Measurement of the CKM matrix element $|V_{ts}|^2$

Dissertation der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Christopher Gerhard Unverdorben
geboren in Bamberg

München, März 2015

Erstgutachter: Prof. Dr. Otmar Biebel
Zweitgutachter: Prof. Dr. Christian Kiesling

Tag der mündlichen Prüfung: 28. April 2015

You never know where a breakthrough might occur,
so please pursue various possibilities for potential research.

Makoto Kobayashi
(Coinventor of the CKM matrix, Nobel Prize laureate)

Abstract

This is the first direct measurement of the CKM matrix element $|V_{ts}|$, using data collected by the ATLAS detector in 2012 at $\sqrt{s} = 8$ TeV pp-collisions with a total integrated luminosity of 20.3 fb^{-1}. The analysis is based on 112 171 reconstructed $t\bar{t}$ candidate events in the *lepton+jets* channel, having a purity of 90.0 %. 183 $t\bar{t} \to W^+W^-b\bar{s}$ decays are expected (charge conjugation implied), which are available for the extraction of the CKM matrix element $|V_{ts}|^2$. To identify these rare decays, several observables are examined, such as the properties of jets, tracks and of b-quark identification algorithms. Furthermore, the s-quark hadrons K^0_S are considered, reconstructed by a kinematic fit. The best observables are combined in a multivariate analysis, called "boosted decision trees". The responses from Monte Carlo simulations are used as templates for a fit to data events yielding a significance value of 0.7σ for $t \to s + W$ decays. An upper limit of

$$|V_{ts}|^2 < 1.74\,\%$$

at 95 % confidence level is set, including all systematic and statistical uncertainties. So this analysis, using a direct measurement of the CKM matrix element $|V_{ts}|^2$, provides the best direct limit on $|V_{ts}|^2$ up to now.

Zusammenfassung

Diese Arbeit beschreibt die erste direkte Messung des CKM-Matrixelements $|V_{ts}|$. Sie basiert auf Daten, die im Jahr 2012 mit dem ATLAS Detektor bei $\sqrt{s} = 8$ TeV pp-Kollisionen und einer integrierten Luminosität von 20.3^{-1} fb aufgezeichnet wurden. Insgesamt sind 112 171 mögliche $t\bar{t}$-Ereignisse im *Lepton+Jets*-Kanal mit einer Reinheit von 90.0 % rekonstruierbar, die für die Hauptanalyse zur Verfügung stehen. Laut Vorhersage zerfallen hiervon 183 Ereignisse über $t\bar{t} \to W^+W^-b\bar{s}$ (inkl. Ladungskonjugation) und können für die Bestimmung des Betragsquadrats $|V_{ts}|^2$ verwendet werden. Für eine Identifikation dieser seltenen Zerfälle werden verschiedenste Observablen untersucht, wie z.B. die Eigenschaften von Jets und Tracks sowie von b-Quark Identifikationsmethoden. Darüber hinaus werden Hadronen von s-Quarks betrachtet, die K^0_S Teilchen, die über einen *kinematischen Fit* rekonstruiert werden. Die Observablen mit den besten Eigenschaften werden anschließend in einer multivariaten Analyse (“Boosted decision trees”) zu einer einzigen zusammengefasst. Die dazugehörigen Monte-Carlo-Simulationen werden dann als Template für eine Beschreibung der Daten verwendet, was als Ergebnis zu einem Signifikanz-Wert von 0.7σ für $t \to s + W$ Zerfälle führt. Insgesamt ergibt sich ein oberes Limit von

$$|V_{ts}|^2 < 1.74\,\%.$$

Dieses wurde bezüglich eines Vertrauensbereiches von 95 % bestimmt, unter Berücksichtigung sämtlicher systematischen und statischen Unsicherheiten. Diese Arbeit, basierend auf einer direkten Messung des CKM-Matrixelements $|V_{ts}|^2$, führt somit zum bisher besten direkten Limit für $|V_{ts}|^2$.

Contents

1. Introduction

The current knowledge of elementary particles and their interactions (except gravity) is summarised in the Standard Model of Particle Physics (SM). In the last decades, multiple high precision tests have been performed, but only small extensions, like neutrino masses, needed to be added to the theory. Despite its huge success, the SM is not capable to describe our surroundings to the full extent. Remaining questions concern e.g. the description of the vacuum energy density, the inclusion of gravity, the imbalance matter vs. anti-matter and the nature of dark energy and dark matter. Those aspects require an extension of the SM or even the formulation of a new theory which comprises the current model.

New physics phenomena beyond the current SM descriptions are expected at higher energies. To detect those, there exist generally two approaches: First, direct searches for new particles or, second, indirect searches based on precision measurements. The experiments at the Large Hadron Collider (LHC) allow both procedures, since with its high centre-of-mass energy, several previously inaccessible regions of physics phase space can now be examined. A typical direct analysis is based on the detection of a new particle in that phase space, while indirect precision measurements evaluate properties of already known particles. An example for the latter are quark-loop processes, mediated by the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM}, which could be modified by contributions from physics beyond the SM.

The main purpose of this matrix is the transformation from quark mass eigenstates to weak eigenstates, by the following equation:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{weak} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass}$$

The matrix elements in a squared form, $|V_{ij}|^2$, represent the rate of the electroweak quark transition $i \to j + W^{\pm}$, where quark i and j can stem from different particle families. Since this matrix is one foundation of the SM, it is subject to intense scrutiny. Seven out of nine elements are directly measurable, while the two matrix elements V_{ts} and V_{td} can, up to now, only be determined indirectly in combination with other elements, e.g. $|V_{ts}^{*} V_{tb}|$.

In this thesis, the first direct measurement of the CKM matrix element $|V_{ts}|$ is now presented. It is based on the rate of the rare decay $t \to s + W$, governed by $|V_{ts}|^2$. The data used were recorded by ATLAS in 2012, from proton-proton-collisions at $\sqrt{s} = 8$ TeV. With about 5 million $t\bar{t}$ events, the number of generated top quarks is much higher than in former experiments. This enhances the chances for an extraction of the $t \rightarrow s + W$ decay, predicted to happen only in 0.2 % of all top decays. The main background for this process is the dominant $t \rightarrow b + W$ decay which is reduced by a multivariate analysis, i.e. a combination of several variables. In the final result, all statistical and systematic uncertainties are included.

Furthermore, during this study, a new method for the identification of muons from cosmic rays was developed, which is presented in addition to the default method.

Conventions

All parameters discussed in this analysis are based on “natural units” which are often used in high energy physics. It implies that the Planck constant $\hbar$ and the speed of light c are normalised to unity

$$\hbar \equiv 1 \quad \text{and} \quad c \equiv 1. \tag{1.1}$$

Furthermore, electrical charges are displayed with respect to the charge of an electron, also known as the elementary charge

$$e \approx 1,6022 \cdot 10^{-19}\ \mathrm{C} \quad [1]. \tag{1.2}$$

For the unit of energy, the notation *electron volt* (eV) is introduced, describing the energy gain of one electron flying through a potential difference of 1 V. With the choice of these notations, several further measurement parameters also adopt new units:

Momentum and mass: eV　　Time and length: 1/eV

2. Theoretical foundations

The development of physics models originates from an interplay of theoretical proposals and consequent experimental falsification. In particle physics, this concept led to the theoretical description of the *Standard Model of Particle Physics* (SM). It was mainly developed about 40 to 50 years ago and since then the core remained basically the same. Only small extensions, like e.g. neutrino oscillations and thus neutrino masses, had to be added. However, there are multiple hints for physics beyond the SM, which mostly arise from observations in cosmology and astrophysics. These indicate that the physics model as known today is not the whole picture and extensions are necessary.

