ GeV}$ ) [215].

## C. Reconstruction of the neutrino momentum

An accurate knowledge of the full neutrino momentum is necessary when entirely reconstructing the  $t\bar{t}Z$  system. As mentioned in section 6.3.1, due to momentum conservation in the transverse plane, only the  $x$ - and  $y$ -components of the neutrino momentum are known by the measurement of the missing transverse momentum.

For the kinematic fit, boundaries are set to the different entries of the neutrino momentum vector, namely  $p_x^\nu$  as well as  $p_y^\nu$  are constrained to stay in the range of  $-100 \text{ GeV} < p_i^\nu < 100 \text{ GeV}$ , whereas for  $p_z^\nu$  the boundaries are extended to  $\pm 1000 \text{ GeV}$ . As a starting value for the transverse momentum components the measured missing energy is used, i.e.  $p_x^\nu = E_x^{\text{miss}}$  and  $p_y^\nu = E_y^{\text{miss}}$ . During the fit they are further constrained by Gaussian transfer functions (cf. appendix B), which confine them to the range of  $E_{x,y}^{\text{miss}} - \sigma < p_{x,y}^\nu < E_{x,y}^{\text{miss}} + \sigma$ , where  $\sigma$  is the deviation due to the finite width of the transfer functions.

However, the  $z$ -component of the neutrino momentum cannot be directly measured, wherefore it has to be calculated separately. The neutrino originates from the leptonic decay of  $W$  boson, thus the invariant mass of the leptonically decaying  $W$  boson is used for the calculation.

Using four-momentum conservation,

$$p^W = p^\nu + p^l \quad . \quad (\text{C.1})$$

Squaring equation (C.1),

$$\begin{aligned} m_W^2 &= (p^\nu + p^l)^2 \\ &= (p^\nu)^2 + (p^l)^2 + 2p^\nu p^l \\ &= m_\nu^2 + m_l^2 + 2(E^\nu E^l - \vec{p}^\nu \vec{p}^l) \\ &\approx m_l^2 + 2(E^\nu E^l) - 2(p_x^\nu p_x^l + p_y^\nu p_y^l + p_z^\nu p_z^l) \quad , \end{aligned} \quad (\text{C.2})$$

where  $m^2 = E^2 - \vec{p}^2 = p^2$  and  $m_\nu \approx 0$  had been made use of.

Defining

$$\alpha := m_W^2 - m_l^2 + 2(p_x^\nu p_x^l + p_y^\nu p_y^l) \quad , \quad (\text{C.3})$$

equation (C.2) can be rewritten as

$$\alpha = 2E^l \sqrt{(p_x^\nu)^2 + (p_y^\nu)^2 + (p_z^\nu)^2} - 2p_z^l p_z^\nu \quad , \quad (\text{C.4})$$

which, after rearranging its terms, reads

$$\frac{\alpha}{2} + p_z^l p_z^\nu = E^l \sqrt{(p_x^\nu)^2 + (p_y^\nu)^2 + (p_z^\nu)^2} \quad . \quad (\text{C.5})$$

Now, squaring equation (C.5) gives

$$\frac{\alpha^2}{4} + (p_z^l)^2 (p_z^\nu)^2 + \alpha \cdot p_z^l p_z^\nu = (E^l)^2 [(p_x^\nu)^2 + (p_y^\nu)^2 + (p_z^\nu)^2] \quad , \quad (\text{C.6})$$

or, sorting after  $p_z^\nu$  ,

$$(p_z^\nu)^2 \left[ (p_z^l)^2 - (E^l)^2 \right] + p_z^\nu \left[ \alpha \cdot p_z^l \right] + \frac{\alpha^2}{4} - (E^l)^2 \left[ (p_x^\nu)^2 + (p_y^\nu)^2 \right] \quad . \quad (\text{C.7})$$

With this quadratic formula of the form

$$A \cdot x^2 + B \cdot x + C = 0 \quad (\text{C.8})$$

the following substitutions are made:

$$A := (p_z^l)^2 - (E^l)^2 \quad (\text{C.9a})$$

$$B := \alpha \cdot p_z^l \quad (\text{C.9b})$$

$$C := \frac{\alpha^2}{4} - (E^l)^2 \left[ (p_x^\nu)^2 + (p_y^\nu)^2 \right] \quad (\text{C.9c})$$

Thus, equation (C.7) is solved by

$$p_z^\nu = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \quad . \quad (\text{C.10})$$

Since the  $z$ -component of neutrino momentum is only quadratically constrained by the  $W$  boson mass (cf. equation (C.2)), the two solutions are equivalent. During all the calculations the mass of the  $W$  boson is treated as a constant as it is well-known from previous measurements.

Equation (C.10) has either zero, one or two solutions depending on the value of the discriminant

$$D := B^2 - 4AC \quad . \quad (\text{C.11})$$

In case of  $D$  being negative, the solutions are imaginary and therefore not taken into account for further calculations. For very small values of typically  $\sqrt{D} < 1 \cdot 10^{-6}$  the whole expression under the square root is set to zero. So only for reasonable positive values of  $D$  both solutions are taken into account. In that case, the maximisation of the likelihood<sup>(1)</sup> is done for both solutions. The one giving rise to the higher likelihood value is finally taken as the  $z$ -component of the neutrino momentum.

At this point everything is prepared and the actual fit begins. In order to receive the highest possible likelihood value of the single event besides all parton energies, the  $x$ - and  $y$ -components of the neutrino momentum are smeared according to their Gaussian transfer function. Since no transfer function exists for  $p_z^\nu$ , it is varied in a way that the mass of the  $W$  boson stays constant under smearing  $p_x^\nu$  and  $p_y^\nu$ .

---

<sup>(1)</sup> Or, equivalently, minimisation of the negative logarithm of the likelihood

## D. Boosting algorithms in TMVA

Boosting is a way of enhancing classification performance of a weak multivariate methods by sequentially applying the multivariate algorithm to the reweighted (“boosted”) training data, i.e. each tree is trained on a different reweighting of the training dataset. This procedure also increases the stability with respect to statistical fluctuations within the training sample. Introduced in the early '90s, it dramatically increases the performance of classification techniques in most cases [205].

### D.1. AdaBoost

The most popular boosting algorithm is the so-called AdaBoost [216, 217] (adaptive boost). Each object in the training set is labelled either with  $y_i = +1$  or  $-1$  and has an initial weight  $w_{(1),i}$ ,  $i = 1, \dots, N$  set to  $\frac{1}{N}$ , where the sum of weights is normalised to unity,

$$\sum_{i=1}^N w_{(1),i} = 1 \quad . \quad (\text{D.1})$$

The idea is to assign higher weights to misclassified events from the recent training, i.e. give them higher priority for the selection cuts in the training of the following tree. Starting with the original event weights when training the first decision tree  $f(\mathbf{x}; w_{(1)})$  with  $\mathbf{x}$  being a tuple of input variables, the subsequent tree is then grown from the the same events as before, but with the new event weights including a boost weight  $\alpha_k$ . The boost weight  $\alpha_k$  is derived from the misclassification rate  $\varepsilon$  from the previous tree  $f(\mathbf{x}; w_{(k)})$ ,

$$\alpha_k := \frac{1}{2} \ln \left( \frac{1 - \varepsilon}{\varepsilon} \right) \quad , \quad (\text{D.2})$$

and is used to derive the new event weights for the subsequent training,

$$w_{(k+1),i} \leftarrow w_{(k),i} \cdot \exp \left[ -\frac{\alpha_k f(\mathbf{x}_i; w_{(k)}) y_i}{Z_k} \right] \quad , \quad (\text{D.3})$$

where  $Z_k$  is a normalisation factor chosen in order to keep the sum of weights constant,

$$\sum_{i=1}^N w_{(k+1),i} = 1 \quad . \quad (\text{D.4})$$

The final BDT is then the weighted average of all individual decision trees,

$$F(\mathbf{x}) = \sum_{k=1}^K \alpha_k \cdot f(\mathbf{x}; w_{(k)}) \quad . \quad (\text{D.5})$$

Defining the result of an individual classifier as  $h(\mathbf{x})$  encoded for signal and background as  $h(\mathbf{x}) = +1$  and  $-1$ , respectively, the boosted event classification is then given by the sum over all

classifiers in the collection,

$$H(\mathbf{x}) = \sum_{k=1}^K \alpha_k \cdot h_k(\mathbf{x}), \quad (\text{D.6})$$

where small (large) values for  $H(\mathbf{x})$  indicate a background (signal) like event.

AdaBoost performs best on weak classifiers, small individual trees with a depth of two or three nodes with almost no discrimination power. The performance can often be further enhanced by enforcing a slow learning and instead allowing a large number of boost steps. The learning rate of the AdaBoost algorithm is therefore controlled by a parameter  $\beta$  which is given as an exponent to the boost weight  $\alpha \rightarrow \alpha^\beta$ .

## D.2. Gradient Boost

Another boosting technique which has been applied in this analysis, is the so-called Gradient Boost [218, 219]. The idea of estimating functions via boosting can be understood by considering an additive expansion approach. The function  $F(\mathbf{x})$  under consideration is assumed to be a weighted sum of parametrised base functions  $f(\mathbf{x}; w_k)$ , so-called weak learners. Each base function in this expansion corresponds to a decision tree

$$F(\mathbf{x}; P) = \sum_{k=0}^K a_k f(\mathbf{x}; w_k); \quad P \in \{a_k; w_k\}_0^K. \quad (\text{D.7})$$

The boosting procedure is employed to adjust the parameters  $P$  in such a way that the deviation between the model response  $F(\mathbf{x})$  and the true value  $y$  obtained from the training sample is minimised. A measure of the deviation is the loss function  $L(F, y)$ , a popular choice being a squared error loss

$$L(F, y) = (F(\mathbf{x}) - y)^2. \quad (\text{D.8})$$

The previous method, AdaBoost, is based on exponential loss,

$$L(F, y) = e^{F(\mathbf{x})y}, \quad (\text{D.9})$$

leading to the reweighting algorithm described in Section D.1, where it explicitly enters in equation (D.3). A shortcoming of exponential loss is the lack of robustness in presence of outliers or mislabelled data points. Therefore, the performance of AdaBoost is expected to worsen within a noisy environment.

In contrast, the Gradient Boost algorithm attempts to overcome this weakness by applying other, potentially more robust loss functions which in addition do not reduce the good out-of-the-box performance. Currently implemented in TMVA for the Gradient Boost is a binomial logarithmic-likelihood loss,

$$L(F, y) = \ln \left( 1 + e^{-2F(\mathbf{x})y} \right). \quad (\text{D.10})$$

The boosting algorithm that corresponds to this loss function cannot be obtained in a straightforward manner, so a steepest descent<sup>(1)</sup> approach has been chosen in order to do the minimisation. In doing so, the current gradient of the loss function is calculated. Afterwards, a regression tree<sup>(2)</sup> whose leaf values are adjusted to match the mean value of the gradient in each region, defined by the tree structure, is grown. Iterating this process leads to the desired set of decision trees

---

<sup>(1)</sup> or gradient descent, hence the name

<sup>(2)</sup> More details on regression techniques can be found in [205].

minimising the loss function. Therefore, Gradient Boost can be adapted to any differentiable loss function.

In the following a schematic of the algorithm will be given. Starting with a training set  $\{(x_i, y_i)\}_{i=1}^N$  of  $N$  input variables  $x_i$  and their corresponding output variables  $y_i$ , the goal is to find an approximation  $\hat{F}(\mathbf{x})$  to a function  $F(\mathbf{x})$  that minimises the expected (or average) value of some specified differentiable loss function  $L(F(\mathbf{x}), y)$  on the training set. The Gradient Boost method assumes a real-valued  $y$  and seeks an approximation  $\hat{F}(\mathbf{x})$  in the form of a weighed sum of functions  $h(\mathbf{x})$ , called weak learner:

$$F(\mathbf{x}) = \sum_{k=1}^K \gamma_k h_k(\mathbf{x}) + \text{const.} \quad (\text{D.11})$$

As a starting point, the model is initialised with a constant value,

$$F_0(\mathbf{x}) = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, \gamma) \quad . \quad (\text{D.12})$$

Henceforth, the model incrementally expands, i.e. for each setp  $k = 1, \dots, K$  a so-called pseudo-residual

$$r_{ik} = - \left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{k-1}(x)} \quad \text{for } i = 1, \dots, N \quad (\text{D.13})$$

is calculated to which a weak learner  $h_k(\mathbf{x})$  is fitted, meaning to train it using  $\{(x_i, r_{ik})\}_{i=1}^N$  as a training set. The multiplier  $\gamma_k$  is computed by solving the following one-dimensional optimisation problem:

$$\gamma_k = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, F_{k-1}(x_i) + \gamma h_k(x_i)) \quad (\text{D.14})$$

After each such step the model gets updated,

$$F_k(\mathbf{x}) = F_{k-1}(\mathbf{x}) + \gamma_k h_k(\mathbf{x}) \quad , \quad (\text{D.15})$$

resulting in the final outcome  $F_K(\mathbf{x})$ .

Just like AdaBoost, Gradient Boost works best on weak classifiers, small individual decision trees with a tree depth of usually at most four nodes. The algorithm's robustness can be enhanced by reducing the learning rate through the shrinkage parameter which controls the the individual tree weights. With a small shrinkage more trees have to be grown, but the accuracy of the prediction in difficult settings can be significantly improved.

Sometimes a bagging-like resampling procedure called stochastic gradient boosting [220], using random subsamples of the training events might be beneficial for the Gradient Boost algorithm for growing decision trees. This technique has been applied on top of Gradient Boost within this analysis, provided by the TMVA framework. Also the sample fraction used in each iteration can be controlled where typically the best results are obtained for values between 0.5 and 0.8.

### D.3. Bagging

Although bagging is not a genuine boosting algorithm, it will be shortly presented in the following, however, at least due to its application in the stochastic gradient boosting. With the term bagging

a resampling technique is denoted in which a classifier is repeatedly trained each time using a randomly chosen subset of the training events. All individual classifiers are then combined to an overall classifier representing an average of the individual ones, in general leading to a significant increase in performance. Within the resampling procedure events are allowed to be randomly picked several times from the parent sample. With this method primarily not an enhancement of a weak classifier in the way of boosting is sought but rather a stabilisation of the classifier's response due to the smearing over statistical representations of the training data.

Just as for boosting, training several classifiers with different resampled training data and combining them afterwards results into an averaged classifier which is more stable with respect to statistical fluctuations within the training data.

## E. Additional plots

### E.1. Signal region definition

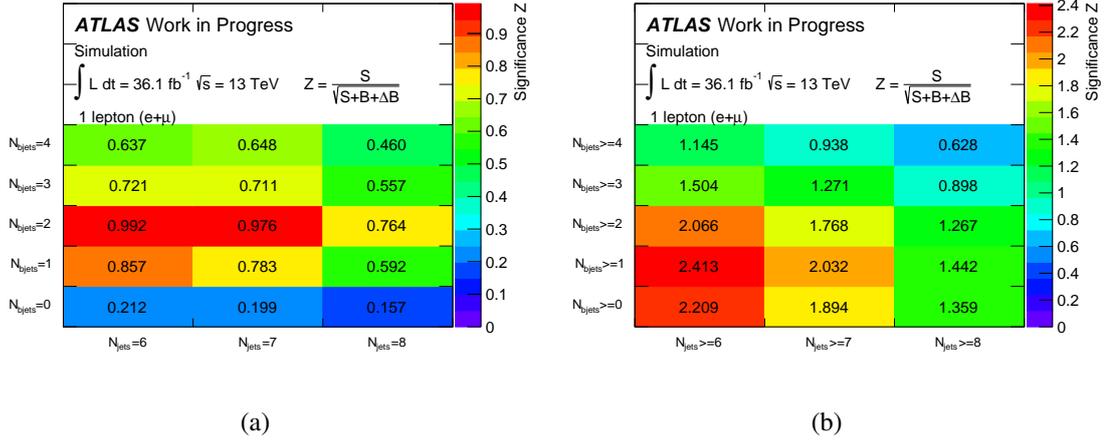


Fig. E.1.: Illustration of the scan through the phase space of ( $b$ -)jet multiplicity, divided into exclusive (cf. Fig. (a)) and inclusive (cf. Fig. (b)) selection criteria. Both images exhibit two promising regions each, in the following referred to as signal regions. The calculation of the respective significance values has been done according to the formula depicted within the images assuming only statistical uncertainties.

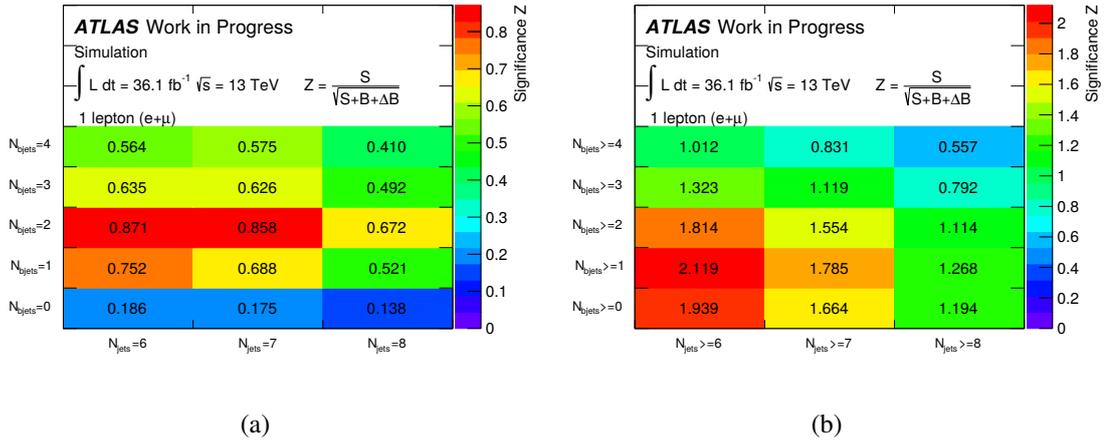


Fig. E.2.: Illustration of the scan through the phase space of ( $b$ -)jet multiplicity, divided into exclusive (cf. Fig. (a)) and inclusive (cf. Fig. (b)) selection criteria. Both images exhibit two promising regions each, in the following referred to as signal regions. The calculation of the respective significance values has been done according to the formula depicted within the images assuming both statistical and systematic uncertainties.

## E.2. Reconstruction performance

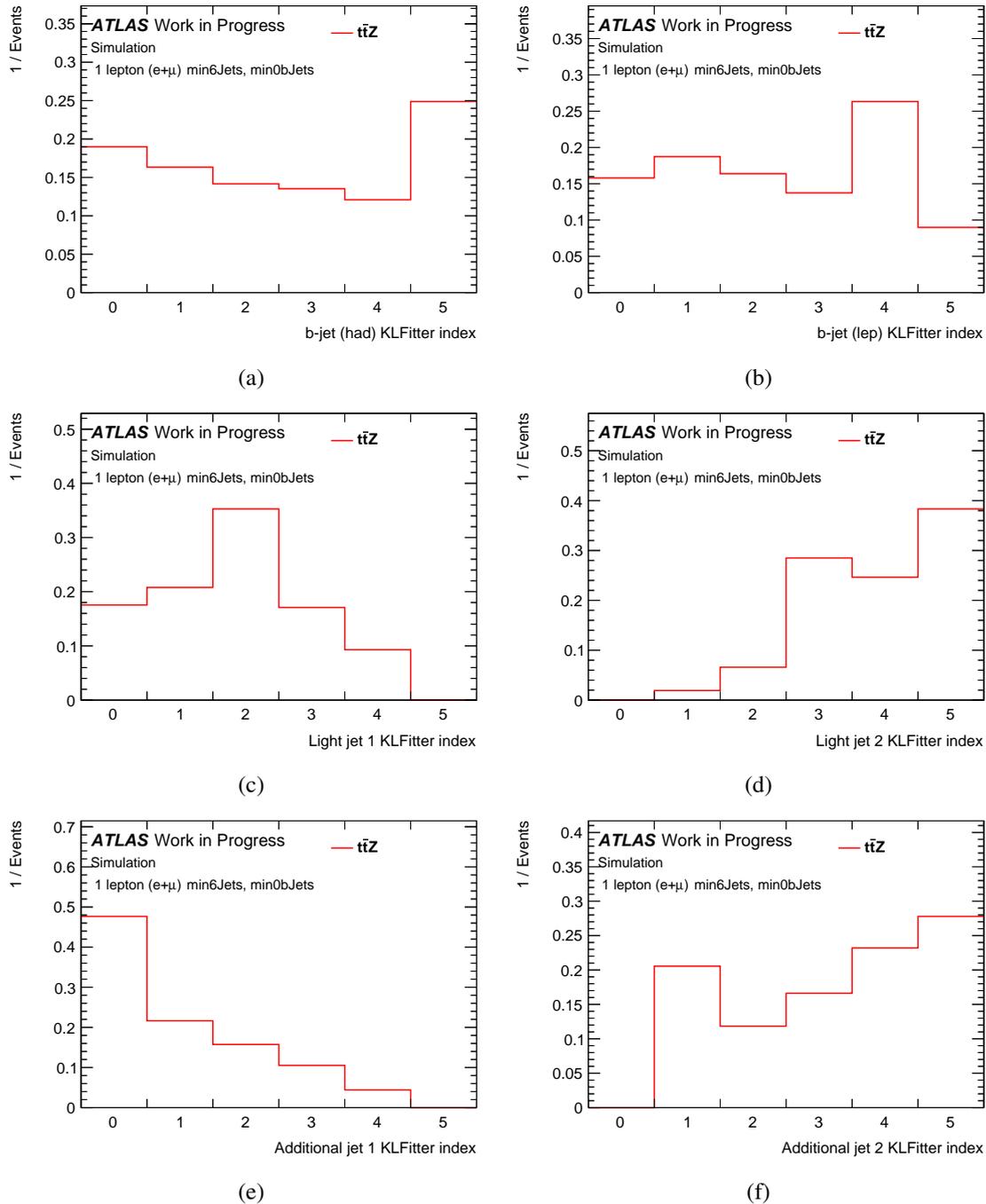


Fig. E.3.: Illustration of the KL Fitter jet indices for all final state particles in the first signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the indices of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

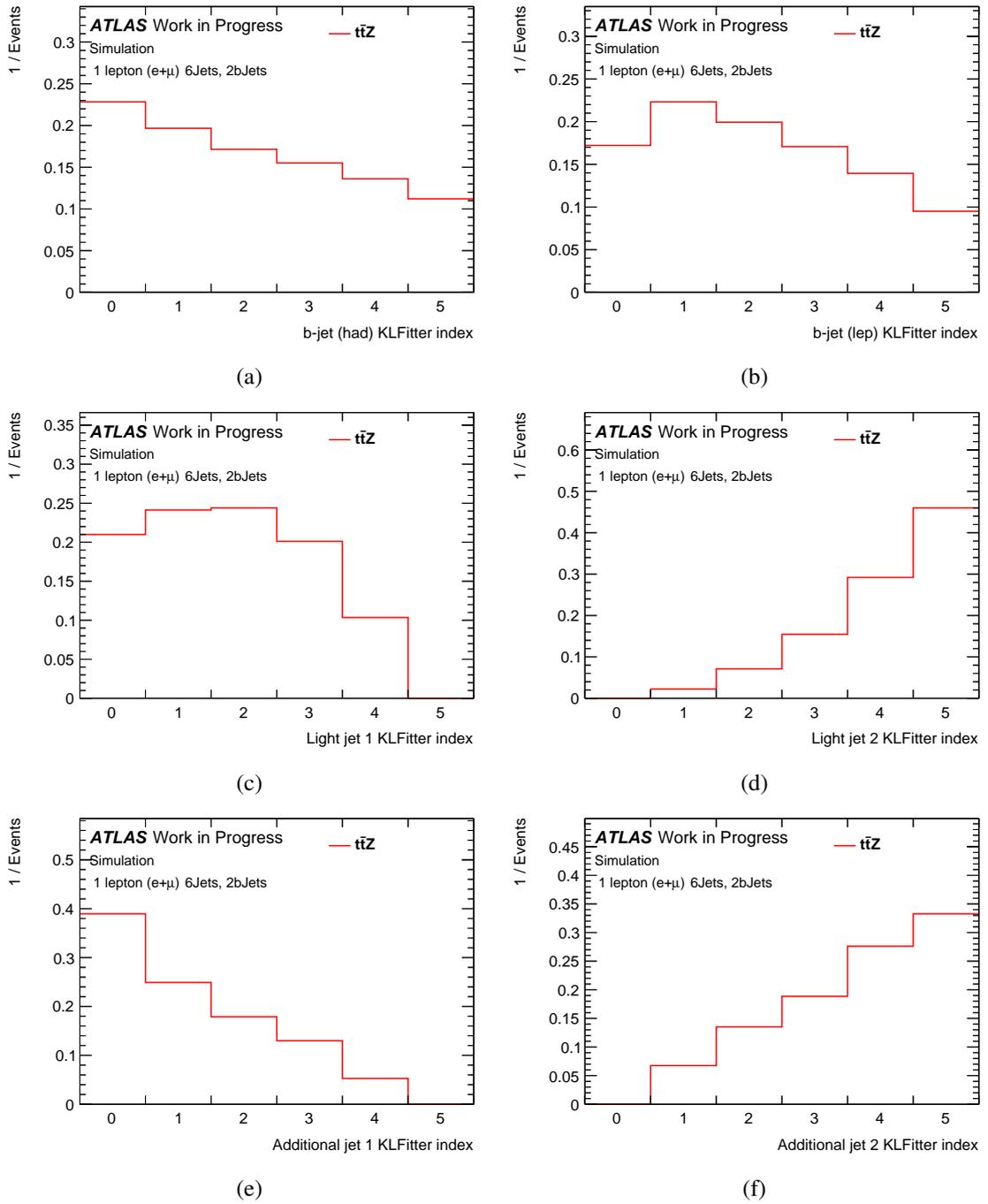


Fig. E.4.: Illustration of the KL Fitter jet indices for all final state particles in the third signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the indices of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E. Additional plots

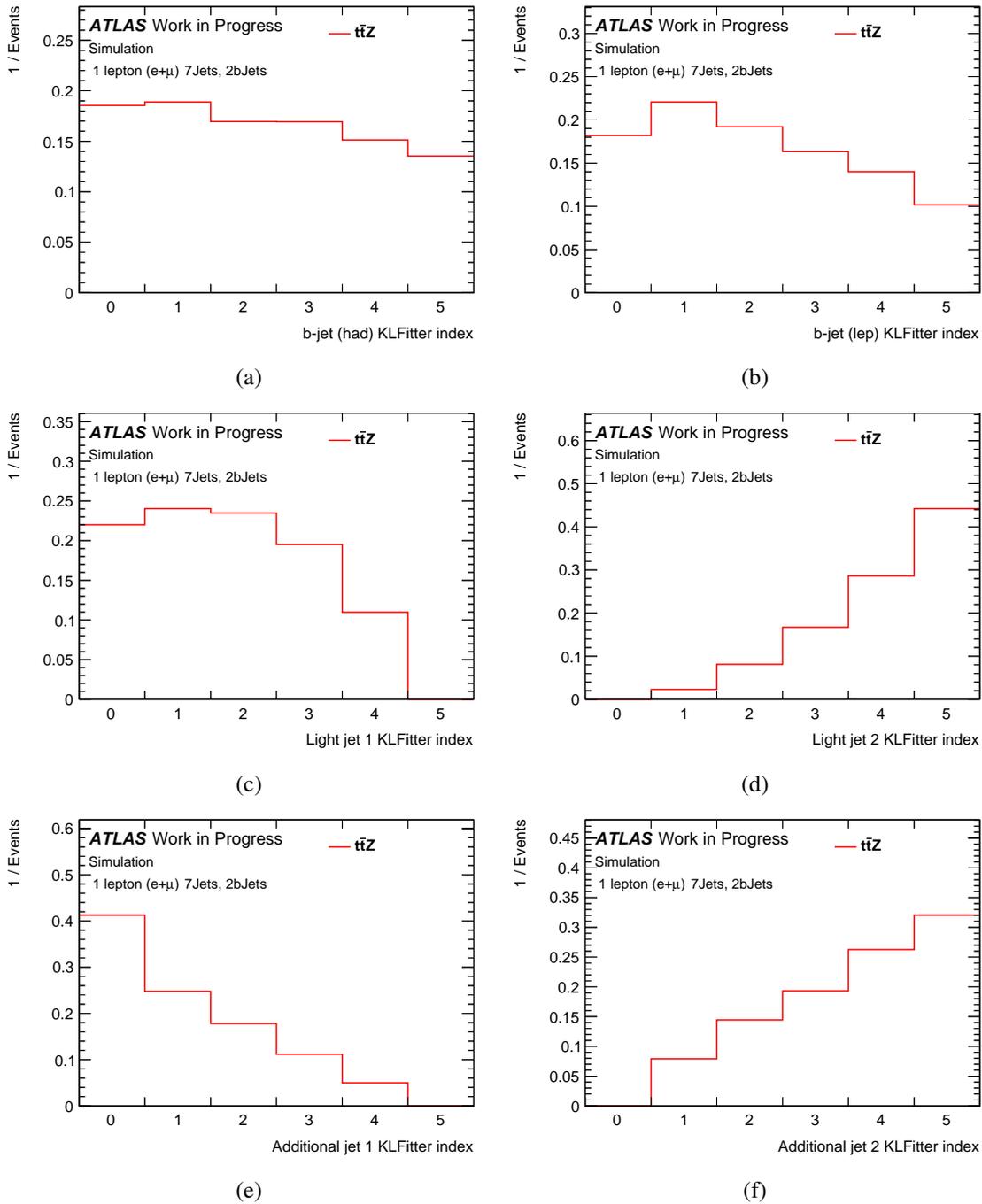


Fig. E.5.: Illustration of the KL Fitter jet indices for all final state particles in the fourth signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the indices of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

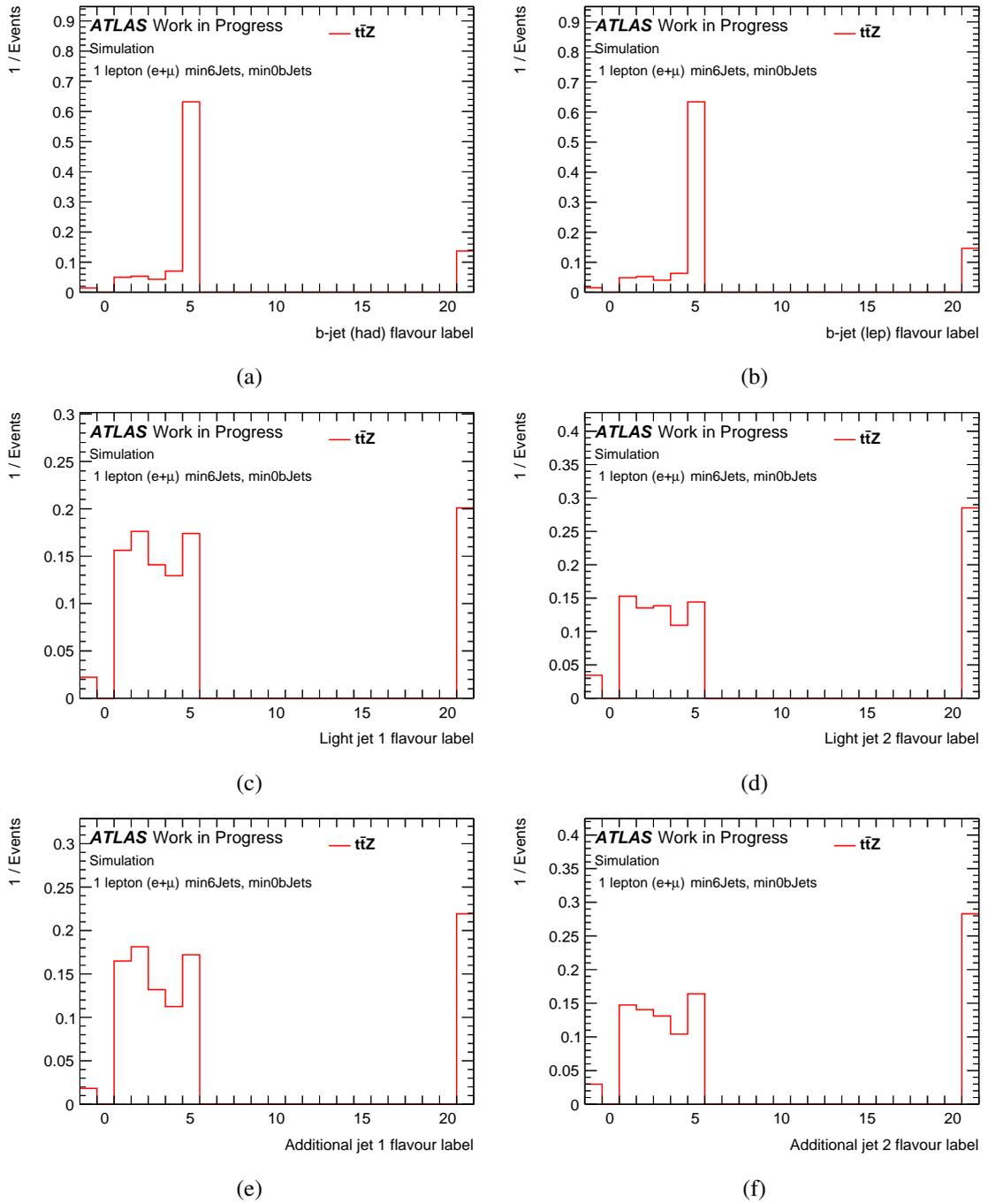


Fig. E.6.: Illustration of the jet flavour labels for all final state particles in the first signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E. Additional plots

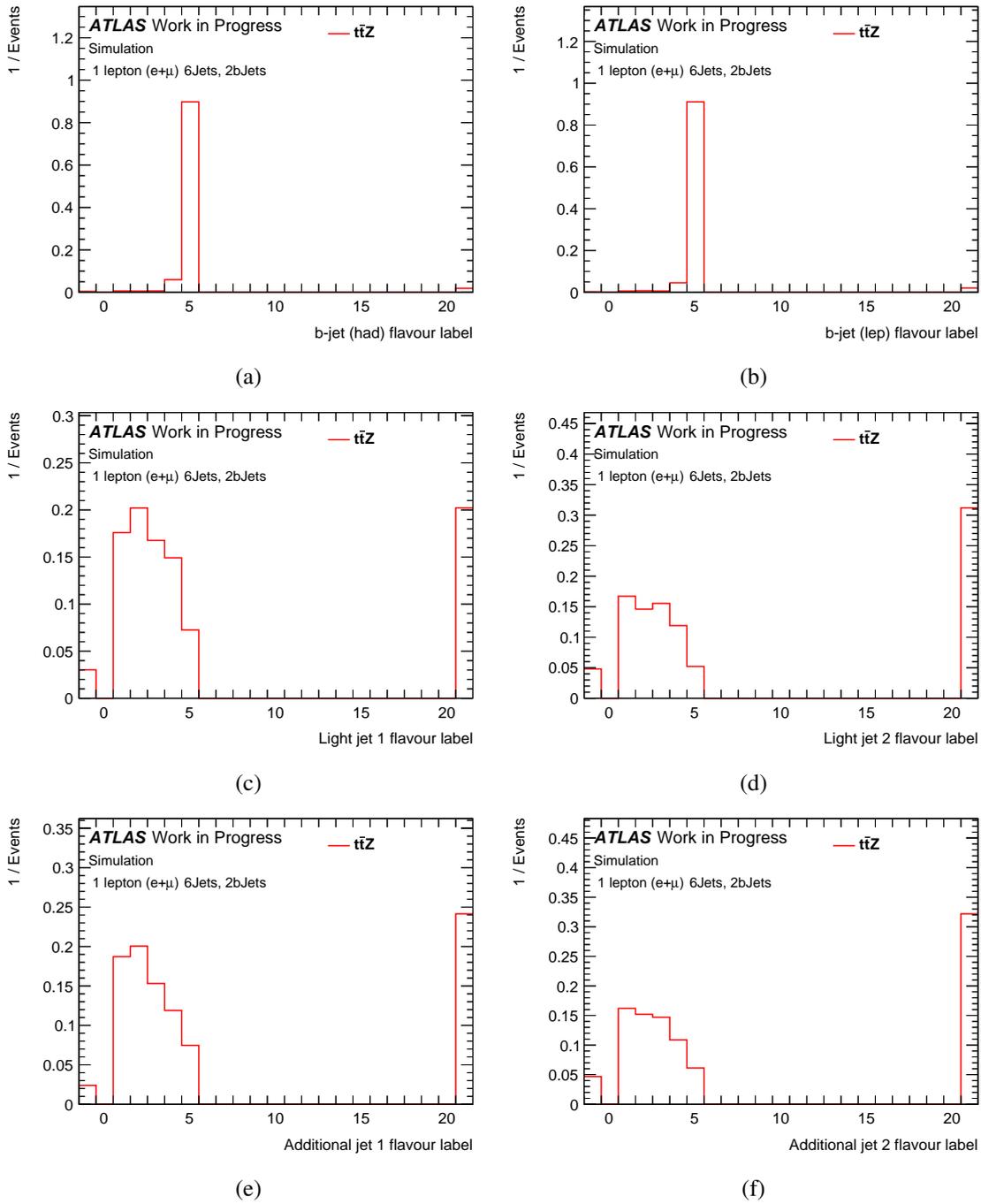


Fig. E.7.: Illustration of the jet flavour labels for all final state particles in the third signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

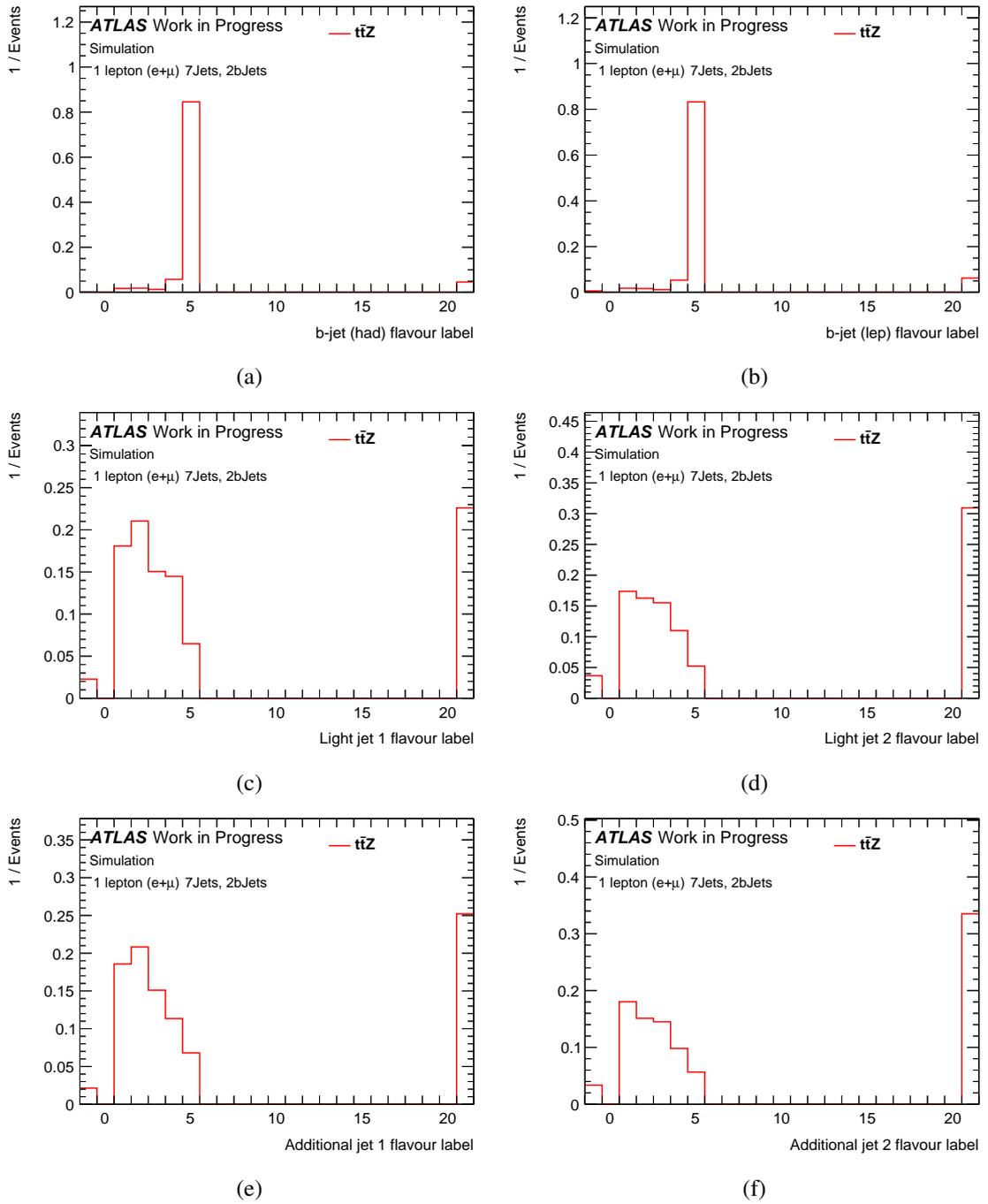


Fig. E.8.: Illustration of the jet flavour labels for all final state particles in the fourth signal region. The two upper images show the two  $b$ -jets originating from the top quark decay (cf. Fig. (a), (b)). In the middle, the labels of two jets of the light quarks from the  $W$  boson decay are depicted (cf. Fig. (c), (d)) and on the bottom the two additional jets arising from the  $Z$  boson decay are shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

### E.3. Discriminating variables for the BDT

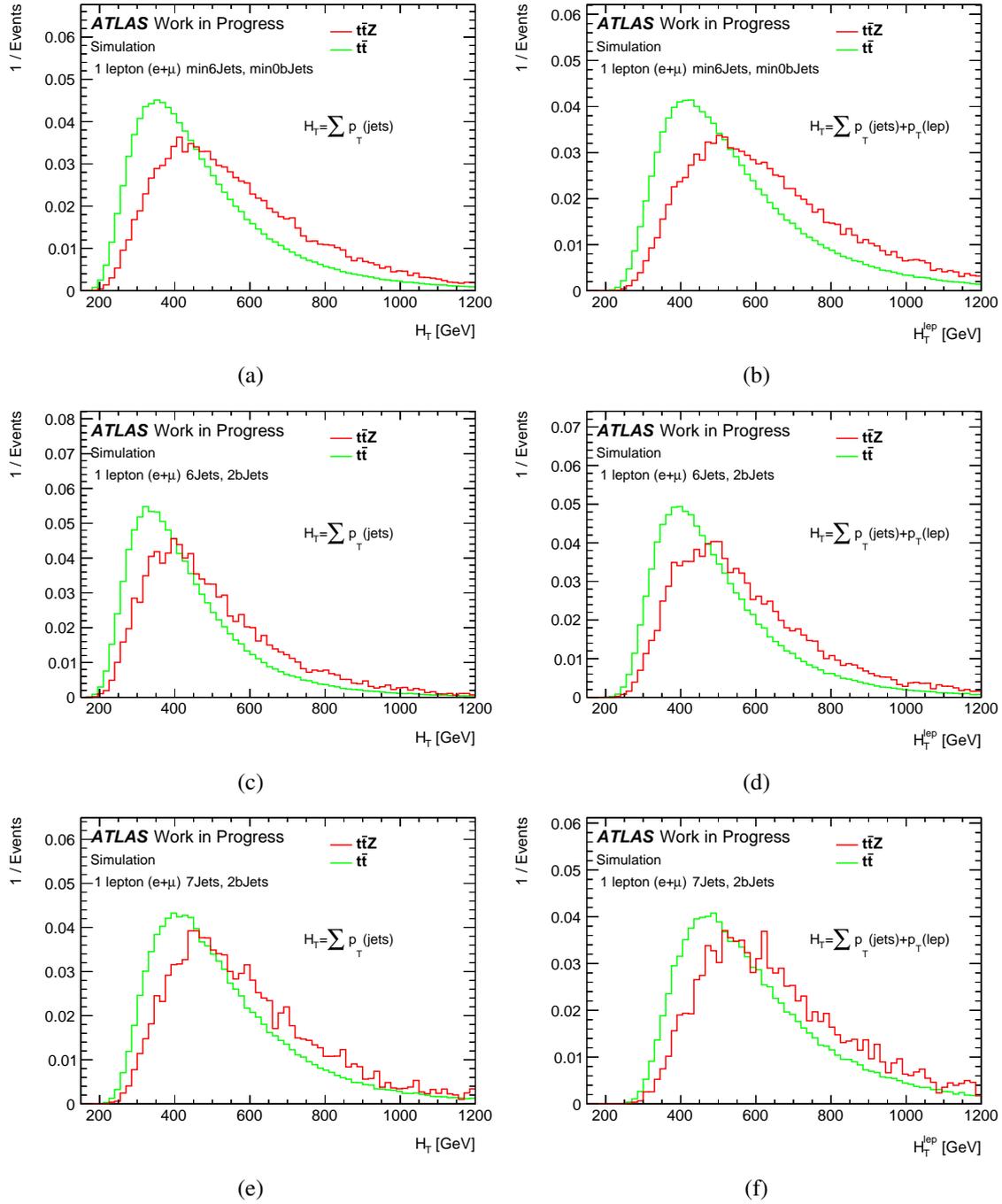


Fig. E.9.: Illustration of  $H_T$  on the left side and of  $H_T^{\text{lep}}$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