In this chapter, the basic concepts of the SM are summarised based on [1, 2, 3, 4], where more detailed explanations can be found. The main focus here lies on the importance of the CKM matrix and its elements.

2.1. Standard Model of Particle Physics: The principles

All elementary particles and their interactions - except gravity - can be described by three fundamental theories: Quantum electrodynamics (QED), quantum chromodynamics (QCD) and electroweak (EW) interactions. These theories compose the *Standard Model of Particle Physics*, described by the gauge symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

with C representing the colour charge of the QCD in SU(3), L the impact on only left-handed particles of the weak isospin group $SU(2)$ and Y the weak hypercharge of $U(1)$. The dynamics in the model are based on the "principle of least action"

$$\delta S = 0 \quad \text{with} \quad S = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.2}$$

with $\mathcal{L}$ representing a Lorentz-invariant Lagrangian density, in the following simply called *Lagrangian*, which describes the aspects of the applied theories, i.e. the particles and fields involved.

The basis of the SM consists of two main aspects: *Invariance under local gauge transformation* and *renormalisability.*

A local gauge transformation is defined as

$$\psi(\boldsymbol{r}, t) \rightarrow \psi'(\boldsymbol{r}, t) = \exp\left[-iq\chi(\boldsymbol{r}, t)\right] \cdot \psi(\boldsymbol{r}, t). \tag{2.3}$$

with a field $\psi(\boldsymbol{r}, t)$ and an associated charge q of the respective theory. The concept of invariance indicates that the application of this transformation does not change any measurable quantities, i.e. the Lagrangian stays invariant. One consequence is formulated in Noether's theorem, which predicts a conserved charge q for each gauge symmetric field. In the SM, hence, there are three fundamental conservations of charges, one for each group theory, like the colour charge of the QCD.

The second aspect, *renormalisation*, relates to the concept of perturbation theory: Most problems in the SM are not directly solvable, therefore, a perturbative approximation is used, which quantifies the theoretical aspects in terms of power series. The leading order term is already an approximative result while additional higher order terms improve the description towards the exact solution, but in general increase also the complexity. Nevertheless, not all terms have to be considered for a valid result, thus yielding huge advantages in calculation. However, due to the approximative concept, energy-scale dependent variables might cause non-physical infinitive values. To still obtain valid results, those infinities are absorbed in measurable quantities (e.g. charge, mass) by using measured effective parameters in the calculations instead of the bare ones. Technically, several further infinities are thus introduced which cancel out the existing ones. This concept is named *renormalisation* since the quantifiable parameters are scaled i.e. renormalised with respect to their effective values. Theories at which this method is applicable are called *renormalisable*, which is true for all theories involved in the SM.

In the following, these SM theories are described by their Lagrangians. The QCD is based on a $SU(3)$ group, characterised by

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^A_{\mu\nu} G^{\mu\nu}_A + \bar{q}_a (i\gamma^\mu D_\mu - m)_{ab} q_b \tag{2.4}$$

$$G^A_{\mu\nu} = \partial_\mu G^A_\nu - \partial_\nu G^A_\mu - g_S f^{ABC} G^B_\mu G^C_\nu \tag{2.5}$$

$$(D_\mu)_{ab} = \partial_\mu \delta_{ab} + i g_S (T^A G^A_\mu)_{ab} \tag{2.6}$$

with $A, B, C = 1\text{-}8$ indicating eight mediating particles of the strong interaction, the *gluons*. The corresponding three colour charges are represented by $a, b = 1\text{-}3$ (red, blue, green), which behave analogue to the theory of additive colour mixing. The equation also contains the term $G^A_{\mu\nu} G^{\mu\nu}_A$ which includes gluon-gluon self-interactions, allowed due to the non-abelian character of the theory. Furthermore, the QCD comprises the effect of *asymptotic freedom*, which implies that coloured partons (gluons or quarks) act as quasi-free particles at high energies (i.e. at small distances) and are subject to strong interactions at low energies (i.e. large distances). Besides, neither free quarks or gluons nor coloured compositions of particles can exist, known as the principle of *confinement*. Thus, according to the colour concept, only two colourless compositions, named *hadrons*, are allowed: Colour and the identical anti-colour states (quark + anti-quark), labelled *meson*, or three quarks of different colours (red + blue + green = white), labelled *baryon*. If a quark is high energetic enough to emerge these colourless hadrons, it can thus not act freely but *hadronises* again, by merging with other quarks from vacuum fluctuations or from gluon radiations (similar to Bremsstrahlung) and build up new colourless compositions. The outcome is a multiplicity of new particles, forming showers, called *jets.* The development of showers ends if the energy of the associated particles reaches the invariant mass of the hadrons, i.e. roughly 1 GeV. This process of *hadronisation* is only describable by approximative models.

The $U(1)_{QED}$ based theory of QED describes all effects of electromagnetism. However, in particle physics an irrefutable correlation to the weak interaction is present. Even though both forces have totally different force strengths, at high energies (100 GeV and above) they can be unified into a single theory. This unification is described by the group $SU(2)_L \otimes U(1)_Y$ with the Lagrangian

$$\begin{aligned} \mathcal{L}_{EW} = & -\frac{1}{4} W^i_{\mu\nu} W^{\mu\nu}_i - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ & + \bar{\psi}_L i \gamma^\mu D^L_\mu \psi_L \\ & + \bar{\psi}_R i \gamma^\mu D^R_\mu \psi_R \end{aligned} \tag{2.7}$$

$$D^L_\mu = \partial_\mu + \frac{1}{2} i g' B_\mu Y + \frac{1}{2} i g_w \tau_i W^i_\mu \tag{2.8}$$

$$D^R_\mu = \partial_\mu + \frac{1}{2} i g' B_\mu Y \tag{2.9}$$

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_w \epsilon^{ijk} W^j_\mu W^k_\nu \tag{2.10}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.11}$$

The *left-handed* (L) and *right-handed* (R) components are defined by the helicity h, with $h = -1$ and $h = +1$, respectively, where $h = \frac{\boldsymbol{s} \cdot \boldsymbol{p}}{|\boldsymbol{s}| \cdot |\boldsymbol{p}|}$ (spin $\boldsymbol{s}$, momentum $\boldsymbol{p}$). Since the weak force interacts exclusively with left-handed fermionic fields ψ_L, the covariant derivatives D_μ exist separately for L and R. The field strength tensor W^i represents the isospin triplet W^1, W^2, W^3, indicating three W bosons from the $SU(2)_L$ group, whereas the singlet B_μ of $U(1)_Y$ indicates a B boson. All four gauge bosons are not directly detectable. However, a combination of spontaneous symmetry breaking with the Higgs mechanism (described below) yields the observable bosons and fields

$$W^\pm_\mu = \tfrac{1}{\sqrt{2}} (W^1_\mu \mp W'^2_\mu) \tag{2.12}$$

$$Z_\mu = \quad B_\mu \cos\theta_w + W^3_\mu \sin\theta_w \tag{2.13}$$

$$A_\mu = -B_\mu \sin\theta_w + W^3_\mu \cos\theta_w \tag{2.14}$$

This describes all weakly interacting gauge bosons $W^\pm$ and Z, and furthermore A_μ, which is the photon field for γ. The here introduced electroweak mixing angle θ_w is a free parameter which can only be determined from experimental measurements.