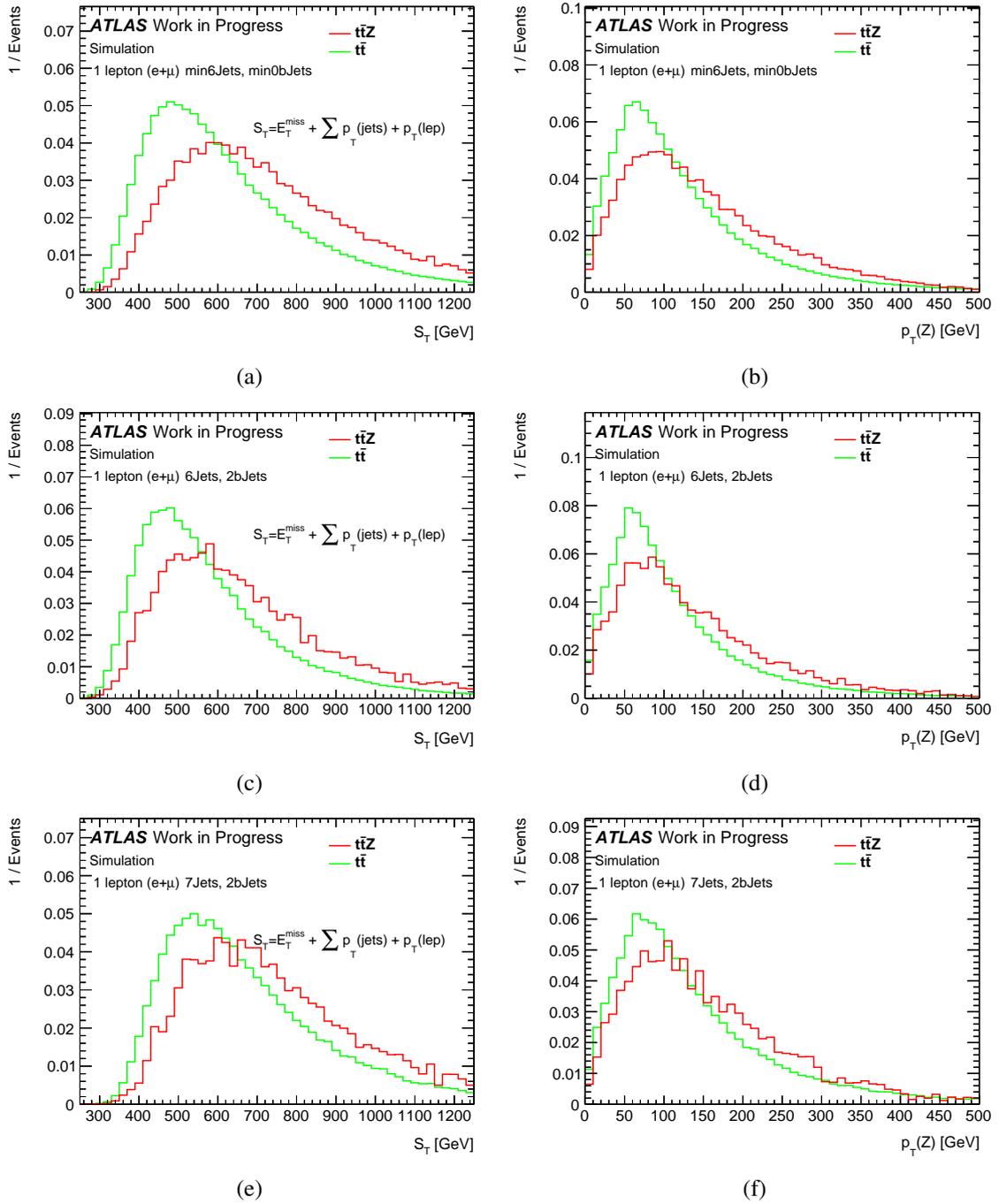


Fig. E.10.: Illustration of  $S_T$  on the left side and of  $p_T^{\text{reco}}(Z)$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E. Additional plots

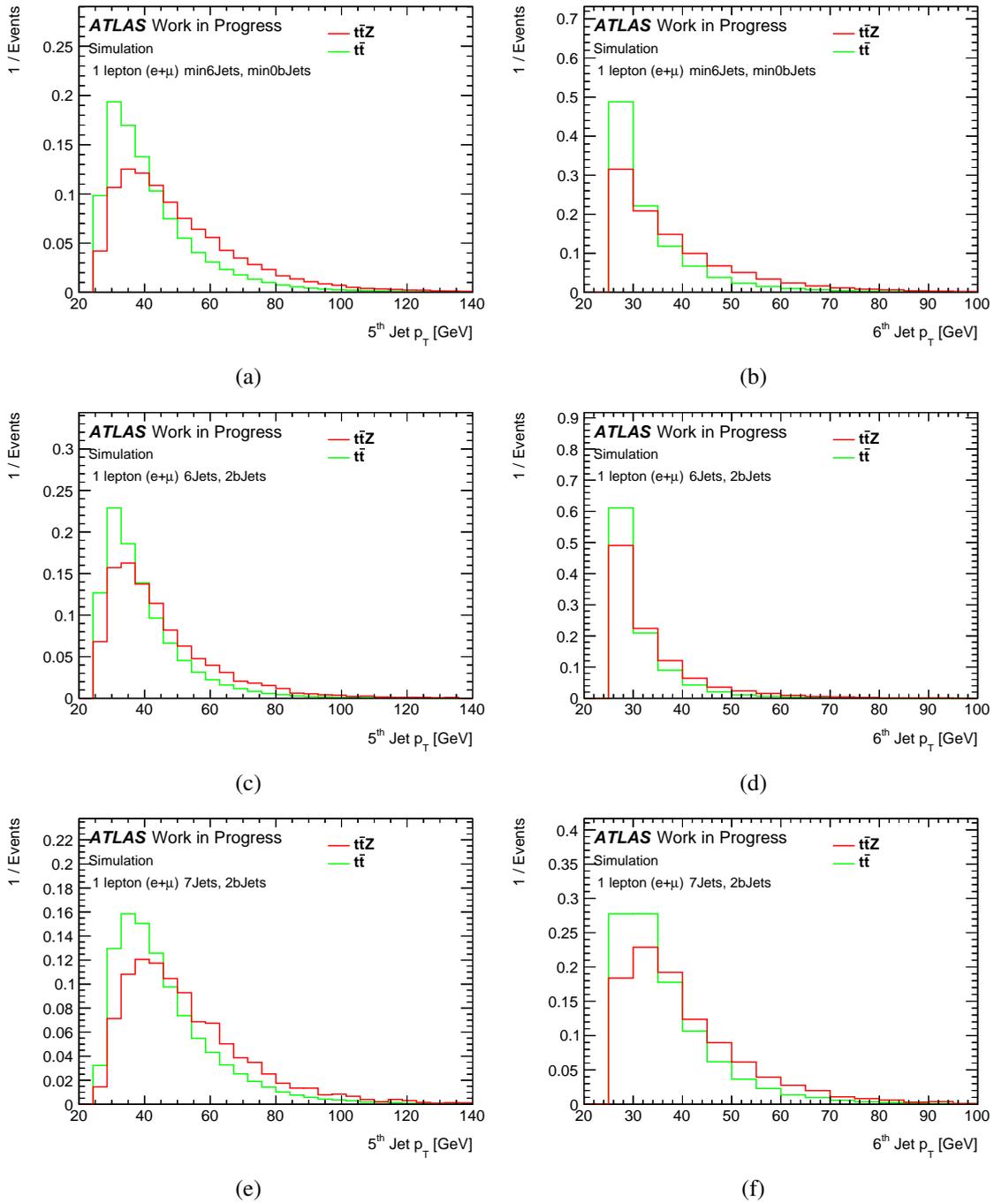


Fig. E.11.: Illustration of  $S_T$  on the left side and of  $p_T(Z)$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

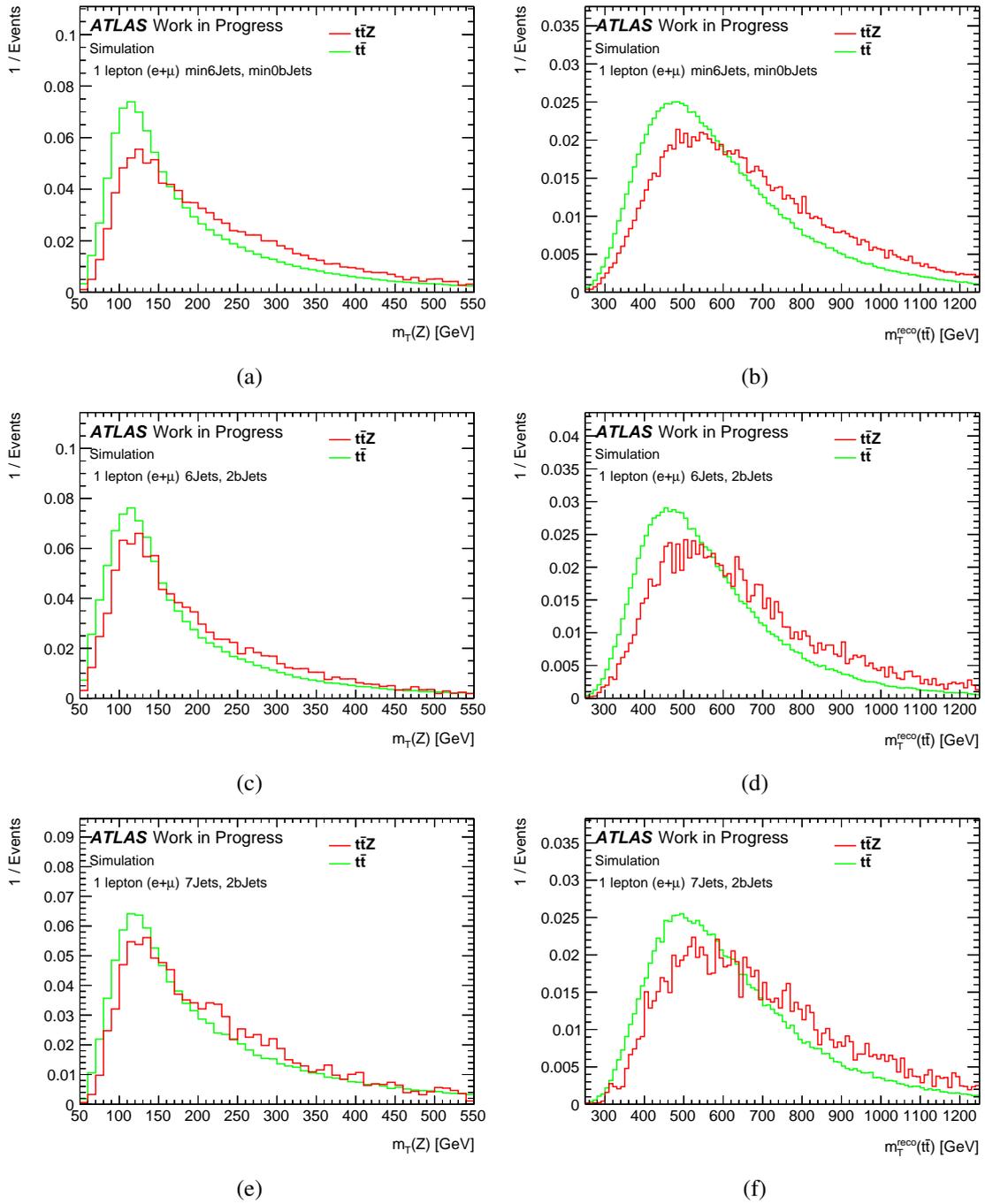


Fig. E.12.: Illustration of  $m_T^{\text{reco}}(Z)$  on the left side and of  $m_T^{\text{reco}}(t\bar{t})$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E. Additional plots

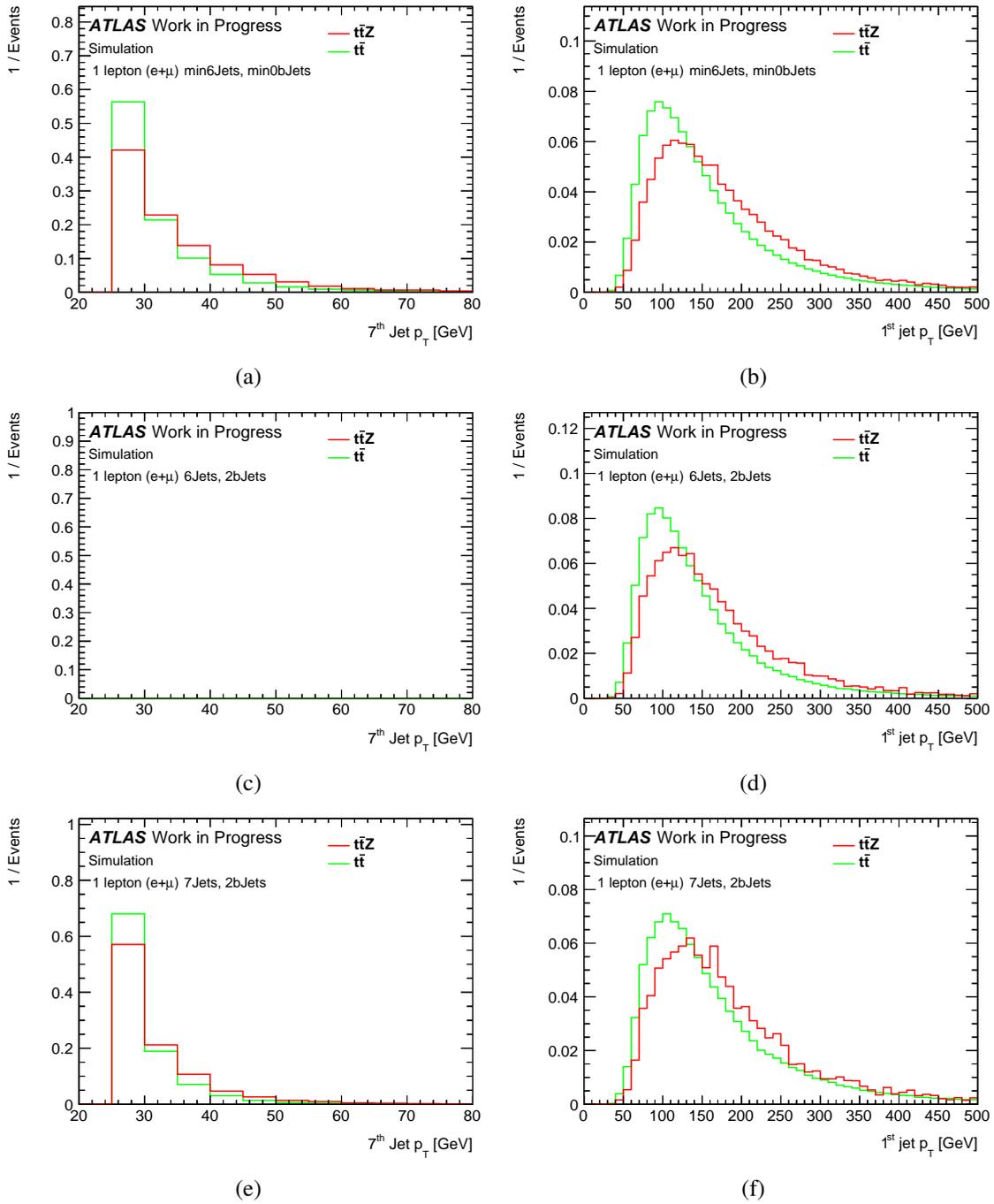


Fig. E.13.: Illustration of  $p_T(7^{\text{th}} \text{ jet})$  on the left side and of  $p_T(1^{\text{st}} \text{ jet})$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

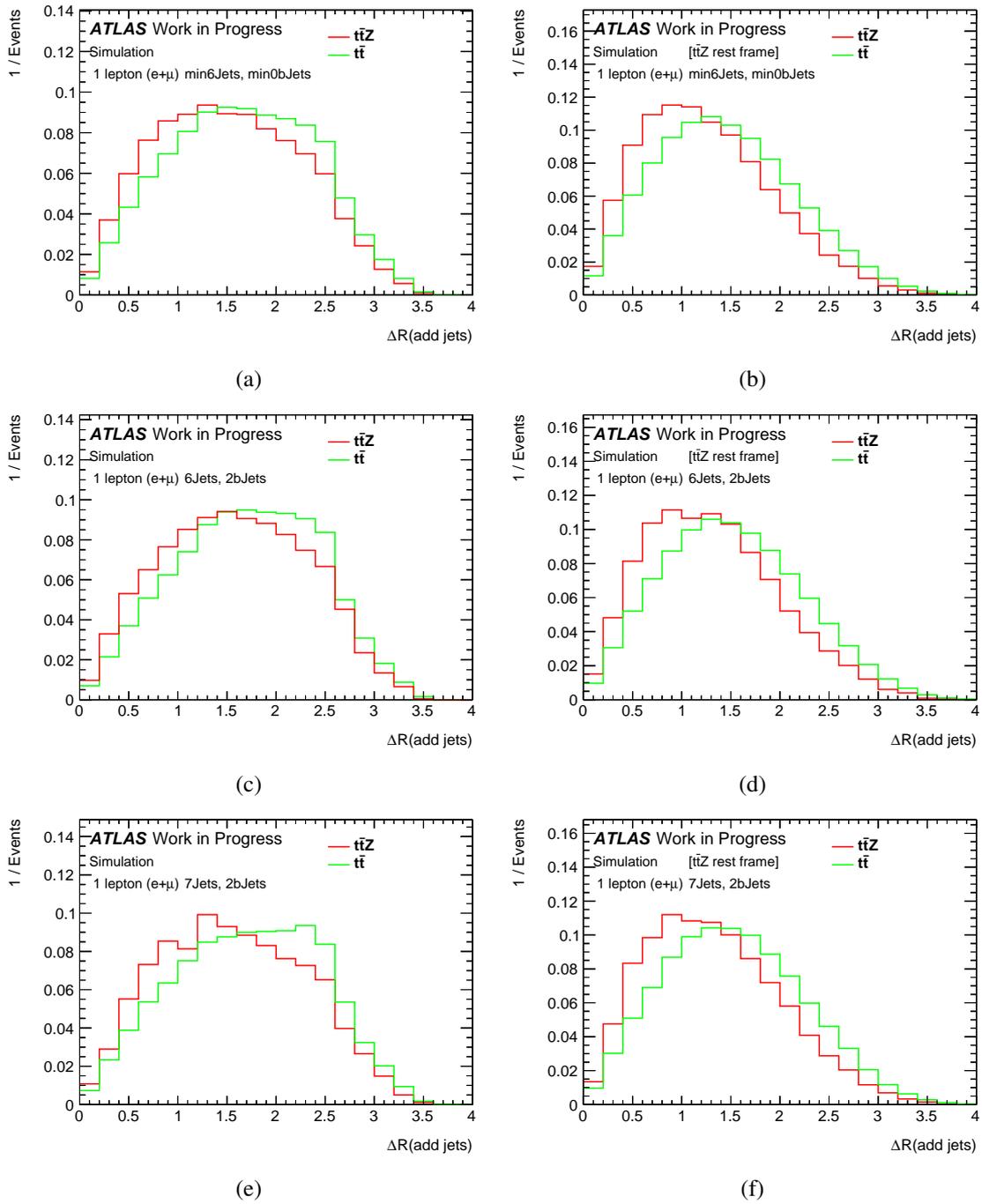


Fig. E.14.: Illustration of  $\Delta R(\text{add. jets})$  on the left side and of  $\Delta R^{ttZ}(\text{add. jets})$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E. Additional plots

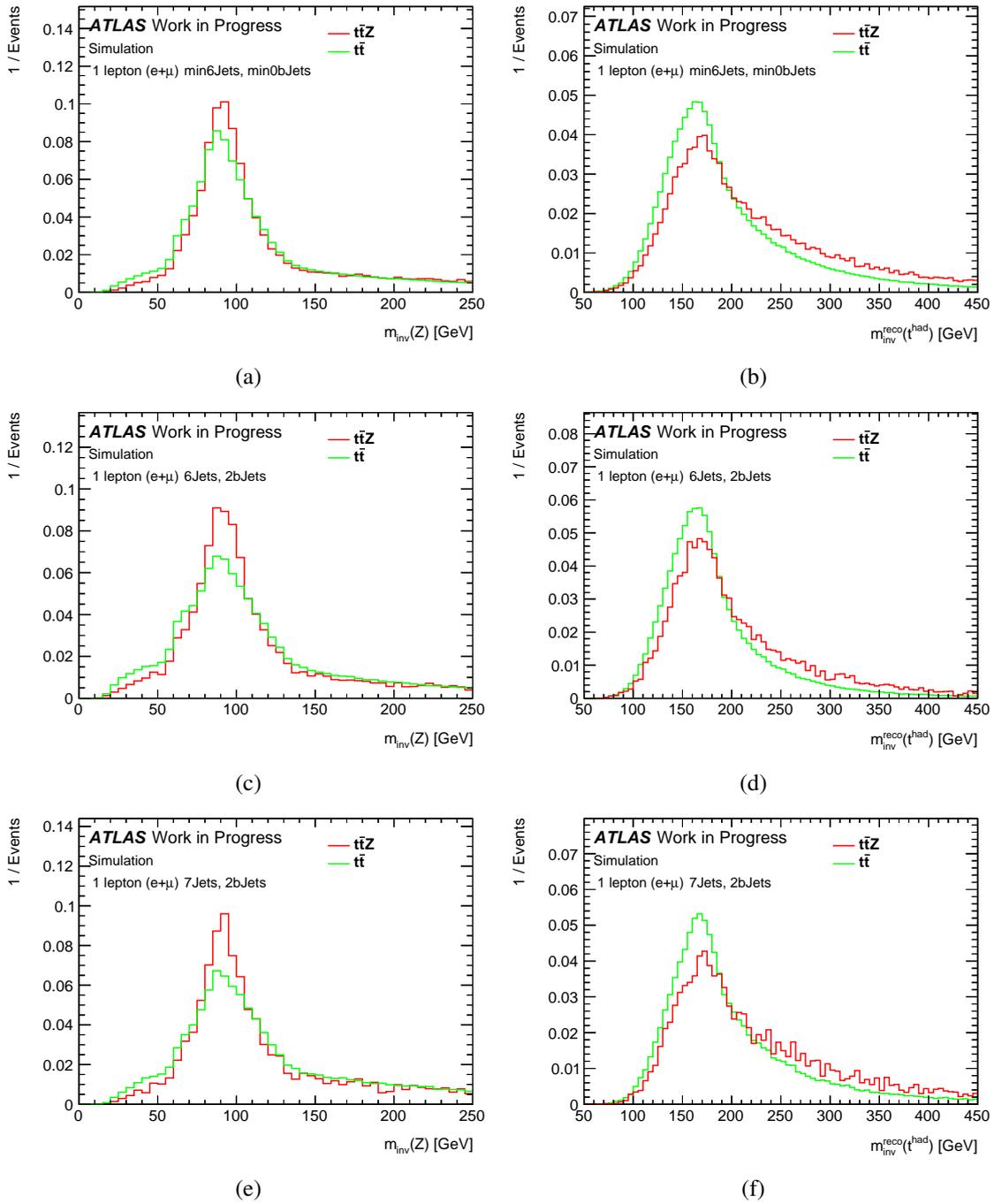


Fig. E.15.: Illustration of  $m_{\text{inv}}^{\text{reco}}(Z)$  on the left side and of  $m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