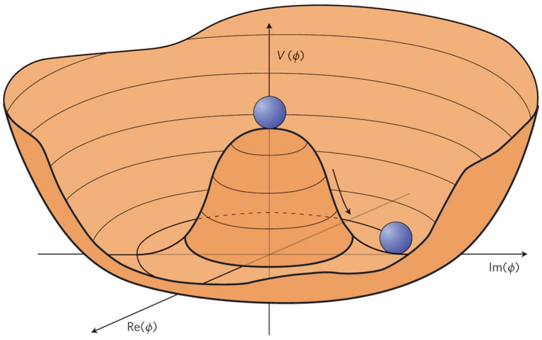


Figure 2.1.: Higgs potential $V(\phi) = \mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4$. Blue spheres indicate the position at the point of spontaneous symmetric breaking and in the ground state ϕ_0. Figure taken from [5].

Even though the main concepts of the SM are now covered, the above Lagrangian would only comprise massless particles. To give those particles mass, the Higgs mechanism is introduced. It states a complex scalar doublet field, the *Higgs field* ϕ, with a potential $V(\phi)$ similar to the shape of a Mexican hat (Figure 2.1). At high energies, the potential appears symmetric, but at a fixed point a spontaneous symmetry breaking occurs, resulting in a non-zero vacuum expectation value for the minimum. This mechanism gives mass to the electroweak gauge bosons Z and $W^\pm$, as well as to fermions. The corresponding Lagrangian is:

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) - g_f\left(\bar{\psi}_R\phi^\dagger\psi_L + \bar{\psi}_L\phi\psi_R\right) \tag{2.15}$$

$$V(\phi) = \mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4 \tag{2.16}$$

$$D_\mu = \partial_\mu + \frac{1}{2}ig'B_\mu Y + \frac{1}{2}ig_w\tau_i W^i_\mu \tag{2.17}$$

with $\mu^2 < 0$, $\lambda > 0$. One detail of the Higgs mechanism is the implication of a new massive boson, the scalar Higgs Boson. In 2012, both the ATLAS and CMS collaborations, discovered a boson with Higgs-like properties, having a mass of m $\simeq$ 125 GeV [1]. This last detected fundamental particle further substantiates the power of the SM.

Adding up the Lagrangians of equations 2.4, 2.7 and 2.15 yields

$$\mathcal{L} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs}, \tag{2.18}$$

representing the concept of the SM. It accounts for 12 bosons mediating the forces (γ, Z, $W^\pm$, 8 gluons) as well as 12 fermions separated in 3 families (up-type quark, down-type quark, charged lepton, neutrino), which are all discussed in the following.[1]

---

[1] For fermions also the corresponding anti-particles, with opposite charge, are described by the SM.

2.2. Elementary particles

The two groups of elementary particles are classified with respect to their spin: Particles with integer spin are called *bosons*, those with half-integer spin *fermions*, described by the Bose-Einstein statistics and the Fermi-Dirac statistics, respectively.

Bosons

Bosons[2], the force carrier particles, mediate the fundamental interactions. They are summarised in Table 2.1.

Table 2.1.: Overview of the three interactions described by the SM, including the mediating bosons and the corresponding charges.

Interaction	Mediating particles	Coupling to
Strong	8 gluons (g)	Colour charge
Weak	$W^{\pm}$, Z^0 bosons	Weak charge
Electromagnetic	Photon (γ)	Electric charge

Gluons are the massless mediators of the QCD, arising from the $SU(3)$ group theory, which generates an octet state. This state describes eight individual gluons carrying a combination of colour and anti-colour, coupling to all coloured particles, with the energy dependent coupling strength $\alpha_s(Q^2)$, implying also self-interacting processes. Even though, the strength of the coupling increases with larger distances, the maximum range is only about 1 fm due to the above described principle of confinement [3].
The particles mediating the weak interaction are the heavy gauge bosons $W^{\pm}$ and Z. Due to their relative high masses of 80.4 GeV and 91.2 GeV, respectively, they only have a short lifetime of $\mathcal{O}\left(10^{-25}\text{ s}\right)$ and a range similar to gluons. Furthermore, they are also self-interactive [1].
The last known boson of the fundamental interactions is the photon γ, which mediates the electromagnetic force by coupling to particles carrying electric charge. Since the γ is massless and no confinement applies, its lifetime is infinite and the propagation velocity equals the maximum velocity, the speed of light c.

---

[2] The Higgs boson is not considered here.

Fermions

The SM describes three families of fundamental fermions. Those of the 1st family build up the surrounding matter as we know it today, whereas the particles of the two further families are unstable and decay to first family fermions[3]. A summary of all fermions is given in Table 2.2.

Table 2.2.: Overview of all fundamental fermions in the SM and their corresponding charges. In addition, for each particle an associated anti-particle exists with opposite charge.

	Family 1st 2nd 3rd	Colour charge	Weak charge	Electric charge
Leptons	ν_e ν_μ ν_τ	-	✓	0
	e^- μ^- τ^-	-	✓	-1
Quarks	u c t	r/b/g	✓	$+2/3$
	d s b	r/b/g	✓	$-1/3$

Fermions can be further divided in two different categories of particles, leptons and quarks.

Leptons do not interact with the strong force of the QCD theory and are thus allowed to exist as free particles. They interact only with the weak force and - for electrical charged leptons (e, μ, τ) - also with the electromagnetic force. The electrically neutral neutrinos ν are thus hard to detect due to their weak interaction. Their masses are supposed to be very small since only upper mass limits have been set yet [1]. Charged leptons, however, have much higher masses than their partners. The values are 511 keV, 106 MeV and 1.78 GeV for e, μ and τ, respectively, and the electric charge allows an easy detection.

The second category of fermions, quarks, perform interactions with all three forces, since they hold all different types of charges known. This includes the colour charge of the $SU(3)_C$ group, which implies that no free quarks are allowed to exist, but only grouped in hadrons. The masses cover a wide range of five orders of magnitude, as summarised in Table 2.3.

It was first assumed that only two families of fermions exist, but the observation of violated CP-symmetry led to the introduction of the CKM-matrix, predicting three particle families (details in section 2.3). This concept was supported by the discovery of the b-quark [7] and finally validated by the evidence of the top quark several years later [8].

---

[3] For neutrinos oscillations between flavour eigenstates have been observed.