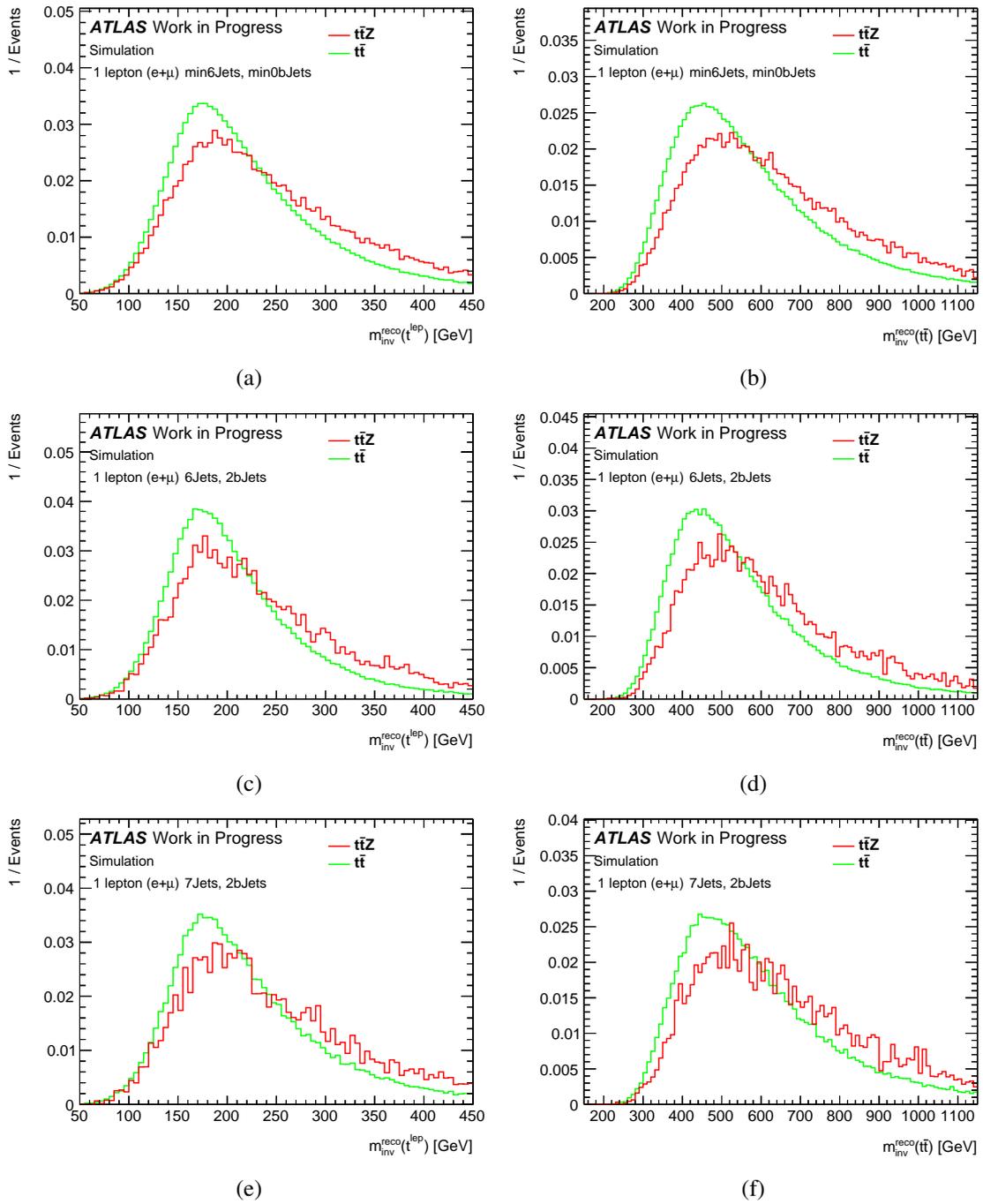


Fig. E.16.: Illustration of  $m_{inv}^{reco}(t^{lep})$  on the left side and of  $m_{inv}^{reco}(t\bar{t})$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

E. Additional plots

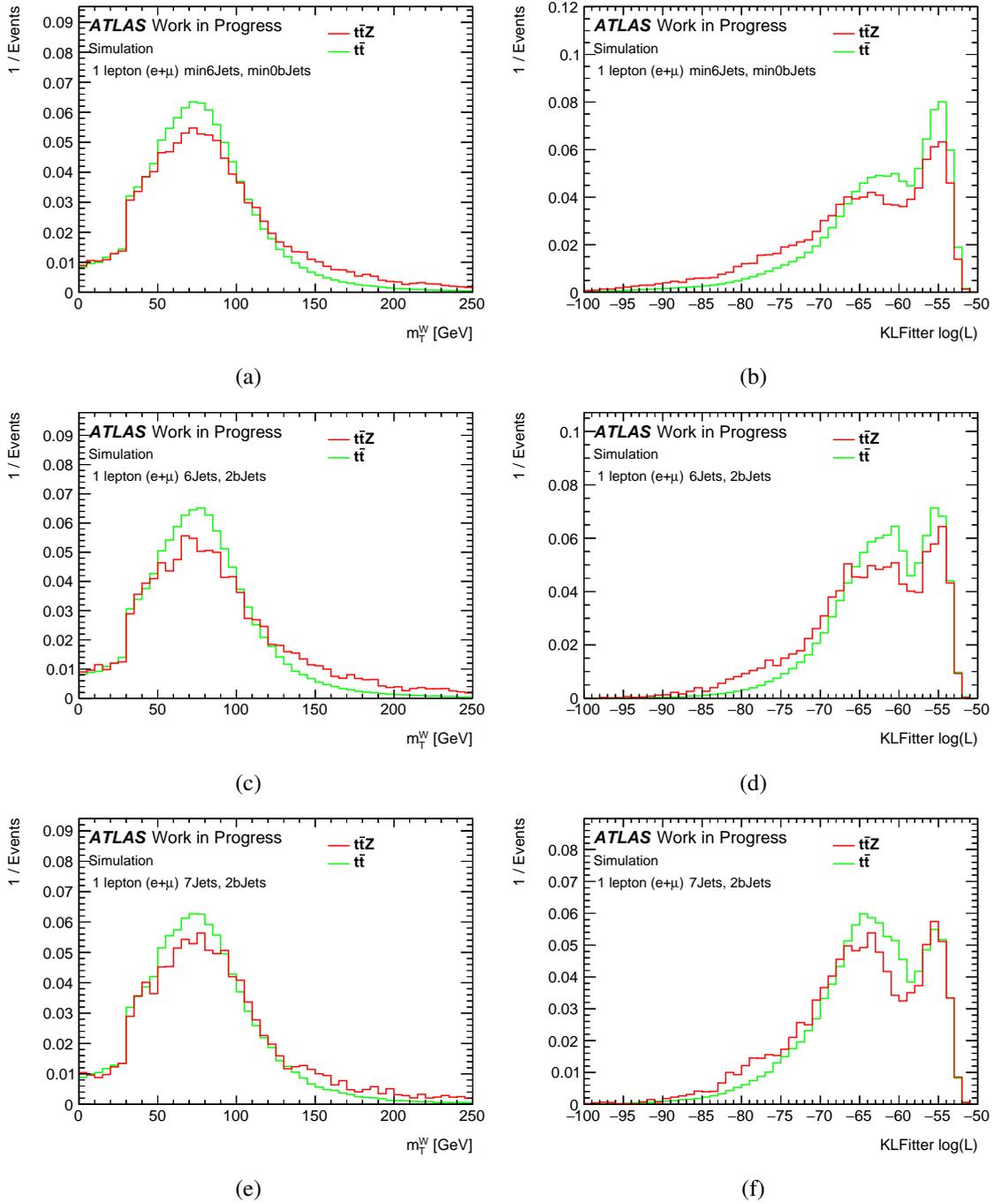


Fig. E.17.: Illustration of  $m_T^W$  on the left side and of  $\log L_{KL Fitter}$  on the right side, respectively. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). All distributions are normalised to total number of events in order to compare their shapes.

## E.4. BDT training results

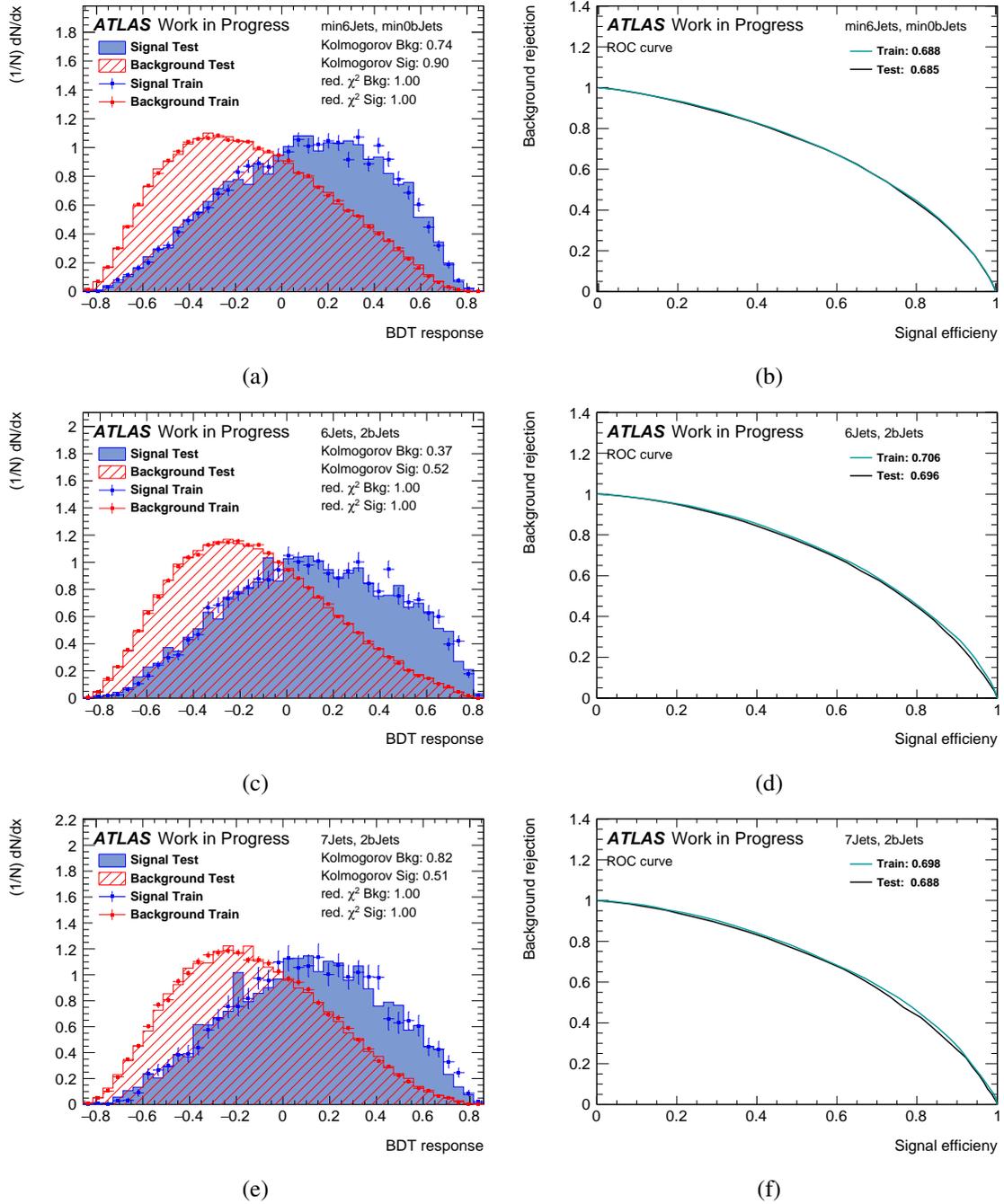


Fig. E.18.: Presentation of the BDT training results depicting the BDT response on the left side for signal and background in blue and red, respectively, and the corresponding ROC curves for training and testing on the right side. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). Both training and testing seems to be in good agreement, thus showing no indication of overtraining.

## E. Additional plots

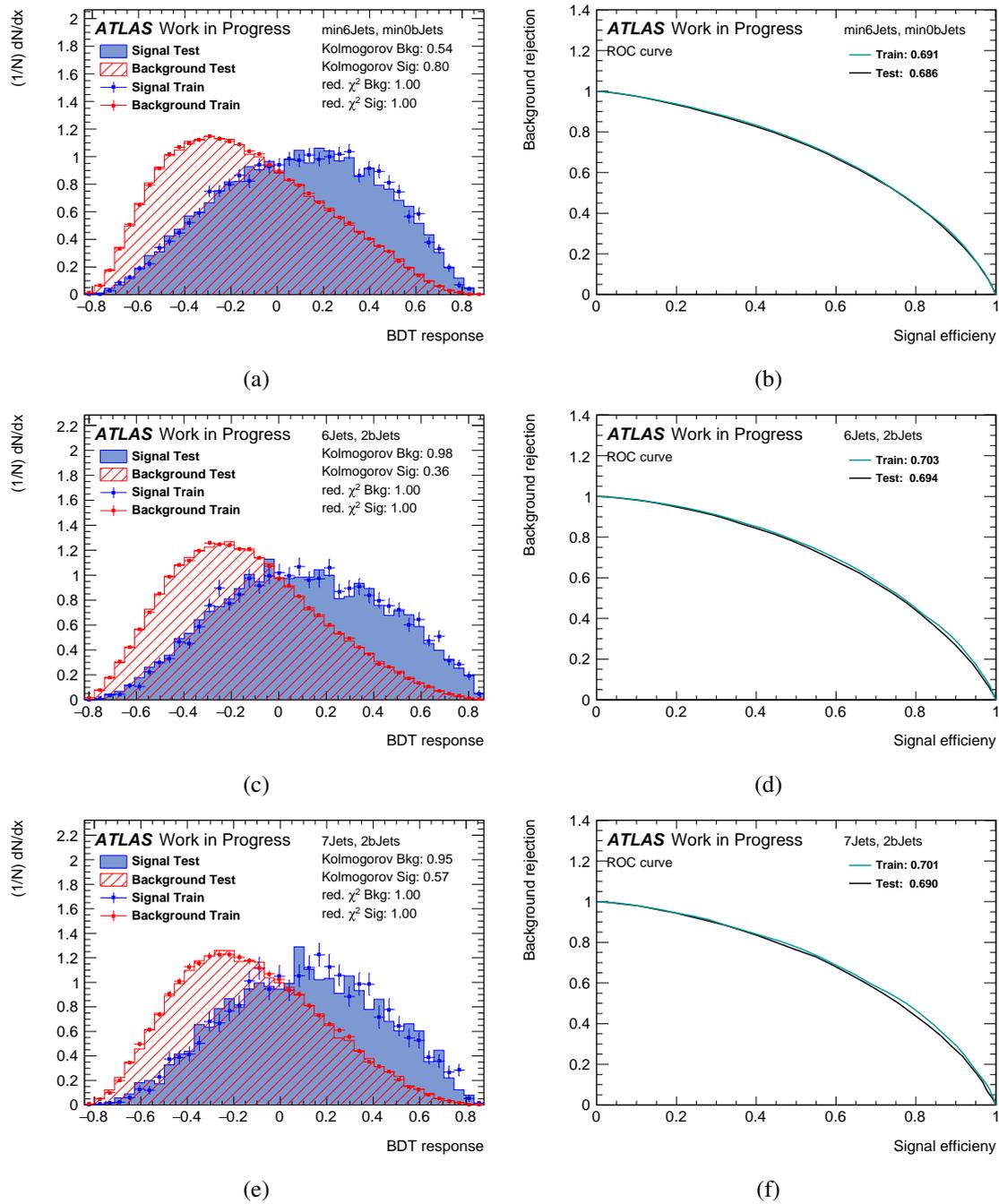


Fig. E.19.: Presentation of the BDT training results depicting the BDT response on the left side for signal and background in blue and red, respectively, and the corresponding ROC curves for training and testing on the right side. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). Both training and testing seems to be in good agreement, thus showing no indication of overtraining. Before the BDT training a variable decorrelation has been performed.

## E.5. BDT evaluation

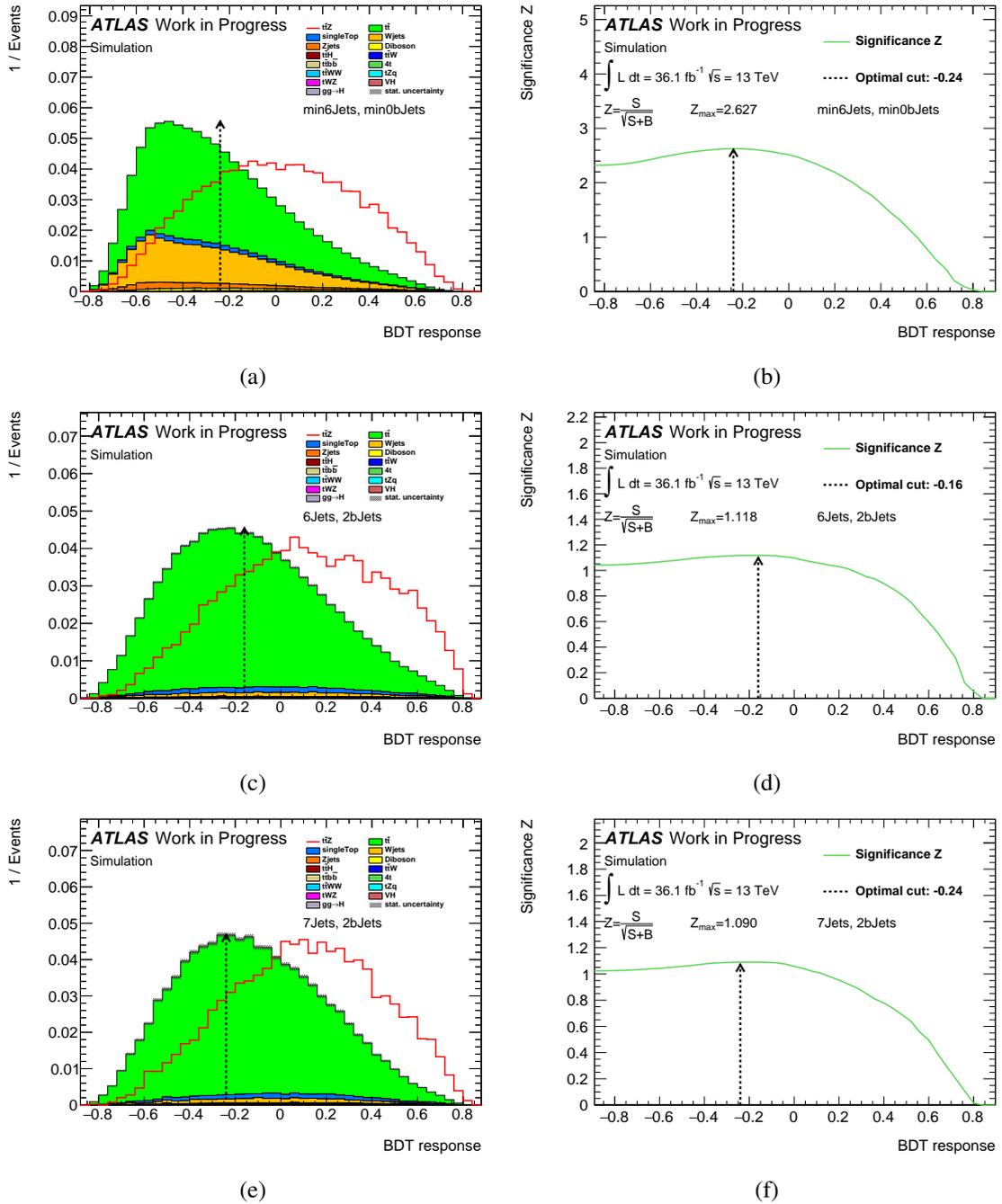


Fig. E.20.: Presentation of the BDT evaluation depicting the evaluated BDT response on the left side and the corresponding significance curve for an integrated luminosity of  $36.1 \text{ fb}^{-1}$  on the right side. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). Both signal and backgrounds are normalised to their respective number of events, thus illustrating the achieved separation. The dashed arrow indicates the optimal cut on the BDT response.