Table 2.3.: Masses of all quarks. Particles of higher families have higher masses. [1, 6]

u-quark:	$2.3^{+0.7}_{-0.5}$ MeV	c-quark:	1.28 ± 3 GeV	t-quark:	173.34 ± 0.76 GeV
d-quark:	$4.8^{+0.5}_{-0.3}$ MeV	s-quark:	95 ± 5 MeV	b-quark:	4.18 ± 0.03 GeV

Top quark

Of special interest in this thesis is the top quark. With $m_{top} = 173.34 \pm 0.76$ GeV [6], it is the heaviest elementary particle known today, roughly equivalent to the mass of a gold atom. Due to that high mass, the properties are quite different from lighter quarks, which e.g. induces a large decay width $\Gamma_{top} = 1.35$ GeV resulting in a very short lifetime of $1/\Gamma_{top} \simeq 5 \cdot 10^{-25}$ s [1]. This value is below the time of hadronisation, indicating that the top quark is the only quark which does not form quarkonium-bound states ($t\bar{t}$-states) before its decay.[4]

Due to this in general missing process, the top's quantum numbers are directly passed on to its decay products which allows a detailed analysis of the properties, like the top spin by examining the angular momentum of the daughter particles. Furthermore, the investigation of the intensive coupling to the Higgs field, induced by the high m_{top}, is important since physics beyond the SM may show strong influences on that. Therefore, precision measurements are a valuable means for the detection of new physics apart from direct searches. In addition, top quarks contribute as a major background to those direct searches and thus its understanding is of utmost interest.

Top quarks can either occur as *single tops*, as a result of weak interactions (e.g. $b \to t + W$), or more often as *top pairs* $t\bar{t}$ (e.g. $g \to t\bar{t}$). The latter combination decays solely via $t\bar{t} \to qqWW$. The products of these two W bosons, either $l\nu$ (leptonic) or $q\bar{q}$ (hadronic), classify the $t\bar{t}$ final states:

- **All hadronic:** Both W bosons decay into $q\bar{q}$ and hadronise subsequently, with an outcome of exclusively jets (6, with possibly further gluon radiation). Since quarks carry three different colour charges, this yields three different decay channels, which makes the process $W \to q\bar{q}$ the dominant one, with 46 % of all $t\bar{t}$ decays.

- **Di-leptonic:** Both W bosons decay into $l\nu$. The consequent final state is composed of 2 charged leptons, 2 neutrinos and (at least) 2 jets. Even though this outcome shows the cleanest signal, only 7 % of the $t\bar{t}$ events end in this state, when ignoring the difficult reconstructable τ leptons.

- **Lepton + Jets:** This channel is a mixture of the two above. The final constituents are (at least) 4 jets, one charged lepton and one neutrino. Ignoring once

---

[4] Nevertheless, weakly binding mechanisms are still active, since a small peak is present in the invariant mass distribution of $t\bar{t}$ [9].

more the τ lepton, 29 % of all $t\bar{t}$ events decay in this channel. With the charged lepton and the missing energy from the neutrino (cf. section 6.3), it is relatively straightforward to reconstruct one W boson, which supports the extraction of a clear signal. Together with the relatively high fraction of $t\bar{t}$ events, it leads to the phrase “golden channel” in top reconstruction.

The neglected τ leptons are not completely lost in the default reconstructions, since about 1/3 are further decaying into either $e + \nu_e$ or $\mu + \nu_\mu$. This allows them to pass the charged lepton reconstruction requirements and yields additional contributions to the respective channels.
The quarks in the $t\bar{t}$ decay not arising from W bosons are dominated by b-quarks, since according to the CKM matrix 99.8 % of all tops decay as $t \rightarrow b + W$.

2.3. The CKM matrix

Based on the work of Nicola Cabibbo from 1963 [10], a 2×2 matrix was formulated [11] to describe the relations between up, down and strange quarks while also predicting the charm quark, which was not yet discovered back then. Furthermore, in 1964, a violation of CP-symmetry was observed in kaon decays [12], which was found to be crucial for the understanding of nature since it can explain - to some extent - the imbalance between matter and anti-matter in the universe. For a theoretical explanation, an extension of the Cabibbo matrix to three particle families was proposed by Makoto Kobayashi and Toshihide Maskawa in 1973 [13]. The resulting Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} has a unitary 3×3 form, which is determined by three real mixing angles θ_{gh} (Euler angles; $g, h = 1, 2, 3$) and one complex phase $i\delta$:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2.19}$$

with $s_{gh} = \sin\theta_{gh}$ and $c_{gh} = \cos\theta_{gh}$. The complex phase $i\delta$ triggers the observed CP violation, but occurs only for at least three families. It yields the unequality $V_{ij} \neq V_{ij}^*$, which causes influences on the CP-symmetry in weak interaction as stated in simplified terms in the following.

CP violation

Assuming a process $i \rightarrow j$, with initial state i and final state j, which is described by the CKM matrix element V_{ij}. A CP-transfer is now applied, which changes simultaneously matter to anti-matter (C: charge flip) as well as the helicity sign (P: parity flip, i.e. point reflection). This results in a transfer from e.g. left-handed quark i to right-handed anti-quark $\tilde{i}$ (same for j). If the rate of $i \rightarrow j$ is the same as $\tilde{i} \rightarrow \tilde{j}$, the processes is CP-invariant. However, the rates of the processes are determined by the complex

transition amplitude $\mathcal{M}$ which is identical for both processes, except for processes that are subjected to interference in conjunction with the matrix element V_{ij} including $i\delta$.[5] Therefore, if different rates for $i \rightarrow j$ and $\tilde{i} \rightarrow \tilde{j}$ are obtained, this stands for a violation of the CP transformation, due to the presence of $i\delta$. [3]
Nevertheless, physical consequences only happen in the case that all three particle families are involved in the process, which is rare (e.g. in loop corrections). Thus, occurrences of CP violation are extremely seldom. [2]

Eigenstate transformation

A more experimental based parametrisation of the CKM matrix is:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{weak} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass} \tag{2.20}$$

Here, the second purpose of the matrix gets clear: It mediates a transformation from quark mass eigenstates to weak eigenstates, where the latter describe quark interactions of the weak force. Observations of quarks are instead always described with respect to the mass eigenstates, also known as the strong eigenstates

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{2.21}$$

However, the weak force couples to

$$\begin{pmatrix} u \\ d' \end{pmatrix} \begin{pmatrix} c \\ s' \end{pmatrix} \begin{pmatrix} t \\ b' \end{pmatrix} \tag{2.22}$$

at which d', s', b' are linear combinations of the observed down-type quarks d, s, b, described by the matrix V_{CKM}. For example, the weak eigenstate d' is written as a combination of the mass eigenstates

$$d' = V_{ud} \cdot d + V_{us} \cdot s + V_{ub} \cdot b. \tag{2.23}$$

Conventionally, up-type quarks stay unchanged in this description. An equivalent formulation also works for up-type quarks, leaving out the down-type quarks. [3]

Since the properties of the weak quark interaction are thus directly contained in the matrix, the weak-mediated transition probabilities of one quark to another can be directly extracted by evaluating the squared value $|V_{ij}|^2$, describing the decay $i \rightarrow j + W^{\pm}$. As indicated by the nonzero off-diagonal entries in the CKM matrix, this also admits transitions between different flavour families.

---

[5] The two transition amplitudes $\mathcal{M}$ are only different, if multiple decay routes ($i \rightarrow x \rightarrow j,\quad i \rightarrow y \rightarrow j$, etc.) exist, resulting in different rates.