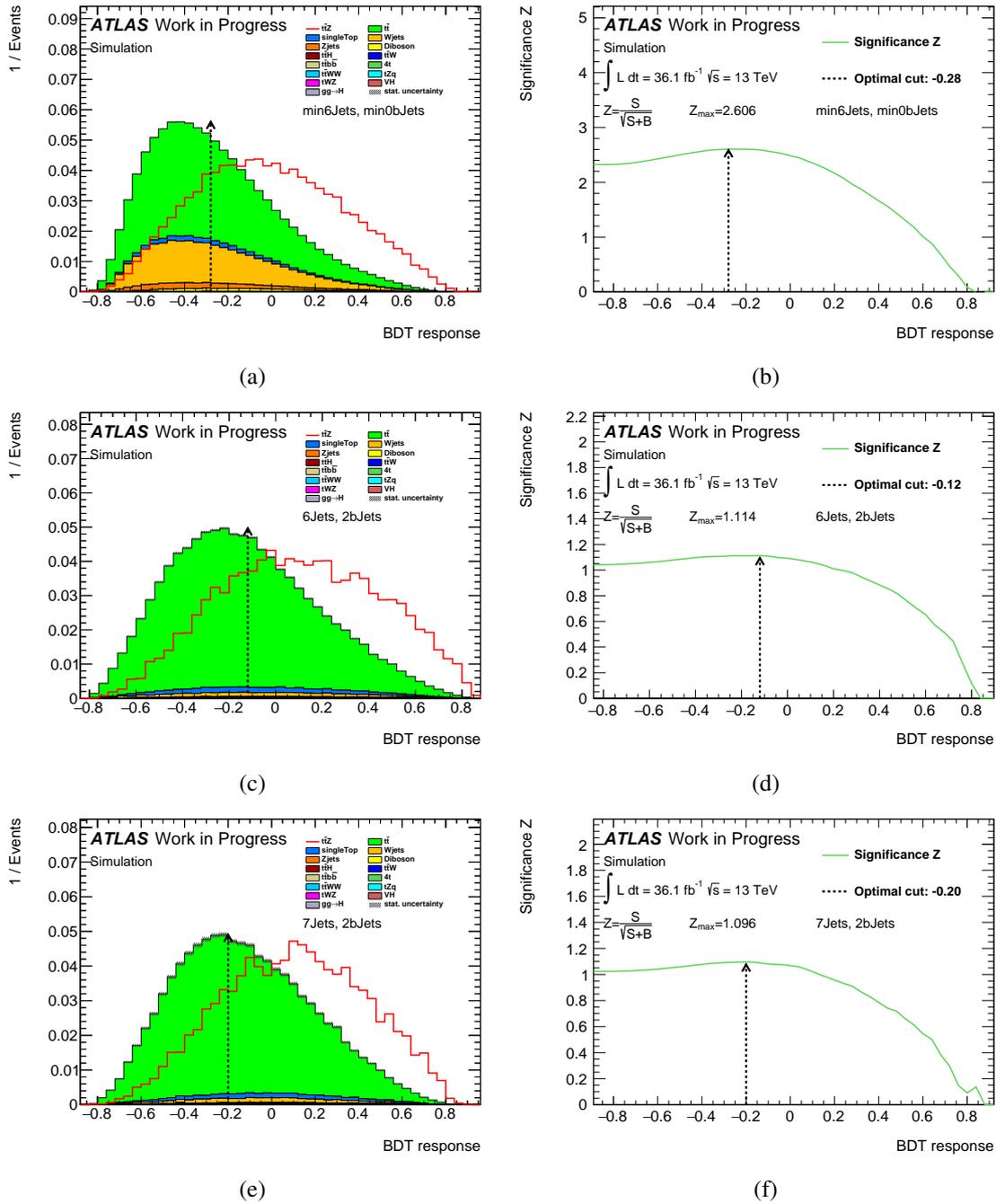


Fig. E.21.: Presentation of the BDT evaluation depicting the evaluated BDT response on the left side and the corresponding significance curve for an integrated luminosity of  $36.1 \text{ fb}^{-1}$  on the right side. Before the BDT training a variable decorrelation has been performed. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). Both signal and backgrounds are normalised to their respective number of events, thus illustrating the achieved separation. The dashed arrow indicates the optimal cut on the BDT response. Before the BDT training a variable decorrelation has been performed.

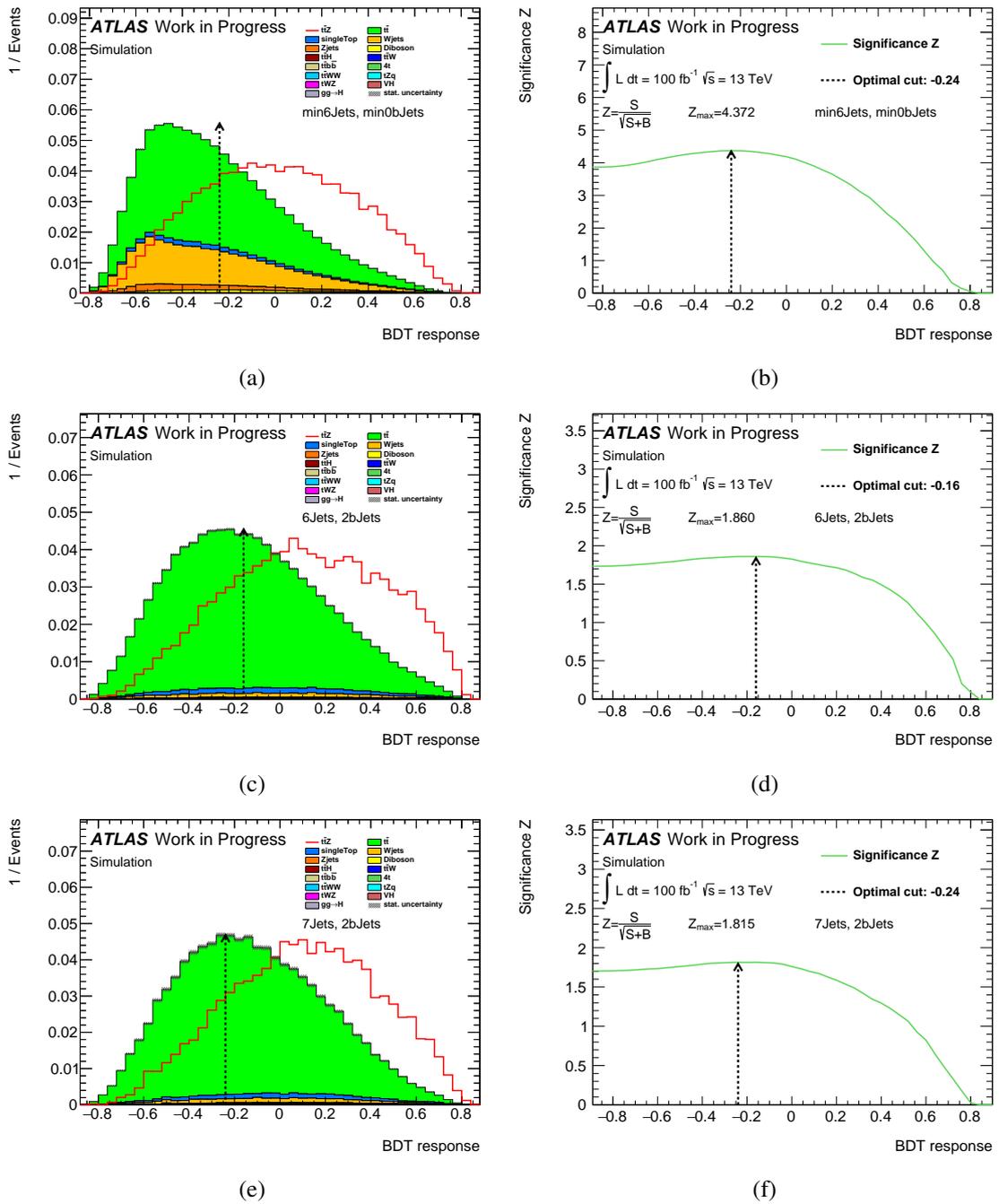


Fig. E.22.: Presentation of the BDT evaluation depicting the evaluated BDT response on the left side and the corresponding significance curve for an integrated luminosity of  $100 \text{ fb}^{-1}$  on the right side. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). Both signal and backgrounds are normalised to their respective number of events, thus illustrating the achieved separation. The dashed arrow indicates the optimal cut on the BDT response.

### E.6. Variable correlations

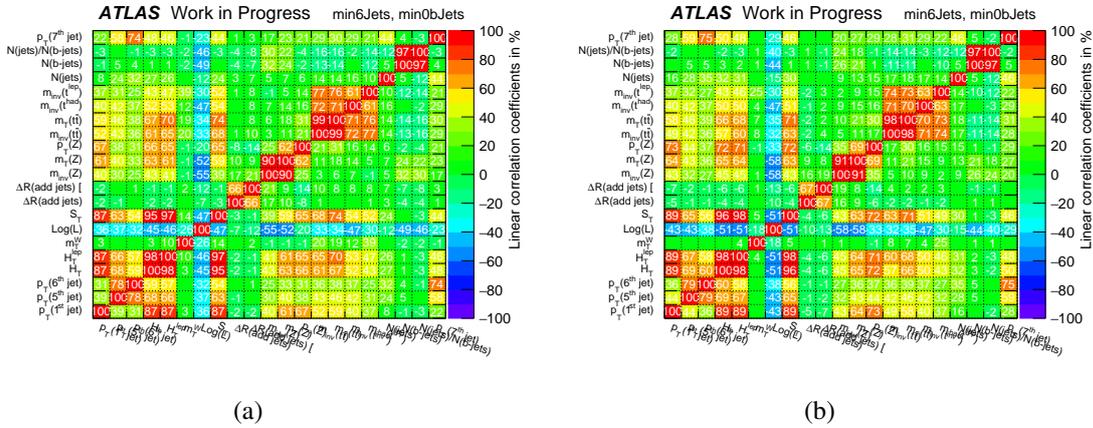


Fig. E.23.: Correlation matrix of the input variables in the first signal region for both the signal (Fig. (a)) and background (Fig. (b)) input sample.

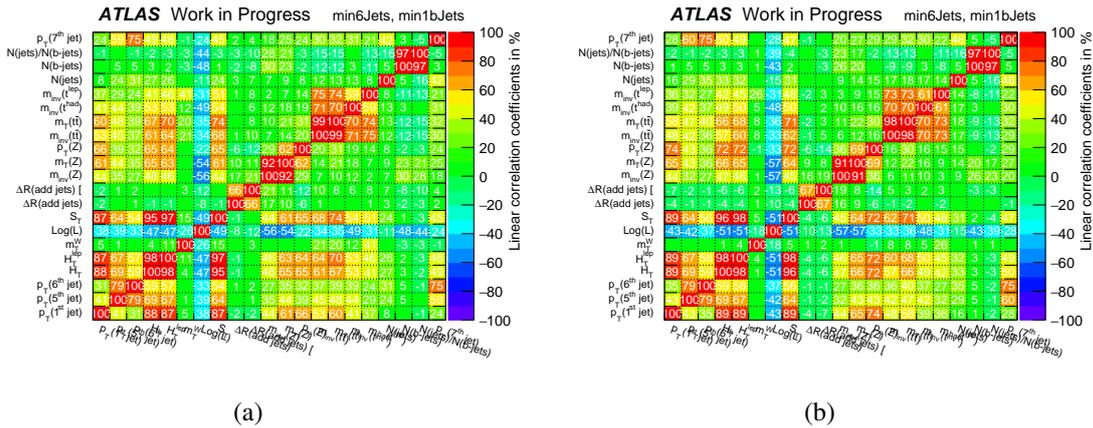


Fig. E.24.: Correlation matrix of the input variables in the first signal region for both the signal (Fig. (a)) and background (Fig. (b)) input sample.

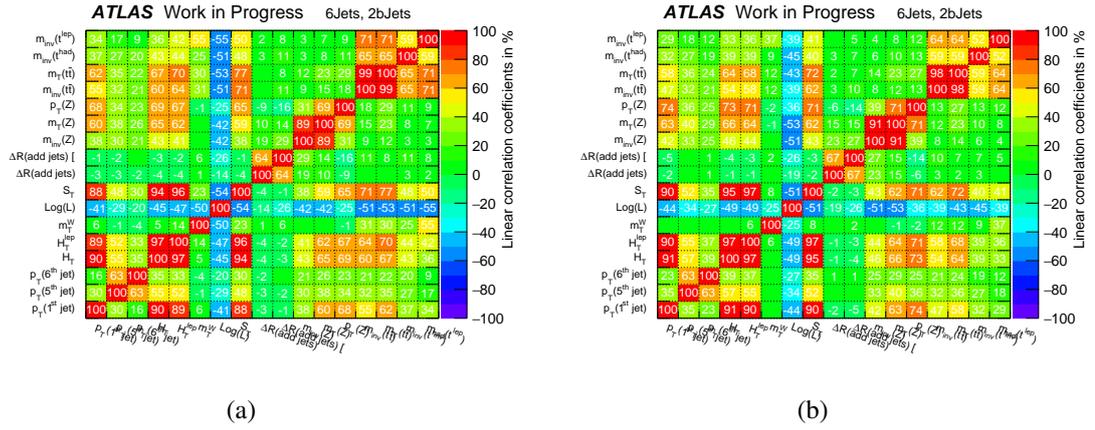


Fig. E.25.: Correlation matrix of the input variables in the first signal region for both the signal (Fig. (a)) and background (Fig. (b)) input sample.

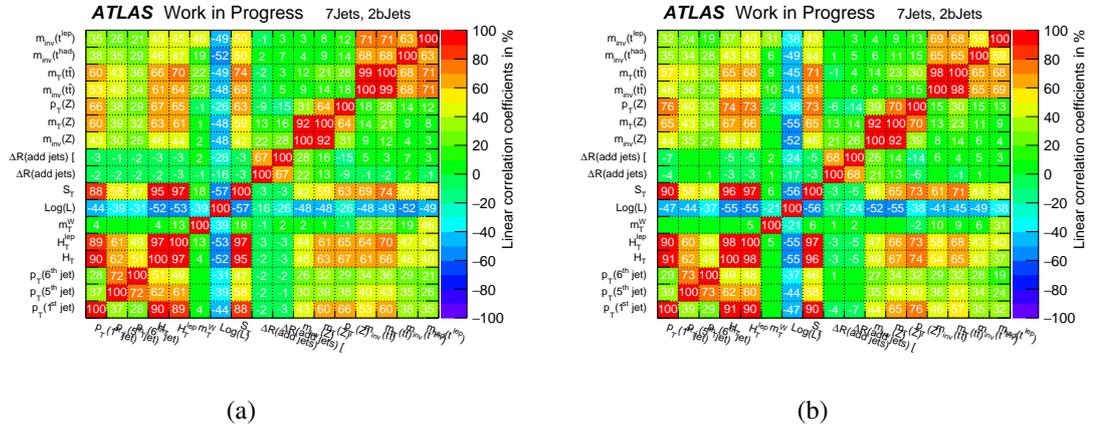


Fig. E.26.: Correlation matrix of the input variables in the first signal region for both the signal (Fig. (a)) and background (Fig. (b)) input sample.



## E.7. Systematic uncertainties

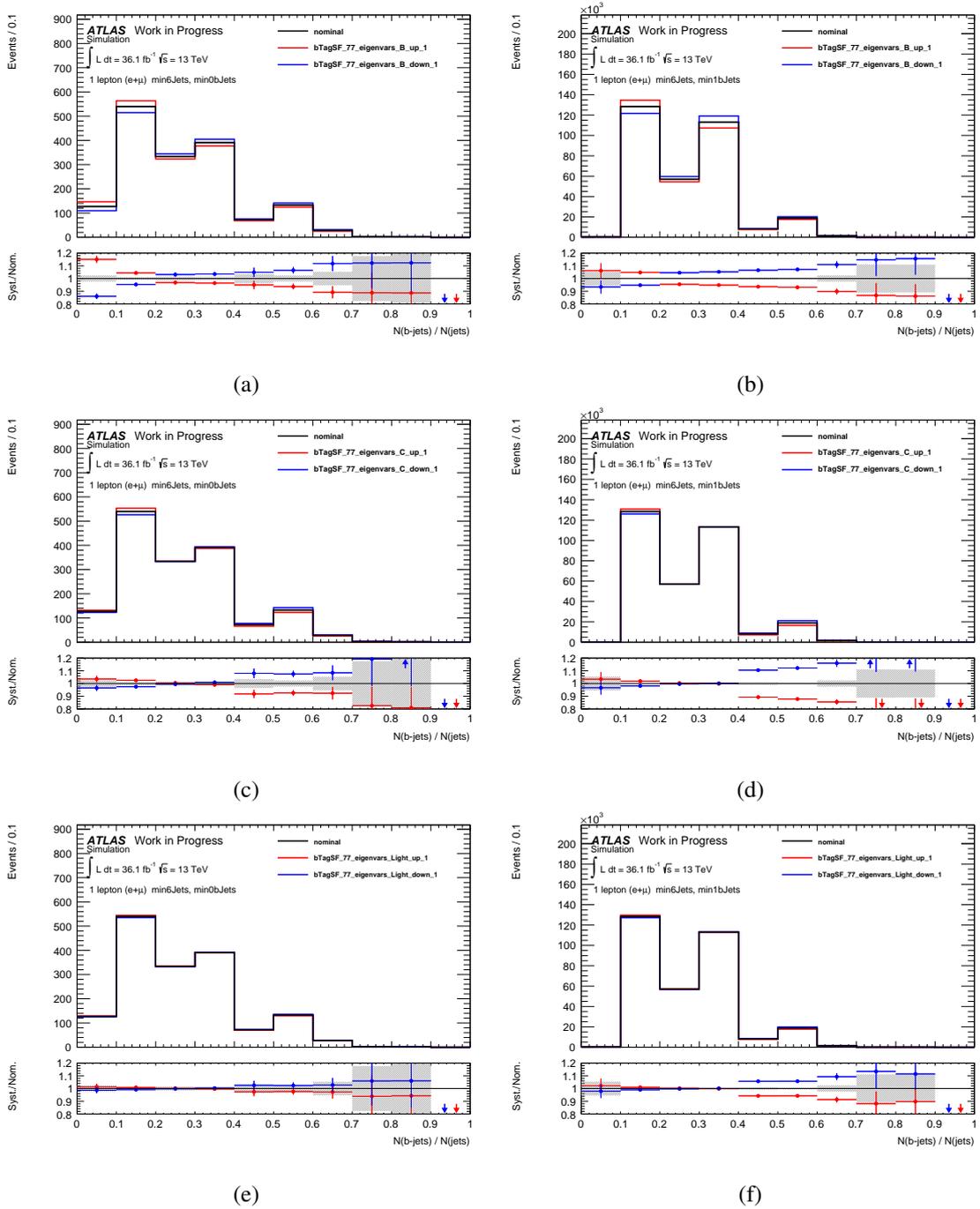


Fig. E.27.: Illustration of the impact of the largest systematic uncertainties on the  $b$ -tagging calibration, shown for the ratio of the  $b$ -jet multiplicity to the number of all jets in an event in SR1. The two upper images show the distributions for the most striking systematic parameter for the  $b$ -jet efficiency (cf. Fig. (a), (b)). In the middle, the situation for the relevant parameter for the  $c$ -jet efficiency is depicted (cf. Fig. (c), (d)) and on the bottom the significant parameter for the light jet mistag rate is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. On the left side the situation for the  $t\bar{t}Z$  signal is illustrated, whereas on the right side the dominant  $t\bar{t}$  background is depicted.

## E. Additional plots

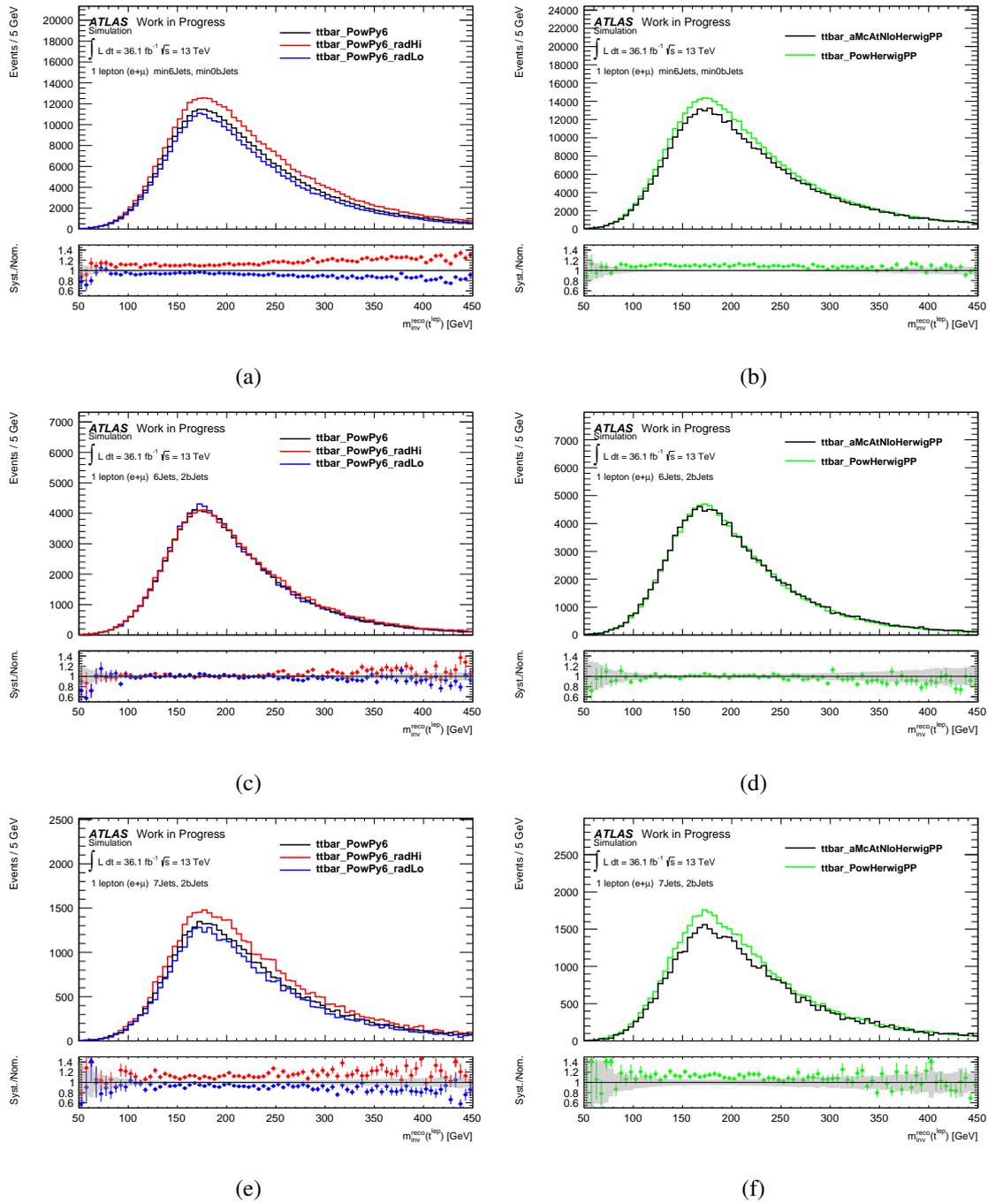


Fig. E.28.: Illustration of the impact of theoretical uncertainties on the additional QCD radiation and the simulation of the hard scatter event for the dominant  $t\bar{t}$  background in SR1, SR3 and SR4, shown for  $m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$ . On the left side the impact of different amounts of QCD radiation is illustrated, whereas on the right side the comparison of different Monte Carlo generators for the hard scatter interaction is depicted. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.