Properties and measurements

Apart from the above discussed aspects, one additional crucial statement of the theory is the unitarity of V_{CKM}, demanding

$$\sum_{i=1}^{3} |V_{ij}|^2 = \sum_{j=1}^{3} |V_{ij}|^2 = 1, \tag{2.24}$$

$$\sum_{k=1}^{3} V_{ik} V_{jk}^* = 0, \tag{2.25}$$

which undergoes intensive tests by several experiments. A proof of erroneousness could e.g. arise if a quark flavour change would be observed, which is not mediated by the electroweak interaction and within the three stated quark families.

However, up to now, no deviations have been found. All tests of the CKM matrix are in principle based on the determination of the magnitudes of all 9 CKM elements, albeit 4 parameters $(\theta_{gh}, i\delta)$ are theoretically sufficient to describe the matrix, which allows further cross-checks. The determination of elements related to u-quarks and c-quarks is achieved directly through processes listed in Table 2.4, while the element V_{tb} is examined from top quark decays with b-quark reconstruction or from single top production processes.

Table 2.4.: Dominant experimental processes for the determination of the 6 matrix elements without involvement of the top quark. [1]

$\lvert V_{ud}\rvert$:	Nuclear β decays	$\lvert V_{us}\rvert$:	Kaon decays	$\lvert V_{ub}\rvert$:	B decays
$\lvert V_{cd}\rvert$:	D decays $+\,\nu$ scatter	$\lvert V_{cs}\rvert$:	D decays	$\lvert V_{cb}\rvert$:	B decays

Thus, seven elements are directly measurable. However, the two CKM matrix elements V_{ts} and V_{td} are only - up to now - determined indirectly, e.g. by studying meson oscillations like B_s^0 mixing, as depicted in Figure 2.2. In these oscillations, the vertices in the box diagrams are described by the matrix elements V_{ts} and V_{tb}, when assuming a top quark as the mediating particle. Thus, $|V_{ts}^* V_{tb}|$ is strongly connected to the oscillation process and its frequency[6], which is proportional to the mass difference Δm_s of the B_s^0 mass eigenstates B_H and B_L, allowing a determination of the matrix element combination $|V_{ts}^* V_{tb}|$ [1]. However, this technique of V_{ts} determination relies on independent measurements of the matrix element V_{tb} from other experiments.

Meson oscillations have also been observed for $K^0 - \bar{K}^0$, $B^0 - \bar{B}^0$ and $D^0 - \bar{D}^0$ systems, which in principle all allow indirect determinations of CKM matrix elements [14]. For example, V_{td} is mainly obtained by $|V_{td}^* V_{tb}|$ from $B^0 - \bar{B}^0$ mixing.

---

[6] Despite the fact that the B_s^0 meson oscillates on average nine times until it decays, an extraction of the frequency is still possible [14].

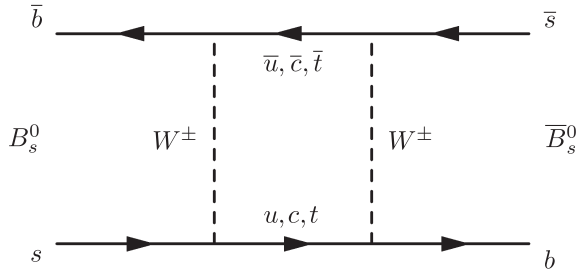


Figure 2.2.: Feynman diagram of B_s^0 mixing. For a mediating top quark, the vertices are determined by the CKM matrix elements V_{ts} and V_{tb}. Figure adopted from [15].

Apart from the B meson oscillations also rare loop-mediated decays are used for the determination of the V_{ts} and V_{td} matrix elements. Two instances are depicted in Figure 2.3, at which the vertices are determined by V_{ts} and V_{tb} if a top quark is the mediating particle.

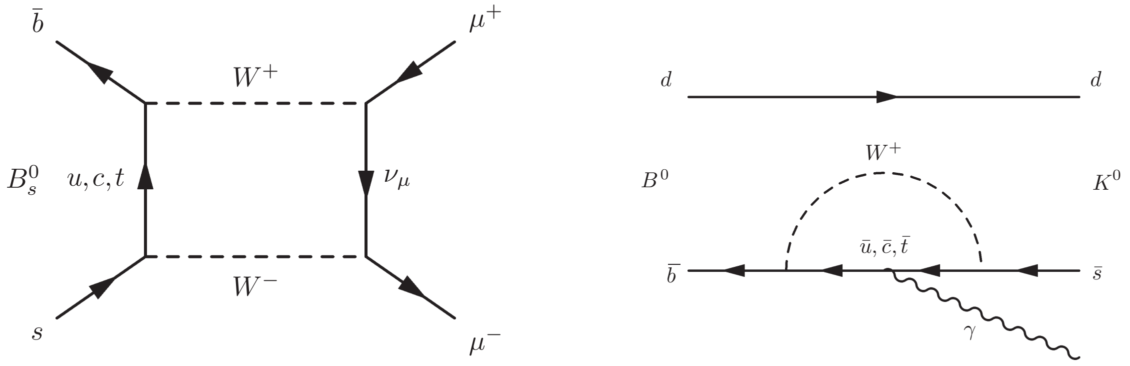


Figure 2.3.: Feynman diagrams for indirect $|V_{ts}|$ extraction. For a mediating top quark, the vertices are determined by the CKM matrix elements V_{ts} and V_{tb}. Figures adopted from [15].

To conclude, a combination of theoretical predictions and available measurements led to the magnitudes of the CKM matrix elements [1].

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22522 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886^{+0.00033}_{-0.00032} & 0.0405^{+0.0011}_{-0.0012} & 0.99914 \pm 0.00005 \end{pmatrix}$$

Since the CKM matrix is determined by only four parameters, but can be tested by several independent measurements (at least 9 due to the individual elements), contributions from physics beyond the SM should clearly be seen through discrepancies between theory and experiments. Several examples of new physics models, which might influence the CKM matrix, are described in [1].

But for a full validity check of the CKM matrix, indirect measurements are not perfect since there are always dependencies on other matrix elements. The first experiment at which the, up to now only indirectly measured, V_{ts} element, can be obtained directly is the LHC. The number of top quarks produced is much higher than in former experiments and so the decay $t \rightarrow s + W$ might be observable which is proportional to $|V_{ts}|^2$.

3. CERN accelerators and detectors

CERN is a research institute located in the border area between France and Switzerland. The origin of the acronym arises from the provisional foundation board "**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire". The major aim of the collaborating physicists and engineers from 21 member states is to gain a better understanding of the structure of nature. To achieve this goal, several collider experiments are operated with particles interacting at a velocity close to the speed of light. About 25 experiments and six accelerators are currently active on site.

The institute was founded in 1954, with the purpose of building European expertise in nuclear physics. In the upcoming years, scientists from CERN made important observations like antinuclei, inner structure of protons, CP violation, discovery of the Z and W bosons as well as the discovery of the Higgs boson. In addition, there were huge steps forward in engineering and computer science, like detector development, superconducting magnet construction, GRID computing and the invention of the World Wide Web. [16]

Today's research activities are dominated by the two general purpose experiments ATLAS and CMS, which are both part of the accelerator complex of the LHC.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) forms a ring with a circumference of about 27 km at a mean depth of 100 meter below surface. The LHC tunnel was originally constructed for the former experiment, the Large Electron Positron collider (LEP) which was operated from 1989 until 2000.

The first plans to build the LHC were announced at the Lausanne workshop in 1984, four years before the LEP tunnel was completed. At that time, the new proton-proton-collider was proposed to have a luminosity of $\mathcal{L} = 10^{33}$ cm^{-2} s^{-1} and a centre-of-mass energy of $\sqrt{s} = 18$ TeV, but already three years later, at the La Thuille workshop this was decreased to $\sqrt{s} = 16$ TeV. At the first "Conceptual Design Report" [17] in 1995 the values were then fixed for construction: $\sqrt{s} = 14$ TeV, $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1} for pp collisions. In addition, also runs with heavy ions at $\sqrt{s} = 1000$ TeV with a luminosity of $\mathcal{L}_{ion} = 10^{27}$ cm^{-2} s^{-1} were announced. In 1998, the construction started. [18]

It was planned to achieve the first collisions in 2008 at lower energies and ramp up to the final design energy of $\sqrt{s} = 14$ TeV in 2009. However, shortly after the first protons circulated the LHC a major technical incident occurred which delayed the start of the physics program by more than one year. The first pp collisions were recorded in March 2010. The LHC fully operated then in 2011 at $\sqrt{s} = 7$ TeV and increased to $\sqrt{s} = 8$ TeV in 2012. Afterwards, upgrades of detector and collider parts for the subsequent Run II were performed from 2013 until 2015. [18]

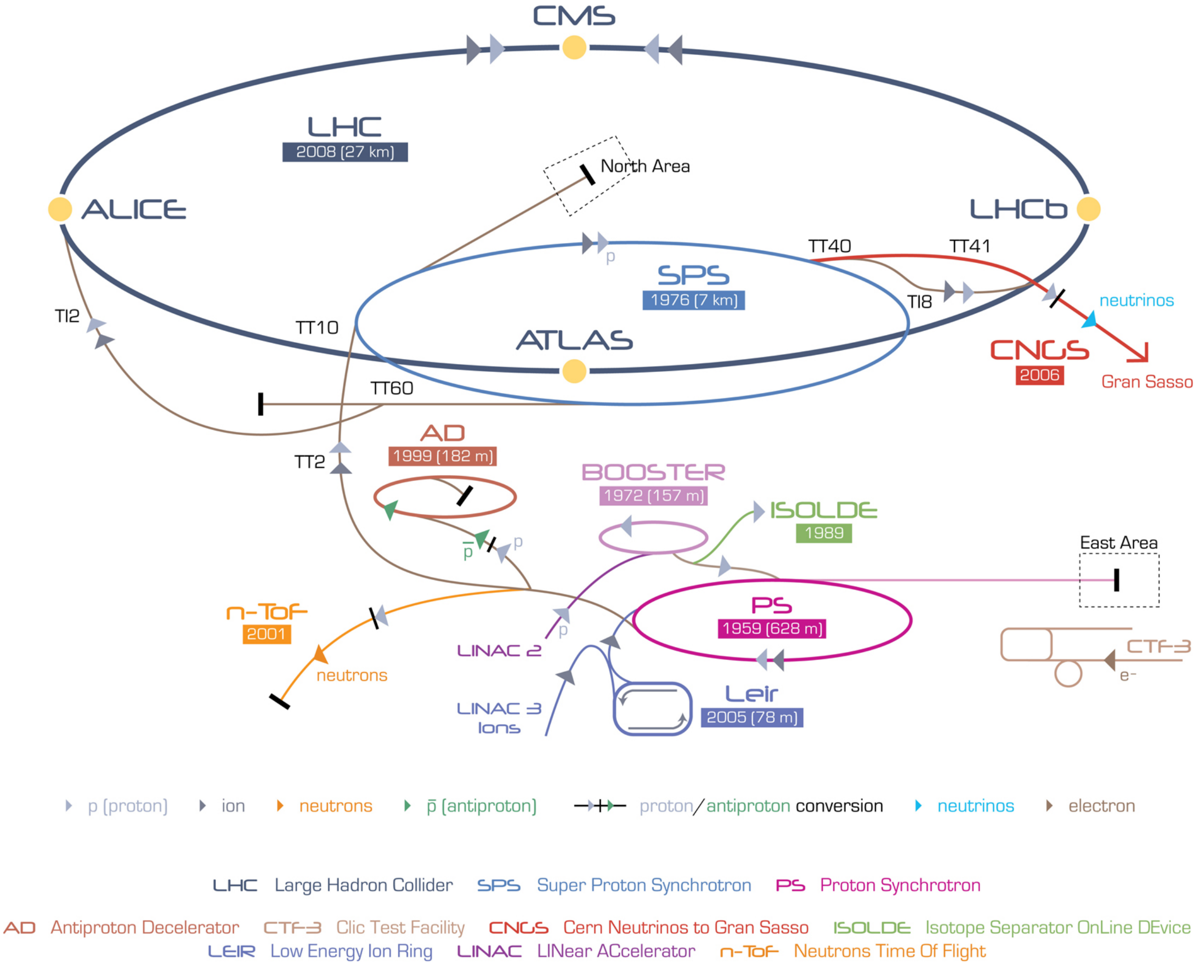


Figure 3.1.: CERN accelerator complex. Figure taken from [16].

The origin of the colliding protons are hydrogen molecules, which are separated in atoms and subsequently ionised. Starting from a 90 keV potential, the protons are successively accelerated in the CERN accelerator complex as depicted in Figure 3.1. This includes the Linac2 (50 MeV), the PS Booster (1.4 GeV), the Proton Synchrotron (25 GeV) and finally the Super Proton Synchrotron (450 GeV). With the transfer injection lines TI2 and TI8, particles are then injected clockwise and anti-clockwise into the LHC, where they are further accelerated to their final collision energy of 3.5 - 4 TeV in Run I (2011, 2012) and 6.5 - 7 TeV in the future Run II. The protons are thus circulating in separate bunches, each containing about 10^{11} protons, with a distance of 25 - 50 ns. [19]

The LHC is not a perfect circle but separated into eight different straight sections and intermediate arcs. The 8 arcs each contain 154 dipole magnets, the 8 straight sections incorporate accelerating cavities, beam dump mechanism, two beam cleaning areas and four collision points where the main experiments are located. [19]
These four experiments consist of two special-purpose detectors and two multi-purpose detectors:

- ALICE[1] is designed for an optimised reconstruction of heavy ion collisions. In these reactions, a quark-gluon-plasma is created in which physics processes still comprise pending questions.
- LHCb[2] has the purpose to study b-quark hadrons at the LHC and thus analyse e.g. the asymmetry between matter and anti-matter in detail.
- CMS[3] is one of the two multi-purpose detectors, with a general approach for the detection of new physics. The name represents nicely its design: It is relatively compact and small (compared to ATLAS), has a strong capability to detect muons and finally incorporates a strong solenoid magnet.
- ATLAS[4] is the second multi-purpose detector and is discussed in detail in section 3.2. It is in principle the counterpart of CMS (and vice versa), since an observation of a new physics aspect always requires the confirmation by a second experiment, to be accepted in the scientific community.