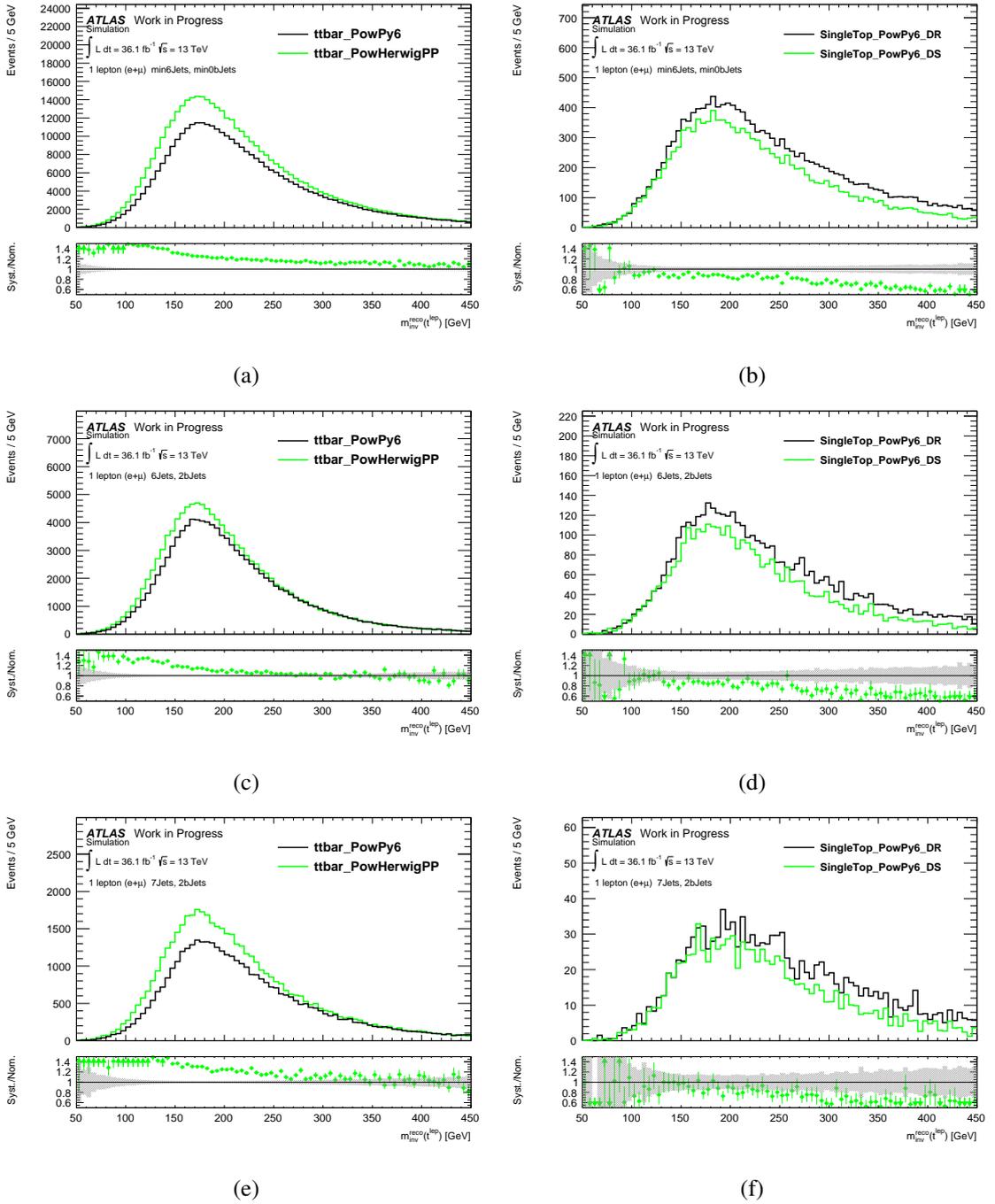


Fig. E.29.: Illustration of the impact of the simulation of the parton shower for  $t\bar{t}$  and the difference between diagram removal and subtraction for the  $Wt$ -channel of single top quark production in SR1, SR3 and SR4, shown for  $m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$ . On the left side the comparison of different Monte Carlo generators for the parton shower modelling is illustrated, whereas on the right side the impact of the difference between diagram removal and subtraction is depicted. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.

## E. Additional plots

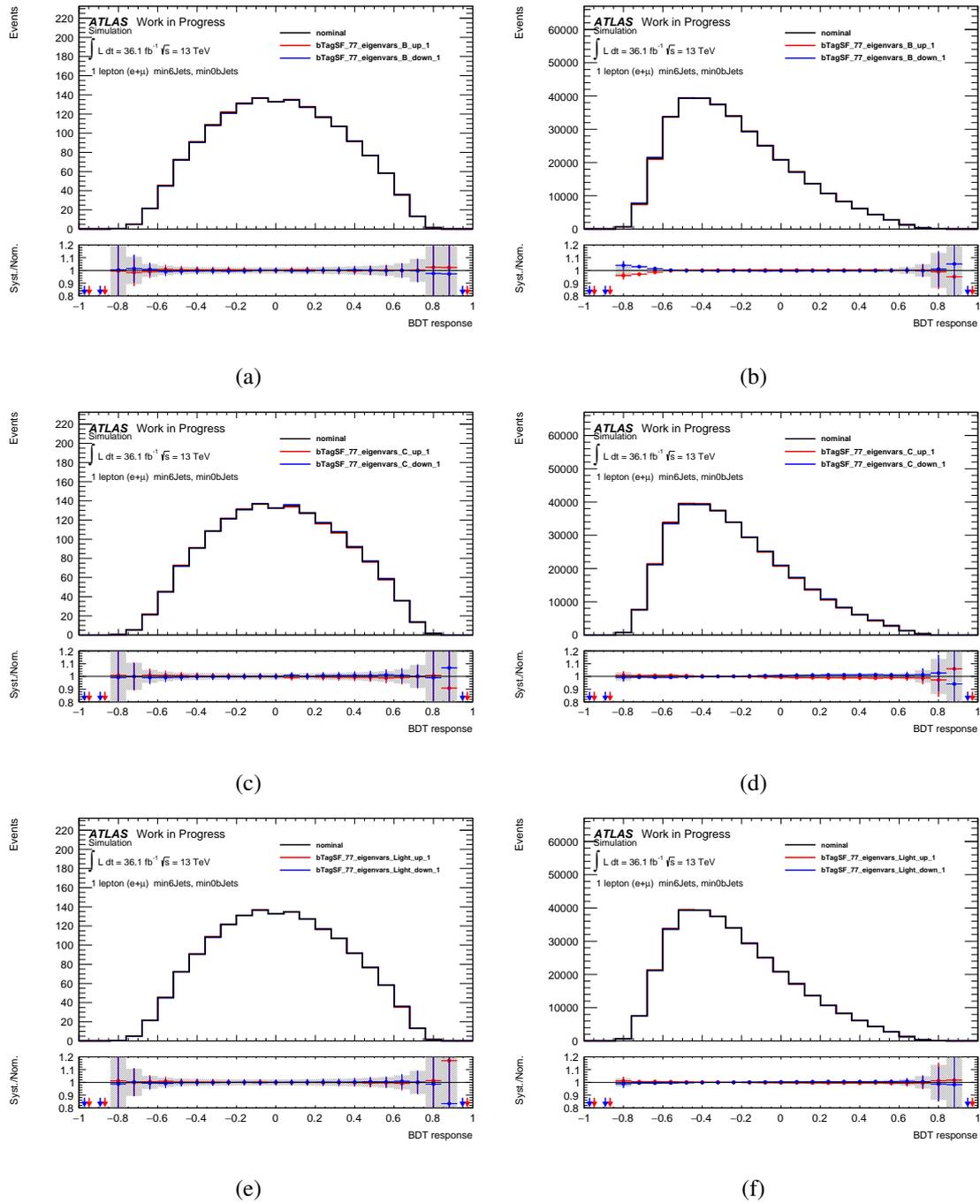


Fig. E.30.: Illustration of the impact of the largest systematic uncertainties on the  $b$ -tagging calibration, shown for the BDT response in SR1. The two upper images show the distributions for the most striking systematic parameter for the  $b$ -jet efficiency (cf. Fig. (a), (b)). In the middle, the situation for the relevant parameter for the  $c$ -jet efficiency is depicted (cf. Fig. (c), (d)) and on the bottom the significant parameter for the light jet mistag rate is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. On the left side the situation for the  $t\bar{t}Z$  signal is illustrated, whereas on the right side the dominant  $t\bar{t}$  background is depicted.

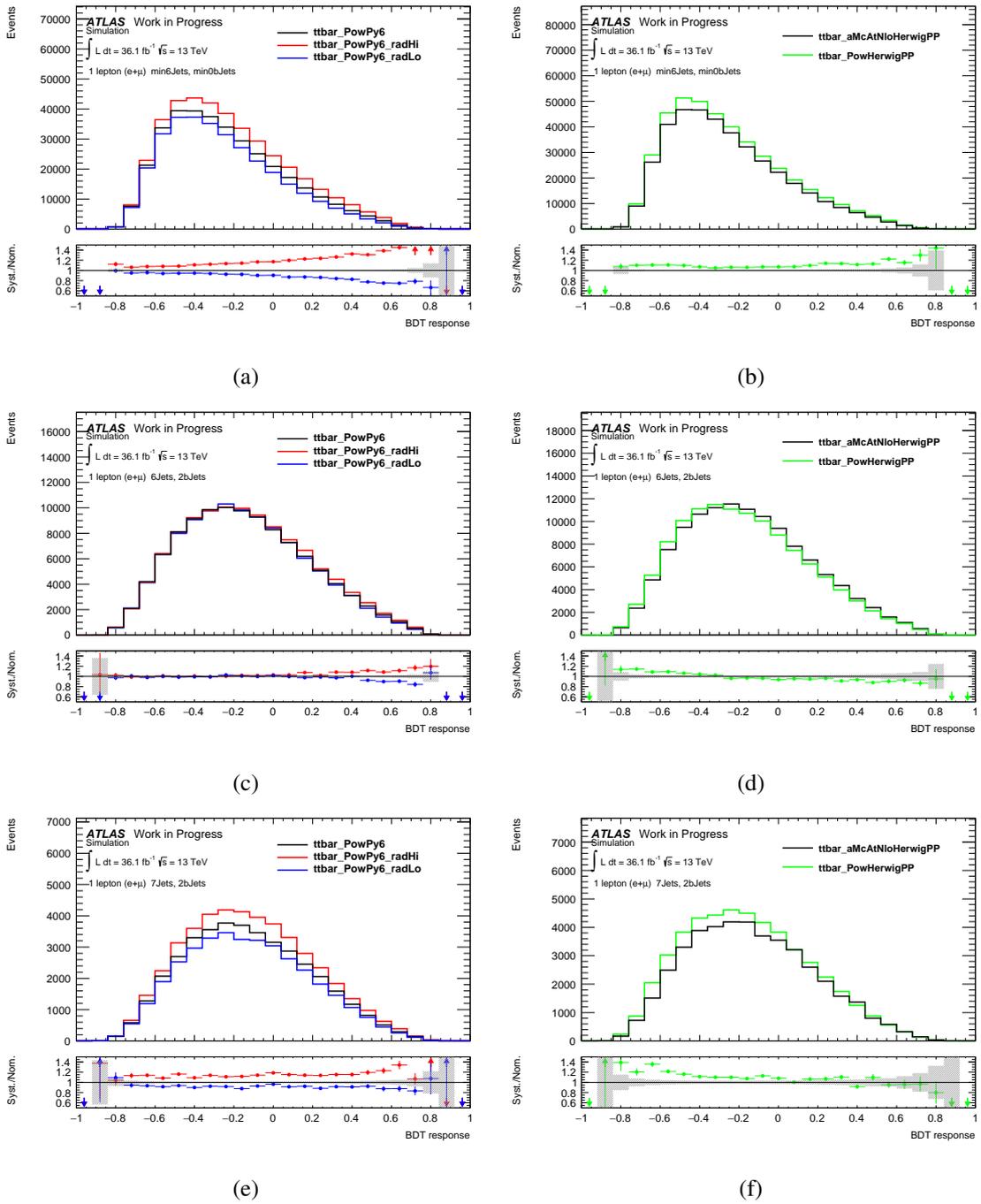


Fig. E.31.: Illustration of the impact of theoretical uncertainties on the additional QCD radiation and the simulation of the hard scatter event for the dominant  $t\bar{t}$  background in SR2, shown for the BDT response. On the left side the impact of different amounts of QCD radiation is illustrated, whereas on the right side the comparison of different Monte Carlo generators for the hard scatter interaction is depicted. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.

## E. Additional plots

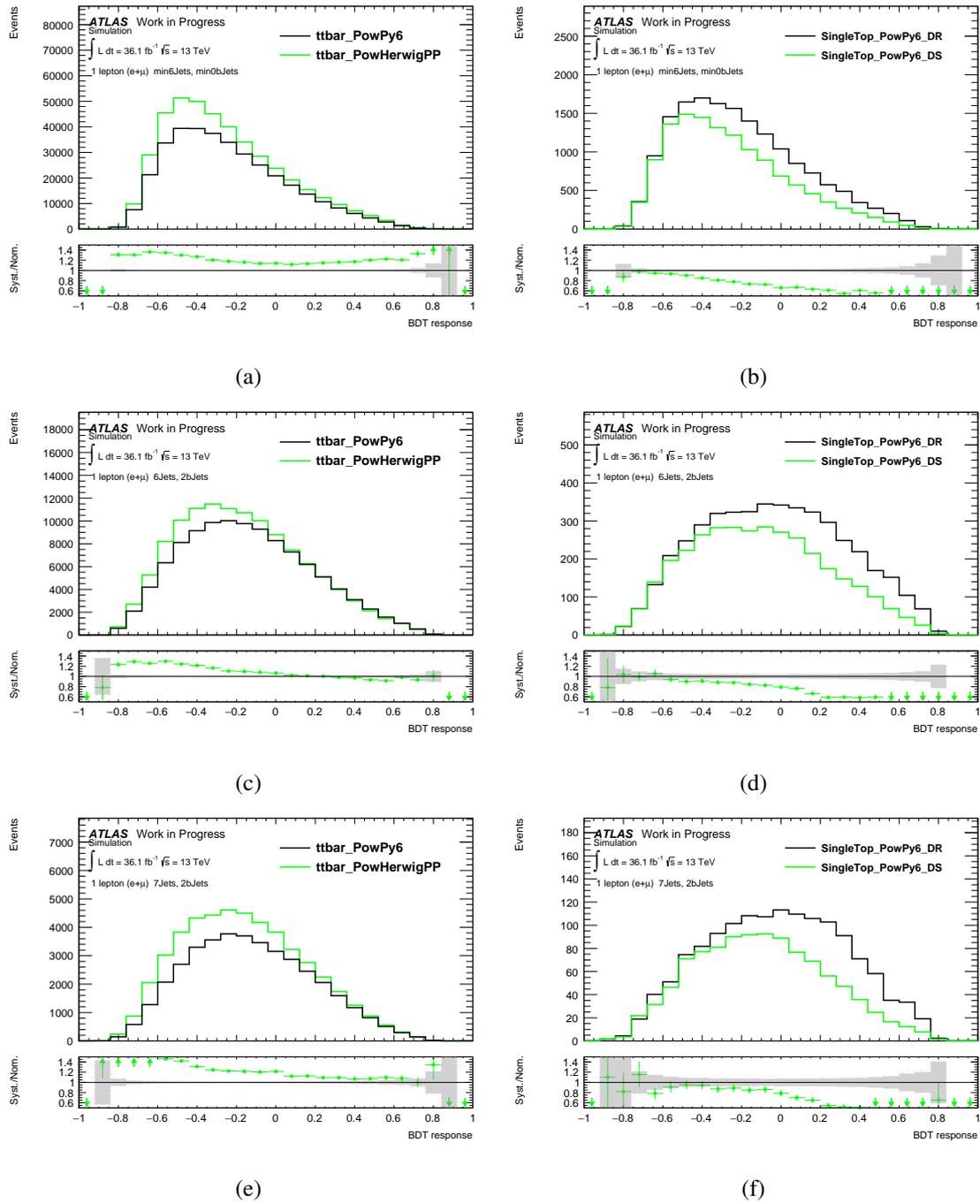


Fig. E.32.: Illustration of the impact of the simulation of the parton shower for  $t\bar{t}$  and the difference between diagram removal and subtraction for the  $Wt$ -channel of single top quark production in SR2, shown for the BDT response. On the left side the comparison of different Monte Carlo generators for the parton shower modelling is illustrated, whereas on the right side the impact of the difference between diagram removal and subtraction is depicted. The two upper images show the distributions for SR1 (cf. Fig. (a), (b)). In the middle, the situation in SR3 is depicted (cf. Fig. (c), (d)) and on the bottom SR4 is shown (cf. Fig. (e), (f)). The black line depicts the nominal distribution, whereas the red and the blue line represent the up and down variation, respectively. The shape depicted in green shows the behaviour of the alternative Monte Carlo generator.

## F. Additional tables

### F.1. BDT rankings

Ranking	SR1	SR2	SR3	SR4
1	$p_T(1^{\text{st}} \text{ jet})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$
2	$p_T(5^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{T}}^W$
3	$m_{\text{inv}}^{\text{reco}}(Z)$	$p_T(1^{\text{st}} \text{ jet})$	$m_{\text{T}}^W$	$p_T(5^{\text{th}} \text{ jet})$
4	$m_{\text{T}}^W$	$m_{\text{T}}^W$	$p_T(5^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(Z)$
5	$H_{\text{T}}$	$p_T(5^{\text{th}} \text{ jet})$	$S_{\text{T}}$	$\Delta R(\text{add. jets})$
6	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\log L_{\text{KLFitter}}$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$\log L_{\text{KLFitter}}$
7	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$N(\text{jets})$	$\log L_{\text{KLFitter}}$	$p_{\text{T}}^{\text{reco}}(Z)$
8	$\log L_{\text{KLFitter}}$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$p_T(6^{\text{th}} \text{ jet})$
9	$N(b\text{-jets})/N(\text{jets})$	$N(b\text{-jets})/N(\text{jets})$	$p_T(6^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$
10	$N(\text{jets})$	$H_{\text{T}}$	$\Delta R(\text{add. jets})$	$p_T(1^{\text{st}} \text{ jet})$
11	$p_T(7^{\text{th}} \text{ jet})$	$p_{\text{T}}^{\text{reco}}(Z)$	$p_{\text{T}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$
12	$p_{\text{T}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$S_{\text{T}}$
13	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$S_{\text{T}}$	$p_T(1^{\text{st}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$
14	$\Delta R(\text{add. jets})$	$m_{\text{T}}^{\text{reco}}(t\bar{t})$	$m_{\text{T}}^{\text{reco}}(Z)$	$H_{\text{T}}$
15	$N(b\text{-jets})$	$p_T(6^{\text{th}} \text{ jet})$	$H_{\text{T}}$	$m_{\text{T}}^{\text{reco}}(t\bar{t})$
16	$S_{\text{T}}$	$p_T(7^{\text{th}} \text{ jet})$	$H_{\text{T}}^{\text{lep}}$	$m_{\text{T}}^{\text{reco}}(Z)$
17	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{T}}^{\text{reco}}(t\bar{t})$	$H_{\text{T}}^{\text{lep}}$
18	$p_T(6^{\text{th}} \text{ jet})$	$\Delta R(\text{add. jets})$	–	–
19	$m_{\text{T}}^{\text{reco}}(Z)$	$N(b\text{-jets})$	–	–
20	$m_{\text{T}}^{\text{reco}}(t\bar{t})$	$m_{\text{T}}^{\text{reco}}(Z)$	–	–
21	$H_{\text{T}}^{\text{lep}}$	$H_{\text{T}}^{\text{lep}}$	–	–

Tab. F.1.: Ranking of the BDT input variables for all signal regions after the training. The rankings are estimated from the frequency of being used in split decision during training. The last four positions for the exclusive signal regions are not occupied due to the smaller amount of input variables.

Ranking	SR1	SR2	SR3	SR4
1	$p_T(5^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$	$S_T$	$S_T$
2	$p_T(6^{\text{th}} \text{ jet})$	$p_T(6^{\text{th}} \text{ jet})$	$H_T^{\text{lep}}$	$p_T(5^{\text{th}} \text{ jet})$
3	$S_T$	$S_T$	$H_T$	$H_T^{\text{lep}}$
4	$H_T$	$H_T$	$m_T^{\text{reco}}(t\bar{t})$	$H_T$
5	$H_T^{\text{lep}}$	$H_T^{\text{lep}}$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_T^{\text{reco}}(t\bar{t})$
6	$p_T(7^{\text{th}} \text{ jet})$	$p_T(7^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$
7	$N(b\text{-jets})/N(\text{jets})$	$N(b\text{-jets})/N(\text{jets})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$p_T(6^{\text{th}} \text{ jet})$
8	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$
9	$m_T^{\text{reco}}(t\bar{t})$	$N(\text{jets})$	$m_T^W$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$
10	$N(\text{jets})$	$m_T^{\text{reco}}(t\bar{t})$	$p_T(1^{\text{st}} \text{ jet})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$
11	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$\log L_{\text{KLFitter}}$	$m_T^W$
12	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$p_T^{\text{reco}}(Z)$	$p_T(1^{\text{st}} \text{ jet})$
13	$\log L_{\text{KLFitter}}$	$\log L_{\text{KLFitter}}$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\log L_{\text{KLFitter}}$
14	$p_T(1^{\text{st}} \text{ jet})$	$m_T^{\text{reco}}(Z)$	$p_T(6^{\text{th}} \text{ jet})$	$p_T^{\text{reco}}(Z)$
15	$m_T^{\text{reco}}(Z)$	$p_T(1^{\text{st}} \text{ jet})$	$m_T^{\text{reco}}(Z)$	$\Delta R(\text{add. jets})$
16	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R(\text{add. jets})$	$m_T^{\text{reco}}(Z)$
17	$p_T^{\text{reco}}(Z)$	$p_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$
18	$m_T^W$	$N(b\text{-jets})$	–	–
19	$N(b\text{-jets})$	$N(b\text{-jets})$	–	–
20	$\Delta R(\text{add. jets})$	$\Delta R(\text{add. jets})$	–	–
21	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$	–	–

Tab. F.2.: Ranking of the BDT input variables for all signal regions before the training. The rankings are determined from the overlapt of the variables shapes. The last four positions for the exclusive signal regions are not occupied due to the smaller amount of input variables.