In addition to these four detectors, there also exists three smaller experiments which do not encircle a collision point. Their measurements are based on forward particles from nearby interaction points.

3.2. The ATLAS detector

ATLAS was designed as a multi-purpose detector, which allows the analysis of *pp* collisions as well as heavy-ion collisions, with the capability of finding new physics in different phase space regions. The major features of the detector are:

- Spatial detection: All detector components, but especially the tracking system, trace the flight paths of particles, which allows the reconstruction of particle decay points as well as the identification of the collision point.
- Momentum measurement: The tracking system and the muon system, combined with the magnetic field, enable the reconstruction of curved tracks from charged particles, which can be interpreted as the particles' momenta.

---

[1] **A** **L**arge **I**on **C**ollider **E**xperiment

[2] **LHC** **b**eauty

[3] **C**ompact **M**uon **S**olenoid

[4] **A** **T**oroidal **L**HC **A**pparatu**S**

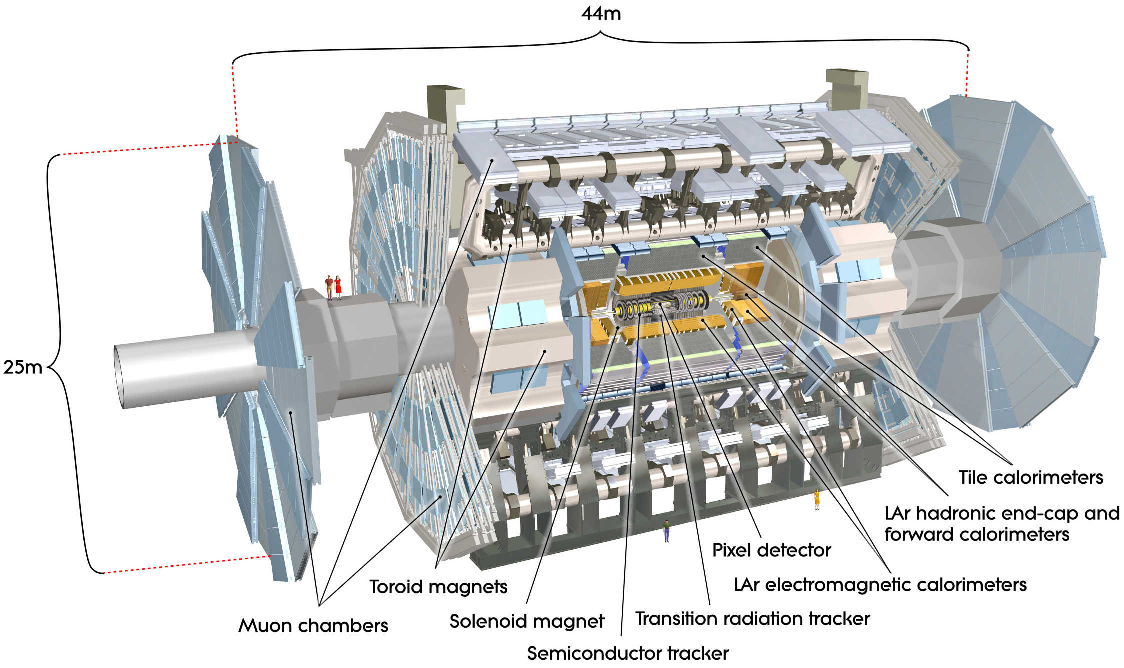


Figure 3.2.: The ATLAS detector and its constituents. Figure taken from [20].

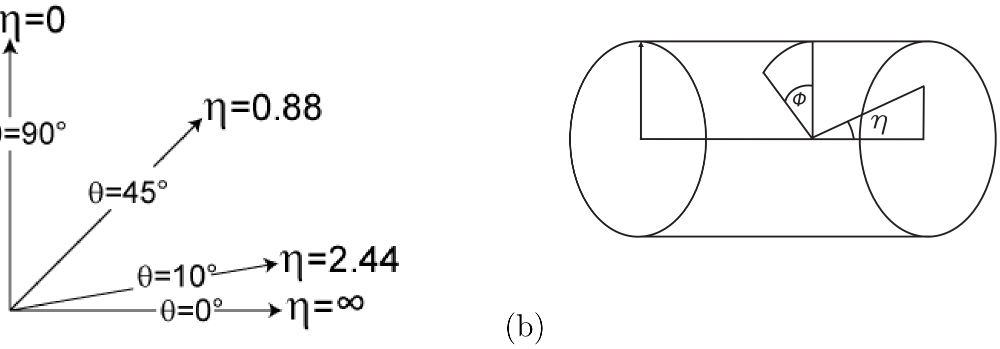
- Energy determination: The energy of most particles is absorbed and evaluated by the calorimeters.

- 4π detection: The detector encircles as a cylinder the collision point as a whole. Thus, no SM particle should leave the detector without a trace, except for neutrinos.

With the cylindrical shape of the detector, it is advantageous to introduce a detector dependent coordinate system. The ATLAS coordinate system is based on a right-handed description, with the origin (0,0,0) at the interaction point. The x-axis is defined as the direction towards the centre of the LHC, while the y-axis points upwards. Consequently, the z-axis is situated parallel to the beam axis. Due to the barrel like structure, a cylindrical coordinate system is used, as depicted in Figure 3.3, where the polar angle θ is replaced by the *pseudorapidity* η:

$$\eta = -\ln[\tan(\frac{\theta}{2})]. \tag{3.1}$$

Due to the high energy of particles, energy and momentum are assumed as being equal. This induces that the spatial pseudorapidity η is approximately identical to the velocity-related rapidity y, which is often used in high-energy physics.



(a) (b)

Figure 3.3.: (a) Relation between the cylindrical angle θ and the pseudorapidity η. (b) Cylindrical coordinate system of ATLAS with the pseudorapidity η and the azimuthal angle ϕ. The beam line runs longitudinally, i.e. here horizontally. Figures taken from [21].

ATLAS consists of four major sub-detectors, which enclose each other in layers (onion-like structure). From the collision point outwards, these are: Inner detector, electromagnetic calorimeter, hadronic calorimeter and muon spectrometer, where all components are permeated by magnetic fields. The following description of these detector systems are mainly based on [22], where more technical details are listed.