Ranking	SR1	SR2	SR3	SR4
1	$N(b\text{-jets})$	$p_T(1^{\text{st}} \text{ jet})$	$p_T(1^{\text{st}} \text{ jet})$	$p_T(1^{\text{st}} \text{ jet})$
2	$p_T(1^{\text{st}} \text{ jet})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$
3	$\log L_{\text{KLFitter}}$	$\log L_{\text{KLFitter}}$	$m_T^W$	$\log L_{\text{KLFitter}}$
4	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$N(b\text{-jets})$	$\log L_{\text{KLFitter}}$	$m_T^W$
5	$H_T$	$H_T$	$p_T^{\text{reco}}(Z)$	$m_T^{\text{reco}}(Z)$
6	$p_T(7^{\text{th}} \text{ jet})$	$m_T^{\text{reco}}(Z)$	$m_T^{\text{reco}}(Z)$	$H_T$
7	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_T^W$	$H_T^{\text{lep}}$	$H_T^{\text{lep}}$
8	$m_T^W$	$p_T(7^{\text{th}} \text{ jet})$	$H_T$	$p_T^{\text{reco}}(Z)$
9	$S_T$	$S_T$	$m_T^{\text{reco}}(t\bar{t})$	$m_T^{\text{reco}}(t\bar{t})$
10	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$\Delta R(\text{add. jets}) [t\bar{t}Z]$
11	$m_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$S_T$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$
12	$p_T(5^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{inv}}^{\text{reco}}(Z)$
13	$p_T(6^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$p_T(5^{\text{th}} \text{ jet})$	$S_T$
14	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$H_T^{\text{lep}}$	$m_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$
15	$H_T^{\text{lep}}$	$p_T(6^{\text{th}} \text{ jet})$	$\Delta R(\text{add. jets})$	$p_T(5^{\text{th}} \text{ jet})$
16	$\Delta R(\text{add. jets}) [t\bar{t}Z]$	$m_T^{\text{reco}}(t\bar{t})$	$p_T(6^{\text{th}} \text{ jet})$	$p_T(6^{\text{th}} \text{ jet})$
17	$N_{\text{jets}}$	$p_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$
18	$m_T^{\text{reco}}(t\bar{t})$	$N(\text{jets})$	–	–
19	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$\Delta R(\text{add. jets})$	–	–
20	$N(b\text{-jets})/N(\text{jets})$	$N(b\text{-jets})/N(\text{jets})$	–	–
21	$p_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	–	–

Tab. F.3.: Ranking of the BDT input variables for all signal regions after the training. Before the BDT training a variable decorrelation has been performed. The rankings are estimated from the frequency of being used in split decision during training. The last four positions for the exclusive signal regions are not occupied due to the smaller amount of input variables.

Ranking	SR1	SR2	SR3	SR4
1	$p_T(5^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$	$S_T$	$S_T$
2	$p_T(6^{\text{th}} \text{ jet})$	$p_T(6^{\text{th}} \text{ jet})$	$H_T^{\text{lep}}$	$p_T(5^{\text{th}} \text{ jet})$
3	$S_T$	$S_T$	$H_T$	$H_T^{\text{lep}}$
4	$H_T$	$H_T$	$m_T^{\text{reco}}(t\bar{t})$	$H_T$
5	$H_T^{\text{lep}}$	$H_T^{\text{lep}}$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_T^{\text{reco}}(t\bar{t})$
6	$p_T(7^{\text{th}} \text{ jet})$	$p_T(7^{\text{th}} \text{ jet})$	$p_T(5^{\text{th}} \text{ jet})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$
7	$N(b\text{-jets})/N(\text{jets})$	$N(b\text{-jets})/N(\text{jets})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$p_T(6^{\text{th}} \text{ jet})$
8	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$
9	$m_T^{\text{reco}}(t\bar{t})$	$N(\text{jets})$	$m_T^W$	$m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$
10	$N(\text{jets})$	$m_T^{\text{reco}}(t\bar{t})$	$p_T(1^{\text{st}} \text{ jet})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$
11	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$m_{\text{inv}}^{\text{reco}}(t\bar{t})$	$\log L_{\text{KLFitter}}$	$m_T^W$
12	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$	$p_T^{\text{reco}}(Z)$	$p_T(1^{\text{st}} \text{ jet})$
13	$\log L_{\text{KLFitter}}$	$\log L_{\text{KLFitter}}$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\log L_{\text{KLFitter}}$
14	$p_T(1^{\text{st}} \text{ jet})$	$m_T^{\text{reco}}(Z)$	$p_T(6^{\text{th}} \text{ jet})$	$p_T^{\text{reco}}(Z)$
15	$m_T^{\text{reco}}(Z)$	$p_T(1^{\text{st}} \text{ jet})$	$m_T^{\text{reco}}(Z)$	$\Delta R(\text{add. jets})$
16	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R^{t\bar{t}Z}(\text{add. jets})$	$\Delta R(\text{add. jets})$	$m_T^{\text{reco}}(Z)$
17	$p_T^{\text{reco}}(Z)$	$p_T^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$
18	$m_T^W$	$N(b\text{-jets})$	–	–
19	$N(b\text{-jets})$	$N(b\text{-jets})$	–	–
20	$\Delta R(\text{add. jets})$	$\Delta R(\text{add. jets})$	–	–
21	$m_{\text{inv}}^{\text{reco}}(Z)$	$m_{\text{inv}}^{\text{reco}}(Z)$	–	–

Tab. F.4.: Ranking of the BDT input variables for all signal regions before the training. Before the BDT training a variable decorrelation has been performed. The rankings are determined from the overlapt of the variables shapes. The last four positions for the exclusive signal regions are not occupied due to the smaller amount of input variables.

## References

- [1] E. Rutherford, *The scattering of alpha and beta particles by matter and the structure of the atom*, *Phil. Mag. Ser.6* **21** (1911) 669–688.
- [2] Professor Orme Masson F.R.S., *XXIV. The constitution of atoms*, *Philosophical Magazine* **41** (1921) 281–285,  
eprint: <http://dx.doi.org/10.1080/14786442108636219>,  
URL: <http://dx.doi.org/10.1080/14786442108636219>.
- [3] J. Chadwick, *The Existence of a Neutron*, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **136** (1932) 692–708,  
eprint: <http://rspa.royalsocietypublishing.org/content/136/830/692.full.pdf>, URL:  
<http://rspa.royalsocietypublishing.org/content/136/830/692>.
- [4] J. Chadwick, *Bakerian Lecture. The Neutron*, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **142** (1933) 1–25,  
eprint: <http://rspa.royalsocietypublishing.org/content/142/846/1.full.pdf>, URL:  
<http://rspa.royalsocietypublishing.org/content/142/846/1>.
- [5] 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>.
- [6] 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>.
- [7] Richard. P. Feynman, *The behavior of hadron collisions at extreme energies*, *Conf. Proc.* **C690905** (1969) 237–258.
- [8] Richard P. Feynman, *Very high-energy collisions of hadrons*, *Phys. Rev. Lett.* **23** (1969) 1415–1417.
- [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>.
- [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>.
- [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>.
- [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>.

- [13] J. D. Bjorken and S. L. Glashow, *Elementary Particles and SU(4)*, *Phys. Lett.* **11** (1964) 255–257.
- [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>.
- [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>.
- [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..>
- [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>.
- [18] F. CDF Collaboration Abe et al., *Observation of Top Quark Production in  $\bar{p}p$  Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74** (14 1995) 2626–2631,  
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [19] S. DØ Collaboration Abachi et al., *Observation of the Top Quark*,  
*Phys. Rev. Lett.* **74** (14 1995) 2632–2637,  
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [20] Sheldon L. Glashow, *The renormalizability of vector meson interactions*,  
*Nucl. Phys.* **10** (1959) 107–117.
- [21] Abdus Salam and John Clive Ward, *Weak and electromagnetic interactions*,  
*Nuovo Cim.* **11** (1959) 568–577.
- [22] 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>.
- [23] 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>.
- [24] 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>.
- [25] 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>.
- [26] M. Baak et al., *The global electroweak fit at NNLO and prospects for the LHC and ILC*,  
*Eur. Phys. J.* **C74** (2014) 3046, arXiv: 1407.3792 [hep-ph].
- [27] Giuseppe Degrand et al.,  
*Higgs mass and vacuum stability in the Standard Model at NNLO*, *JHEP* **08** (2012) 098,  
arXiv: 1205.6497 [hep-ph].

- [28] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, arXiv: 1207.7214 [hep-ex].
- [29] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, arXiv: 1207.7235 [hep-ex].
- [30] Hideki Yukawa, Shoichi Sakata, and Mitsuo Taketani, *On the Interaction of Elementary Particles. III*, *Progress of Theoretical Physics Supplement* **1** (1955) 24–45, URL: [+http://dx.doi.org/10.1143/PTPS.1.24](http://dx.doi.org/10.1143/PTPS.1.24).
- [31] Stephen P. Martin, *A Supersymmetry primer*, (1997), [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv: [hep-ph/9709356](http://arxiv.org/abs/hep-ph/9709356) [hep-ph].
- [32] 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*, *JHEP* **11** (2015) 172. 31 p, Comments: 31 pages plus author list + cover pages (47 pages total), 11 figures, 8 tables, submitted to JHEP, All figures including auxiliary figures are available at [http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/ATLAS-CONF-2015-024/](http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/ATLAS-CONF-2015-024), URL: <https://cds.cern.ch/record/2053330>.
- [33] CMS Collaboration, *Observation of top quark pairs produced in association with a vector boson in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*, *JHEP* **01** (2015) 096. 52 p, Replaced with published version. Added journal reference and DOI, URL: <http://cds.cern.ch/record/2057196>.
- [34] C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.
- [35] C. Berger, *Elementarteilchenphysik, Von den Grundlagen zu den modernen Experimenten*, 3rd ed., Springer Spektrum, 2014, ISBN: 978-3540231431.
- [36] D. Griffiths, *Introduction to Elementary Particles*, 2nd ed., WILEY-VCH Verlag GmbH & Co. KGaA, 2008, ISBN: 978-3527406012.
- [37] Mark Thomson, *Modern Particle Physics*, 3rd ed., Cambridge University Press, 2015, ISBN: 978-1107034266.
- [38] 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](http://oupanbackfile/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>.
- [39] J. Schwinger, *Quantum Electrodynamics. I. A Covariant Formulation*, *Physical Review* **74** (1948) 1439–1461.
- [40] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, *Physical Review* **80** (1950) 440–457.
- [41] 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>.

- [42] 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.
- [43] K. Bethge and U. E. Schröder, *Elementarteilchen und ihre Wechselwirkungen*, 3rd ed., WILEY-VCH Verlag GmbH & Co. KGaA, 2006, ISBN: 978-3527405879.
- [44] 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](https://arxiv.org/abs/hep-ex/0610037) [[hep-ex](#)], URL: <http://www.slac.stanford.edu/econf/C060706/pdf/0610037.pdf>.
- [45] Siegfried Bethke, *World Summary of  $\alpha_s$  (2012)*, (2012), Nucl. Phys. Proc. Suppl. 234, 229(2013), arXiv: [1210.0325](https://arxiv.org/abs/1210.0325) [[hep-ex](#)].
- [46] 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>.
- [47] 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>.
- [48] Jeff Greensite, *An introduction to the Confinement Problem*, Springer-Verlag, 2011, ISBN: 978-3-642-14381-6.
- [49] Reinhard Alkofer and J. Greensite, *Quark Confinement: The Hard Problem of Hadron Physics*, *J. Phys.* **G34** (2007) S3, arXiv: [hep-ph/0610365](https://arxiv.org/abs/hep-ph/0610365) [[hep-ph](#)].
- [50] Nathan Philip Jurik and Tomasz Skwarnicki, *Observation of  $J/\psi$   $p$  resonances consistent with pentaquark states in  $\Lambda_b^0 \rightarrow J/\psi K^0 p$  decays*, Presented 01 Aug 2016, 2016, URL: <https://cds.cern.ch/record/2206806>.
- [51] G. Dissertori, I. Knowles, and M. Schmelling, *Quantum Chromodynamics*, 1st ed., Oxford University Press, 2003, ISBN: 978-0199566419.
- [52] Timo van Ritbergen and Robin G. Stuart, *Complete 2-Loop Quantum Electrodynamics 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>.
- [53] 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](https://arxiv.org/abs/hep-ph/9909436) [[hep-ph](#)].
- [54] 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** (4 2011) 041803, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.106.041803>.
- [55] H. D. Perkins, *Introduction to High Energy Physics*, 4th ed., Cambridge University Press, 2000, ISBN: 978-0521621960.
- [56] Peter Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*, 2nd ed., Springer-Verlag, 1995, ISBN: 978-3540584865.
- [57] Nicola Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533, [648(1963)].

- [58] 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)].
- [59] 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)].
- [60] Guido Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298–318.
- [61] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv: 0901.0002 [hep-ph].
- [62] ATLAS, CDF, CMS, and DØ Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass*, tech. rep. arXiv:1403.4427. ATLAS-CONF-2014-008. CDF-NOTE-11071. CMS-PAS-TOP-13-014. D0-NOTE-6416, Comments: 34 pages, 7 figures: CERN, 2014, URL: <https://cds.cern.ch/record/1669819>.
- [63] 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].
- [64] LHCTopWG, *Summary plots from the ATLAS Top physics group*, 2017, URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/> (visited on 08/29/2017).
- [65] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06** (2015) 184, arXiv: 1504.03446 [hep-ph].
- [66] ATLAS Collaboration, *Measurement of the  $t\bar{t}Z$  and  $t\bar{t}W$  production cross sections in multilepton final states using  $3.2 \text{ fb}^{-1}$  of  $pp$  collisions at 13 TeV at the LHC*, tech. rep. ATLAS-CONF-2016-003, CERN, 2016, URL: <https://cds.cern.ch/record/2138947>.
- [67] CMS Collaboration, *Measurement of top pair-production in association with a  $W$  or  $Z$  boson in  $pp$  collisions at 13 TeV*, tech. rep. CMS-PAS-TOP-17-005, CERN, 2017, URL: <https://cds.cern.ch/record/2264544>.
- [68] ATLAS Collaboration, *Search for  $t\bar{t}Z$  production in the three lepton final state with  $4.7 \text{ fb}^{-1}$  of  $\sqrt{s} = 7 \text{ TeV}$   $pp$  collision data collected by the ATLAS detector*, tech. rep. ATLAS-CONF-2012-126, CERN, 2012, URL: <https://cds.cern.ch/record/1474643>.
- [69] ATLAS Collaboration, *Measurement of the  $t\bar{t}W$  and  $t\bar{t}Z$  production cross sections in  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$  with the ATLAS detector*. Tech. rep. ATLAS-CONF-2015-032, CERN, 2015, URL: <https://cds.cern.ch/record/2038143>.
- [70] ATLAS Collaboration, *Measurement of the  $t\bar{t}Z$  and  $t\bar{t}W$  production cross sections in multilepton final states using  $3.2 \text{ fb}^{-1}$  of  $pp$  collisions at 13 TeV at the LHC*, tech. rep. ATLAS-CONF-2016-003, CERN, 2016, URL: <https://cds.cern.ch/record/2138947>.
- [71] CMS Collaboration, *Measurement of top quark-antiquark pair production in association with a  $W$  or  $Z$  boson in  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$* , Eur. Phys. J. C **74** (2014) 3060. 26 p, Comments: Replaced with published version. Added journal reference and DOI, URL: <https://cds.cern.ch/record/1712680>.

- [72] 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)].
- [73] 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.
- [74] P. Bagnaia et al., *Evidence for  $Z^0 \rightarrow e^+ e^-$  at the CERN anti- $p p$  Collider*, *Phys. Lett.* **129B** (1983) 130–140.
- [75] ALEPH Collaboration et al., *Precision Electroweak Measurements on the Z Resonance*, *Phys. Rep.* **427** (2005) 257. 302 p,  
URL: <https://cds.cern.ch/record/892831>.
- [76] Cinzia De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, 2016, URL: <https://cds.cern.ch/record/2197559>.
- [77] 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).
- [78] Oliver Sim Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, 2004, URL: <http://cds.cern.ch/record/782076>.
- [79] ATLAS Collaboration, *Improved luminosity determination in  $pp$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector at the LHC*, *Eur. Phys. J.* **C73** (2013) 2518, arXiv: 1302.4393 [hep-ex].
- [80] Simon White, *Luminosity Scans at the LHC. Luminosity Scans at LHC*, (2011), URL: <https://cds.cern.ch/record/1357865>.
- [81] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [82] Joao Pequeno, *Computer generated image of the whole ATLAS detector*, 2008, URL: <http://cds.cern.ch/record/1095924>.
- [83] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*, Technical Design Report ATLAS, CERN, 1997, URL: <http://cds.cern.ch/record/338080>.
- [84] ATLAS Collaboration, *ATLAS central solenoid: Technical Design Report*, Technical Design Report ATLAS, Electronic version not available, CERN, 1997, URL: <http://cds.cern.ch/record/331067>.
- [85] ATLAS Collaboration, *ATLAS barrel toroid: Technical Design Report*, Technical Design Report ATLAS, Electronic version not available, CERN, 1997, URL: <http://cds.cern.ch/record/331065>.
- [86] ATLAS Collaboration, *ATLAS end-cap toroids: Technical Design Report*, Technical Design Report ATLAS, Electronic version not available, CERN, 1997, URL: <http://cds.cern.ch/record/331066>.
- [87] R. L. Gluckstern, *Uncertainties in track momentum and direction, due to multiple scattering and measurement errors*, *Nucl. Instrum. Meth.* **24** (1963) 381–389.
- [88] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, Volume 1*, Technical Design Report ATLAS, CERN, 1997, URL: <http://cds.cern.ch/record/331063>.

- 
- [89] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, Volume 2*, Technical Design Report ATLAS, CERN, 1997, URL: <http://cds.cern.ch/record/331064>.
- [90] Joao Pequena, *Computer generated image of the ATLAS inner detector*, 2008, URL: <http://cds.cern.ch/record/1095926>.
- [91] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*, Technical Design Report ATLAS, CERN, 1998, URL: <http://cds.cern.ch/record/381263>.
- [92] ATLAS Collaboration, *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>.
- [93] Joao Pequena, *Computer Generated image of the ATLAS calorimeter*, 2008, URL: <http://cds.cern.ch/record/1095927>.
- [94] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical Design Report ATLAS, CERN, 1996, URL: <http://cds.cern.ch/record/331061>.
- [95] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*, Technical Design Report ATLAS, CERN, 1996, URL: <http://cds.cern.ch/record/331062>.
- [96] Joao Pequena, *Computer generated image of the ATLAS Muons subsystem*, 2008, URL: <http://cds.cern.ch/record/1095929>.
- [97] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, Technical Design Report ATLAS, CERN, 1997, URL: <http://cds.cern.ch/record/331068>.
- [98] Marco Bruschi, *The new ATLAS/LUCID detector*, tech. rep. ATL-FWD-PROC-2015-002, CERN, 2015, URL: <https://cds.cern.ch/record/2025000>.
- [99] Abdel Khalek et al., *The ALFA Roman Pot Detectors of ATLAS*, JINST **11** (2016) P11013. 36 p, 37 pages, 22 figures, final version published in JINST, URL: <http://cds.cern.ch/record/2212402>.
- [100] ATLAS Collaboration, *The ATLAS Trigger System: Ready for Run 2*, tech. rep. ATL-DAQ-PROC-2015-038, CERN, 2015, URL: <https://cds.cern.ch/record/2058218>.
- [101] ATLAS Collaboration, *The Run-2 ATLAS Trigger System*, tech. rep. ATL-DAQ-PROC-2016-003, CERN, 2016, URL: <https://cds.cern.ch/record/2133909>.
- [102] ATLAS Collaboration, *Operation of the Upgraded ATLAS Level-1 Central Trigger System*, *J. Phys. Conf. Ser.* **664** (2015) 082013.
- [103] ATLAS Collaboration, *Upgrade of the ATLAS Central Trigger for LHC Run-2*, tech. rep. ATL-DAQ-PROC-2014-042. 02, CERN, 2014, URL: <https://cds.cern.ch/record/1969488>.

- [104] ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, Technical Design Report ATLAS, CERN, 2003, URL: <https://cds.cern.ch/record/616089>.
- [105] ATLAS Collaboration, *Operation of the upgraded ATLAS Central Trigger Processor during the LHC Run 2*, tech. rep. ATL-DAQ-PROC-2015-050. 02, CERN, 2015, URL: <https://cds.cern.ch/record/2063035>.
- [106] Christoph Eck et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Technical Design Report LCG, CERN, 2005, URL: <http://cds.cern.ch/record/840543>.
- [107] ATLAS Collaboration, *Simulation of Pile-up in the ATLAS Experiment*, *J. Phys. Conf. Ser.* **513** (2014) 022024.
- [108] 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>.
- [109] The ATLAS Collaboration, *2012  $Z \rightarrow \mu\mu$  event with high pileup*, 2012, URL: [https://twiki.cern.ch/twiki/pub/AtlasPublic/EventDisplayStandAlone/2012\\_highPileup.png](https://twiki.cern.ch/twiki/pub/AtlasPublic/EventDisplayStandAlone/2012_highPileup.png) (visited on 08/11/2017).
- [110] The ATLAS Collaboration, *Luminosity results for Run-2*, 2012, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (visited on 08/11/2017).
- [111] Andy Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, arXiv: 1101.2599 [hep-ph].
- [112] Glen Cowan, *Statistical Data Analysis*, Oxford University Press, 1989, ISBN: 978-0198501558.
- [113] Stefan Höche, *Introduction to parton-shower event generators*, (2014), arXiv: 1411.4085 [hep-ph], URL: <https://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf>.
- [114] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31–145.
- [115] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys.* **B238** (1984) 492–528.
- [116] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063, arXiv: hep-ph/0109231 [hep-ph].
- [117] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, *JHEP* **08** (2002) 015, arXiv: hep-ph/0205283 [hep-ph].
- [118] F. Caravaglios, Michelangelo L. Mangano, M. Moretti, and R. Pittau, *A New approach to multijet calculations in hadron collisions*, *Nucl. Phys.* **B539** (1999) 215–232, arXiv: hep-ph/9807570 [hep-ph].

- 
- [119] 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].
- [120] 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].
- [121] Paolo Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, **JHEP** **11** (2004) 040, arXiv: hep-ph/0409146 [hep-ph].
- [122] 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].
- [123] 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].
- [124] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer, *MadGraph 5 : Going Beyond*, **JHEP** **06** (2011) 128, arXiv: 1106.0522 [hep-ph].
- [125] Stefano Frixione and Bryan R. Webber, *Matching NLO QCD computations and parton shower simulations*, **JHEP** **06** (2002) 029, arXiv: hep-ph/0204244 [hep-ph].
- [126] T. Gleisberg et al., *Event generation with SHERPA 1.1*, **JHEP** **02** (2009) 007, arXiv: 0811.4622 [hep-ph].
- [127] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands, *PYTHIA 6.4 Physics and Manual*, **JHEP** **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].
- [128] Torbjörn Sjöstrand, 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].
- [129] M. Bahr et al., *Herwig++ Physics and Manual*, **Eur. Phys. J.** **C58** (2008) 639–707, arXiv: 0803.0883 [hep-ph].
- [130] K. Arnold et al., *Herwig++ 2.6 Release Note*, (2012), arXiv: 1205.4902 [hep-ph].
- [131] D. J. Lange, *The EvtGen particle decay simulation package*, **Nucl. Instrum. Meth.** **A462** (2001) 152–155.
- [132] Richard D. Ball et al., *Parton distributions for the LHC Run II*, **JHEP** **04** (2015) 040, arXiv: 1410.8849 [hep-ph].
- [133] Andrew Buckley, *ATLAS Pythia 8 tunes to 7 TeV data*, tech. rep. ATL-PHYS-PROC-2014-273, CERN, 2014, URL: <https://cds.cern.ch/record/1974411>.
- [134] Richard D. Ball et al., *Parton distributions with LHC data*, **Nucl. Phys.** **B867** (2013) 244–289, arXiv: 1207.1303 [hep-ph].

- [135] Stefano Carrazza, Stefano Forte, and Juan Rojo, *Parton Distributions and Event Generators, Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*, 2013 89–96, arXiv: [1311.5887 \[hep-ph\]](#), URL: <https://inspirehep.net/record/1266070/files/arXiv:1311.5887.pdf>.
- [136] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012, arXiv: [hep-ph/0201195 \[hep-ph\]](#).
- [137] Peter Z. Skands, *The Perugia Tunes, Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08): Perugia, Italy, October 27-31, 2008*, 2009 284–297, arXiv: [0905.3418 \[hep-ph\]](#), URL: [http://lss.fnal.gov/cgi-bin/find\\_paper.pl?conf-09-113](http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-09-113).
- [138] Peter Zeiler Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, *Phys. Rev.* **D82** (2010) 074018, arXiv: [1005.3457 \[hep-ph\]](#).
- [139] Federico Demartin, Benedikt Maier, Fabio Maltoni, Kentarou Mawatari, and Marco Zaro, *tWH associated production at the LHC*, *Eur. Phys. J.* **C77** (2017) 34, arXiv: [1607.05862 \[hep-ph\]](#).
- [140] Hung-Liang Lai et al., *New parton distributions for collider physics*, *Phys. Rev.* **D82** (2010) 074024, arXiv: [1007.2241 \[hep-ph\]](#).
- [141] 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 \[hep-ph\]](#).
- [142] Fabio Cascioli, Philipp Maierhöfer, Niccolo Moretti, Stefano Pozzorini, and Frank Siegert, *NLO matching for  $t\bar{t}b\bar{b}$  production with massive b-quarks*, *Phys. Lett.* **B734** (2014) 210–214, arXiv: [1309.5912 \[hep-ph\]](#).
- [143] M. R. Whalley, D. Bourilkov, and R. C. Group, *The Les Houches accord PDFs (LHAPDF) and LHAGLUE, HERA and the LHC: A Workshop on the implications of HERA for LHC physics. Proceedings, Part B*, 2005 575–581, arXiv: [hep-ph/0508110 \[hep-ph\]](#).
- [144] Andy Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132, arXiv: [1412.7420 \[hep-ph\]](#).
- [145] ATLAS Collaboration, *Measurement of the  $Z/\gamma^*$  boson transverse momentum distribution in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, *JHEP* **09** (2014) 145, arXiv: [1406.3660 \[hep-ex\]](#).
- [146] Stefan Höche, Frank Krauss, Marek Schonherr, and Frank Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](#).
- [147] Tanju Gleisberg and Stefan Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](#).
- [148] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](#).
- [149] Stefan Gieseke, Christian Röhr, and Andrzej Siódmok, *Colour reconnections in Herwig++*, *The European Physical Journal C* **72** (2012) 2225, URL: <https://doi.org/10.1140/epjc/s10052-012-2225-5>.

- 
- [150] Bo Andersson, *The Lund Model*, 7th ed., Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, 1998, ISBN: 0521017343, URL: <http://lccn.loc.gov/2006274063>.
- [151] 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].
- [152] Andreas P Contogouris and Z Merebashvili, *Approximate Next-to-Leading Order and Next-to-Next-to-Leading Order Corrections*, *Int. J. Mod. Phys. A* **18** (2002) 957–966. 8 p, URL: <https://cds.cern.ch/record/552950>.
- [153] G. Aad et al., *The ATLAS Simulation Infrastructure*, *Eur. Phys. J.* **C70** (2010) 823–874, arXiv: 1005.4568 [physics.ins-det].
- [154] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303.
- [155] ATLAS Collaboration, *ATLAS Computing: technical design report*, Technical Design Report ATLAS, CERN, 2005, URL: <https://cds.cern.ch/record/837738>.
- [156] I Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, tech. rep. CERN-LHCC-2014-014. LCG-TDR-002, 2014, URL: <https://cds.cern.ch/record/1695401>.
- [157] 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>.
- [158] Morten Dam Jørgensen, *ATLAS Data Quality Monitoring*, 2016, URL: [https://atlasdqm.web.cern.ch/atlasdqm/grlgen/All\\_Good/data15\\_13TeV.periodAllYear\\_DetStatus-v79-repro20-02\\_DQDefects-00-02-02\\_PHYS\\_StandardGRL\\_All\\_Good\\_25ns.xml](https://atlasdqm.web.cern.ch/atlasdqm/grlgen/All_Good/data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml) (visited on 08/16/2017).
- [159] Morten Dam Jørgensen, *ATLAS Data Quality Monitoring*, 2017, URL: [https://atlasdqm.web.cern.ch/atlasdqm/grlgen/All\\_Good/data16\\_13TeV.periodAllYear\\_DetStatus-v88-pro20-21\\_DQDefects-00-02-04\\_PHYS\\_StandardGRL\\_All\\_Good\\_25ns.xml](https://atlasdqm.web.cern.ch/atlasdqm/grlgen/All_Good/data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml) (visited on 08/16/2017).
- [160] Joao Pequeno, *Event Cross Section in a computer generated image of the ATLAS detector*. 2008, URL: <https://cds.cern.ch/record/1096081>.
- [161] ATLAS Collaboration, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, 2007, URL: <https://cds.cern.ch/record/1020106>.
- [162] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics*, CERN, 2009, URL: <https://cds.cern.ch/record/1125884>.

- [163] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s}=13$  TeV*, *Eur. Phys. J.* **C76** (2016) 292, arXiv: 1603.05598 [hep-ex], URL: <https://cds.cern.ch/record/2139897>.
- [164] 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>.
- [165] 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>.
- [166] 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>.
- [167] 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>.
- [168] 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>.
- [169] J. Alitti et al., *Inclusive jet cross-section and a search for quark compositeness at the CERN  $\bar{p}p$  collider*, *Phys. Lett.* **B257** (1991) 232–240.
- [170] CDF Collaboration, F. Abe, et al., *Topology of three-jet events in  $\bar{p}p$  collisions at  $\sqrt{s}=1.8$  TeV*, *Phys. Rev. D* **45** (5 1992) 1448–1458, URL: <https://link.aps.org/doi/10.1103/PhysRevD.45.1448>.
- [171] DØ Collaboration, S. Abachi, et al., *Studies of topological distributions of inclusive three- and four-jet events in  $\bar{p}p$  collisions at  $\sqrt{s}=1800$  GeV with the DØ detector*, *Phys. Rev. D* **53** (11 1996) 6000–6016, URL: <https://link.aps.org/doi/10.1103/PhysRevD.53.6000>.
- [172] Jeff Tseng and Hannah Evans, *Sequential recombination algorithm for jet clustering and background subtraction*, *Phys. Rev.* **D88** (2013) 014044, arXiv: 1304.1025 [hep-ph].
- [173] S Catani, Yu L Dokshitzer, Michael H Seymour, and Bryan R Webber, *Longitudinally-invariant  $k_{\perp}$ -clustering algorithms for hadron-hadron collisions*, *Nucl. Phys. B* **406** (1993) 187–224. 38 p, URL: <http://cds.cern.ch/record/246812>.
- [174] S D Ellis and Davison Eugene Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160–3166. 15 p, URL: <http://cds.cern.ch/record/248558>.
- [175] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, arXiv: hep-ph/9707323 [hep-ph].

- 
- [176] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, (1998) 270–279, arXiv: [hep-ph/9907280](https://arxiv.org/abs/hep-ph/9907280) [hep-ph], URL: [https://inspirehep.net/record/484872/files/arXiv:hep-ph\\_9907280.pdf](https://inspirehep.net/record/484872/files/arXiv:hep-ph_9907280.pdf).
- [177] Gavin P. Salam, *Towards Jetography*, *Eur. Phys. J.* **C67** (2010) 637–686, arXiv: [0906.1833](https://arxiv.org/abs/0906.1833) [hep-ph].
- [178] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The Anti- $k(t)$  jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph].
- [179] 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>.
- [180] 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>.
- [181] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005, arXiv: [0802.1188](https://arxiv.org/abs/0802.1188) [hep-ph].
- [182] ATLAS Collaboration, *ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2*, tech. rep. ATL-PHYS-PROC-2017-045, CERN, 2017, URL: <https://cds.cern.ch/record/2263777>.
- [183] 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>.
- [184] *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>.
- [185] ATLAS Collaboration, *Determination of the jet energy scale and resolution at ATLAS using  $Z/\gamma$ -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>.
- [186] 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>.
- [187] 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>.
- [188] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, tech. rep. ATLAS-CONF-2015-029, CERN, 2015, URL: <https://cds.cern.ch/record/2037702>.

- [189] 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>.
- [190] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, tech. rep. ATLAS-CONF-2011-102, CERN, 2011, URL: <https://cds.cern.ch/record/1369219>.
- [191] ATLAS Collaboration, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, Journal of Physics: Conference Series **119** (2008) 032032, URL: <http://stacks.iop.org/1742-6596/119/i=3/a=032032>.
- [192] 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>.
- [193] 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>.
- [194] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at  $\sqrt{s} = 13$  TeV*, tech. rep. ATL-PHYS-PUB-2015-027, CERN, 2015, URL: <https://cds.cern.ch/record/2037904>.
- [195] 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>.
- [196] 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>.
- [197] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, tech. rep. ATLAS-CONF-2014-058, CERN, 2014, URL: <https://cds.cern.ch/record/1951336>.
- [198] Frederic Derue, *Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at  $\sqrt{s} = 13$  TeV with AnalysisTop-2.3.41*, tech. rep. ATL-COM-PHYS-2016-198, CERN, 2016, URL: <https://cds.cern.ch/record/2135116>.
- [199] ATLAS Collaboration, *Measurements of the  $t\bar{t}$  production cross-section in the dilepton and lepton-plus-jets channels and of the ratio of the  $t\bar{t}$  and Z boson cross-sections in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, tech. rep. ATLAS-CONF-2015-049, CERN, 2015, URL: <https://cds.cern.ch/record/2052605>.
- [200] Johannes Erdmann, Kevin Kröniger, Olaf Nackenhorst, and Arnulf Quadt, *Kinematic fitting of  $t\bar{t}$  events using a likelihood approach – the KL Fitter package*, Nucl. Instrum. Meth. **A748** (2014) 18–25, arXiv: 1312.5595 [hep-ex].

- [201] Johannes Erdmann et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, *Nucl. Instrum. Meth.* **A748** (2014) 18–25, arXiv: 1312.5595 [hep-ex].
- [202] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81–86.
- [203] Allen Caldwell, Daniel Kollar, and Kevin Kroninger, *BAT: The Bayesian Analysis Toolkit*, *Comput. Phys. Commun.* **180** (2009) 2197–2209, arXiv: 0808.2552 [physics.data-an].
- [204] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343–367.
- [205] Andreas Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv: physics/0703039 [PHYSICS].
- [206] 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>.
- [207] 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>.
- [208] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, *Journal of Educational Psychology* **24** (933) 417–441.
- [209] ATLAS Collaboration, *Luminosity determination in pp collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **76** (2016) 653. 71 p, Comments: 53 pages plus author list + cover pages (71 pages total), 19 figures, 9 tables, submitted to EPJC, All figures including auxiliary figures are available at <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/DAPR-2013-01>, URL: <https://cds.cern.ch/record/2208146>.
- [210] 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>.
- [211] ATLAS Collaboration, *Simulation of top quark production for the ATLAS experiment at  $\sqrt{s} = 13$  TeV*, tech. rep. ATL-PHYS-PUB-2016-004, CERN, 2016, URL: <https://cds.cern.ch/record/2120417>.
- [212] ATLAS Collaboration, *Comparison of Monte Carlo generator predictions for gap fraction and jet multiplicity observables in top-antitop events*, tech. rep. ATL-PHYS-PUB-2014-005, CERN, 2014, URL: <https://cds.cern.ch/record/1703034>.
- [213] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J.* **C77** (2017) 466, arXiv: 1703.10485 [hep-ex].
- [214] Boris Lemmer, *KLFitterTF*, 2012, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/KLFitterTF> (visited on 06/11/2017).

- [215] Boris Lemmer et al., *Transfer Functions for the TopKLFitter - Supporting Plots*, 2012,  
URL: <https://klfittertf.web.cern.ch/klfittertf/> (visited on 07/11/2017).
- [216] 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>.
- [217] 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>.
- [218] Llew Mason, Jonathan Baxter, Peter Bartlett, and Marcus Frean,  
*Boosting Algorithms As Gradient Descent in Function Space*, *NIPS* **7** (1999) 512–518,  
URL: <http://dl.acm.org/citation.cfm?id=3009657.3009730>.
- [219] 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>.
- [220] 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](http://dx.doi.org/10.1016/S0167-9473(01)00065-2).

## List of Figures

2.1. Electron-positron scattering . . . . .	7
2.2. Running coupling constants . . . . .	8
2.3. Gluon self-coupling . . . . .	9
2.4. Parton distribution function . . . . .	14
2.5. Single top quark production . . . . .	16
2.6. Top quark pair production at LO . . . . .	17
2.7. Top quark pair production at NLO . . . . .	18
2.8. Top quark pair production cross section measurements at 13 TeV . . . . .	19
2.9. The dileptonic $t\bar{t}$ decay . . . . .	20
2.10. The lepton+jets $t\bar{t}$ decay . . . . .	20
2.11. The fully hadronic $t\bar{t}$ decay . . . . .	21
2.12. Top quark pair production in association with a $Z$ boson . . . . .	22
2.13. The $t\bar{t}Z$ 1-lepton channel . . . . .	22
3.1. The CERN accelerator complex . . . . .	26
3.2. The ATLAS detector . . . . .	29
3.3. The ATLAS magnet system . . . . .	31
3.4. The ATLAS inner detector . . . . .	32
3.5. The ATLAS inner detector parts . . . . .	33
3.6. The ATLAS calorimeter system . . . . .	35
3.7. The ATLAS muon spectrometer . . . . .	37
3.8. The ATLAS level-1 trigger . . . . .	39
3.9. High pile-up environment . . . . .	40
3.10. Average number of interactions . . . . .	41
4.1. Schematic illustration of a proton-proton collision . . . . .	44
6.1. Kinematic Monte Carlo distributions . . . . .	66
6.2. Jet-related Monte Carlo distributions . . . . .	67
6.3. Composite Monte Carlo distributions . . . . .	68
6.4. Number of valid permutations . . . . .	70
6.5. KLFilter logarithmic likelihood distribution . . . . .	73
6.6. Significances in jet phase space . . . . .	74
6.7. Reconstructed mass distributions in SR1 . . . . .	77
6.8. Reconstructed mass distributions in SR2 . . . . .	78
6.9. Reconstructed mass distributions in SR3 . . . . .	79
6.10. Reconstructed mass distributions in SR4 . . . . .	80
6.11. $Z$ boson mass distribution for different ( $b$ -)jet multiplicities . . . . .	81
6.12. KLFilter logarithmic likelihood distribution for different ( $b$ -)jet multiplicities . . . . .	82
6.13. KLFilter jet indices in SR2 . . . . .	83
6.14. Correct KLFilter jet indices for different ( $b$ -)jet multiplicities . . . . .	85
6.15. Jet flavour labels in SR2 . . . . .	87
6.16. KLFilter jet indices for inclusive selection criteria . . . . .	88
6.17. KLFilter jet indices for exclusive selection criteria . . . . .	89

7.1. Illustration of a decision tree . . . . .	92
7.2. TMVA separation methods . . . . .	93
7.3. Discriminating ( $b$ -)jet multiplicity-related variables . . . . .	101
7.4. Discriminating variables in SR2 I . . . . .	102
7.5. Discriminating variables in SR2 II . . . . .	103
7.6. Discriminating variables in SR2 III . . . . .	104
7.7. BDT training results for SR2 . . . . .	105
7.8. BDT evaluation results for SR2 at $36.1 \text{ fb}^{-1}$ . . . . .	105
7.9. BDT training results for SR2 with variable decorrelation with variable decorrelation . . . . .	106
7.10. BDT evaluation results for SR2 at $36.1 \text{ fb}^{-1}$ . . . . .	106
8.1. BDT evaluation results for SR2 at $100 \text{ fb}^{-1}$ . . . . .	108
9.1. Systematic parameter for $b$ -tagging efficiency for SR1 . . . . .	112
9.2. Theoretical uncertainties on input variables in SR2 . . . . .	113
9.3. Systematic uncertainties for $b$ -tagging efficiency in SR2 . . . . .	114
9.4. Theoretical uncertainties on the BDT response in SR2 . . . . .	115
B.1. KLfitter transfer function parametrisation . . . . .	124
E.1. Significances with statistical uncertainties . . . . .	131
E.2. Significances with statistical and systematic uncertainties . . . . .	131
E.3. KLfitter jet indices in SR1 . . . . .	132
E.4. KLfitter jet indices in SR3 . . . . .	133
E.5. KLfitter jet indices in SR4 . . . . .	134
E.6. Jet flavour labels in SR1 . . . . .	135
E.7. Jet flavour labels in SR3 . . . . .	136
E.8. Jet flavour labels in SR4 . . . . .	137
E.9. Discriminating variables $H_T$ and $H_T^{\text{lep}}$ . . . . .	138
E.10. Discriminating variables $S_T$ and $p_T^{\text{reco}}(Z)$ . . . . .	139
E.11. Discriminating variables $S_T$ and $p_T(Z)$ . . . . .	140
E.12. Discriminating variables $m_T^{\text{reco}}(Z)$ and $m_T^{\text{reco}}(t\bar{t})$ . . . . .	141
E.13. Discriminating variables $p_T(7^{\text{th}} \text{ jet})$ and $p_T(1^{\text{st}} \text{ jet})$ . . . . .	142
E.14. Discriminating variables $\Delta R(\text{add. jets})$ and $\Delta R^{t\bar{t}Z}(\text{add. jets})$ . . . . .	143
E.15. Discriminating variables $m_{\text{inv}}^{\text{reco}}(Z)$ and $m_{\text{inv}}^{\text{reco}}(t^{\text{had}})$ . . . . .	144
E.16. Discriminating variables $m_{\text{inv}}^{\text{reco}}(t^{\text{lep}})$ and $m_{\text{inv}}^{\text{reco}}(t\bar{t})$ . . . . .	145
E.17. Discriminating variables $m_T^W$ and $\log L_{\text{KLfitter}}$ . . . . .	146
E.18. BDT training in SR1, SR3 and SR4 . . . . .	147
E.19. BDT training in SR1, SR3 and SR4 . . . . .	148
E.20. BDT evaluation in SR1, SR3 and SR4 at $36.1 \text{ fb}^{-1}$ . . . . .	149
E.21. BDT evaluation with variable decorrelation in SR1, SR3 and SR4 at $36.1 \text{ fb}^{-1}$ . . . . .	150
E.22. BDT evaluation in SR1, SR3 and SR4 at $100 \text{ fb}^{-1}$ . . . . .	151
E.23. Correlation matrix SR1 . . . . .	152
E.24. Correlation matrix SR2 . . . . .	152
E.25. Correlation matrix SR3 . . . . .	153
E.26. Correlation matrix SR4 . . . . .	153
E.27. Systematic uncertainties for $b$ -tagging efficiency on input variables in SR1 . . . . .	155
E.28. Theoretical uncertainties on input variables in SR1, SR3 and SR4 I . . . . .	156
E.29. Theoretical uncertainties on input variables in SR1, SR3 and SR4 II . . . . .	157

E.30. Systematic uncertainties for $b$ -tagging efficiency on BDT response in SR1 . . . . .	158
E.31. Theoretical uncertainties on BDT response in SR1, SR3 and SR4 I . . . . .	159
E.32. Theoretical uncertainties on BDT response in SR1, SR3 and SR4 II . . . . .	160



## List of Tables

2.1. Fundamental particles of the Standard Model – bosons . . . . .	3
2.2. Fundamental particles of the Standard Model – fermions . . . . .	4
2.3. Branching fractions for the $W$ boson . . . . .	19
2.4. Branching fractions for the $Z$ boson . . . . .	23
4.1. Monte Carlo datasets . . . . .	51
6.1. Event yields after selection . . . . .	62
6.2. Event yield ratios and singificance . . . . .	63
6.3. Used triggers . . . . .	64
6.4. Number of valid permutations . . . . .	70
6.5. Event yields for all signal regions . . . . .	75
6.6. Event ratio and significances at $36.1 \text{ fb}^{-1}$ . . . . .	76
6.7. Jet flavour labels . . . . .	84
7.1. BDT separation techniques . . . . .	92
7.2. BDT configuration parameters . . . . .	96
7.3. BDT input variables . . . . .	97
7.4. Significances after BDT evaluation ar $36.1 \text{ fb}^{-1}$ . . . . .	100
7.5. Significances after BDT evaluation with variable decorrelation . . . . .	100
8.1. Event yield ratios and singificances at $100 \text{ fb}^{-1}$ . . . . .	107
8.2. Significances after BDT evaluation at 100 % . . . . .	108
10.1. Summary of significances . . . . .	118
F.1. BDT variable ranking after training . . . . .	161
F.2. BDT variable ranking before training . . . . .	162
F.3. BDT variable ranking after training with variable decorrelation . . . . .	163
F.4. BDT variable ranking before training with variable decorrelation . . . . .	164



## Acknowledgement

I would like to take the opportunity to thank all those people who have accompanied me so far!

In particular I want to name:

- My family, who have always supported me without hesitation
- My supervisors PD Dr. Jeannine Wagner-Kuhr and Prof. Dr. Otmar Biebel, who always took the time to discuss and answer my questions
- Prof. Dr. Dorothee Schaile, who gave me the opportunity to do my Master's thesis at her chair
- My colleagues of the chair for the excellent atmosphere and their unlimited readiness to help me whenever I had problems, particularly I want to mention David Handl, Dr. Friedrich Hönig, Dr. Thomas Maier, Balthasar Schachtner, Nikolai Hartmann, Andrea Matic and Dr. Michael Bender
- Elke Grimm-Zeidler for her continuous and friendly support with any administrative issue
- Daniela Köck and Paola Arrubarrena, who invited me from time to time to a ping-pong match when I really needed a break
- Clara Leitgeb, who took the trouble to proofread this thesis



## **Erklärung**

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den 15. September 2017

Unterschrift