3.2.1. Magnetic system

In particle detectors, the momenta of charged particles are extracted from their curvatures in magnetic fields. Since the particles at ATLAS are high energetic, strong magnetic fields are required to still obtain reasonable bending. Therefore, all niobium-titanium magnets are cooled with liquid helium to reach a superconductive state.
The region of the inner detector is thus filled with a homogeneous 2 T magnetic field, generated by a solenoid. The typical bending power for a charged particle is 2.1 Tm, which allows a precise momentum and charge reconstruction of all charged particles in question. [23]
For the outer detector region, i.e. for the muon spectrometer, eight barrel toroid magnets and two end-cap toroid magnets are installed. Typically, values up to 3 Tm in the barrel region and up to 6 Tm in the end-cap region are achieved, which is essential for detecting the high energetic muons. [24, 25]

3.2.2. Inner detector

The first element of the detector, starting from the central collision point, is the inner detector. With an outer diameter of 2.3 m and a total length of 7 m, it covers a region of $|\eta| < 2.5$. Three components with different technologies compose this sub-detector (Figure 3.4):

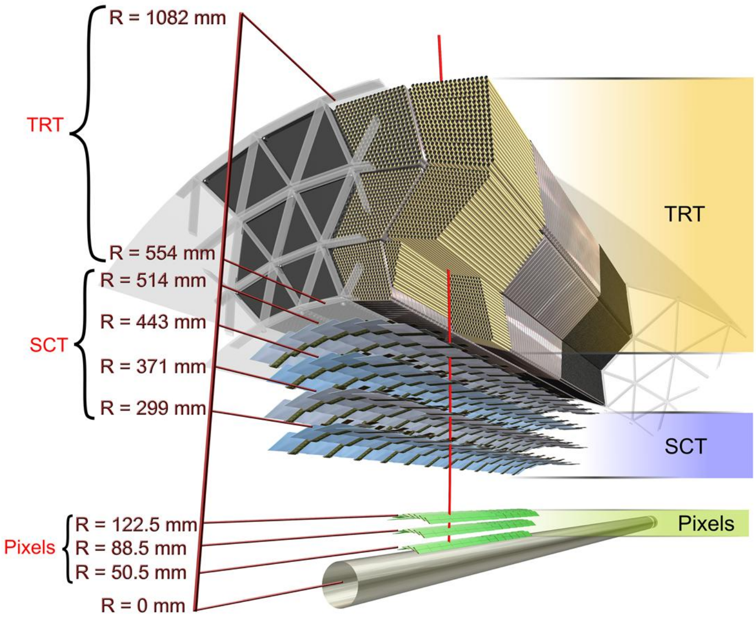


Figure 3.4.: The ATLAS inner detector consisting of pixel detector, semiconductor tracker (SCT) and transition radiation tracker (TRT). Figure taken from [20].

- Pixel detector:
  With a distance of only 5.05 cm from the collision point, it is the most central detector component. The particle reconstruction is performed by means of pixels with a size of 50 μm × 400 μm and a resolution of 12 μm. Both, in the barrel and end-cap region, they are organised in three separate layers.

- Semiconductor tracker (SCT):
  The second instrument consists of silicon strip layers with a size of 80 μm × 12 cm. In the barrel region, four layers are installed, whereas in the end-cap section nine disks are situated on each side.

- Transition radiation tracker (TRT):
  The third component contains about 300 000 straws with a size of 4 mm. They are made of Kapton and filled with a Xenon-based gas mixture. The transition radiation is sensitive to E/m which allows a differentiation of electrons and pions.

The combination of these three components results in a high resolution for track reconstruction as well as a very good vertex identification. These are crucial ingredients for the reconstruction of particle decay points and b-hadrons. [26, 27, 28]

3.2.3. Electromagnetic calorimeter

In the electromagnetic calorimeter, electrons and photons are subject to a complete loss of energy via pair production and bremsstrahlung. Their absorbed energy, which equals their initial energy, can thus be measured. The layout is based on alternating layers of massive sections (lead) for the energy absorption and of sensitive sections for the energy determination (liquid argon). The latter component also led to its common name "LAr calorimeter". As depicted in Figure 3.5 in light orange, the calorimeter consists of two regions, the *LAr electromagnetic barrel* for particles with $|\eta| < 1.475$ and the *LAr electromagnetic end-cap* with $1.375 < |\eta| < 3.2$. An energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ can be achieved. [29]

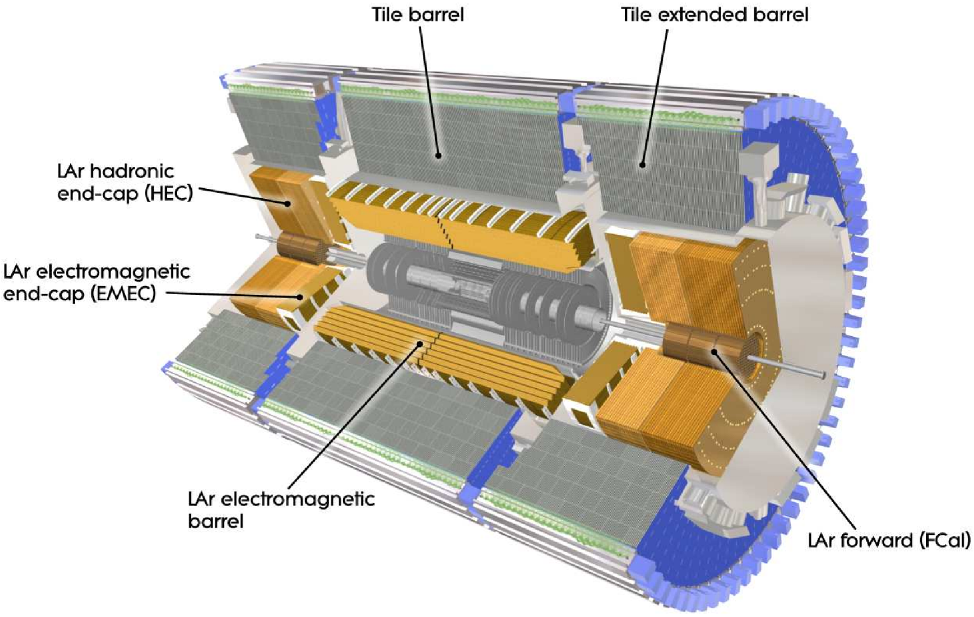


Figure 3.5.: The ATLAS calorimeters. Figure taken from [20].

3.2.4. Hadronic calorimeter

The hadronic calorimeter is similar to the electromagnetic calorimeter, in layout, function as well as in its purpose, which is the absorption of all hadrons and a concurrently measure of their energy. In the barrel region (grey/green in Figure 3.5), the massive sections are fabricated in iron, while the sensitive sections use scintillating plastic tiles, which led to the common name "Tile calorimeter". When particle showers are passing through, light is emitted which can easily be recorded. The coverage of the barrel tile calorimeter is $|\eta| < 1.7$.

For the hadronic end-caps a coverage of $1.5 < |\eta| < 3.2$ is achieved (*HEC*), with alternating sections of copper and liquid argon, similar to the electromagnetic calorimeter.

The specifications of the full hadronic calorimeter hold a length of 12 m, a diameter of 8.5 m and an energy resolution of $\sigma_E/E = 50\%/\sqrt{E} \oplus 0.7\%$. [29, 30]

Apart from the individual hadronic and electromagnetic calorimeter systems, a combined LAr calorimeter is installed in the forward region (*FCal*), covering all particles with $3.1 < |\eta| < 4.9$ (Figure 3.5). For particles with a much larger pseudorapidity η, there exist also several forward detectors outside the ATLAS cavern with special purposes. One example is the “Zero Degree Calorimeter” for particles with $|\eta| > 8.3$. [29, 31]

One major advantage of the calorimeter system used is the absorption of all SM particles, except muons and neutrinos. Combined with the muon spectrometer, this allows for a measurement of missing transverse energy $\not{E}_T$, which can be used for the reconstruction of neutrino energies.

3.2.5. Muon spectrometer

The final sub-detector is the muon spectrometer. It is designed to reconstruct muons which traverse the previous detector parts without being stopped. Since all other interacting particles have already been absorbed, the registered signals can be supposed to stem from muons.

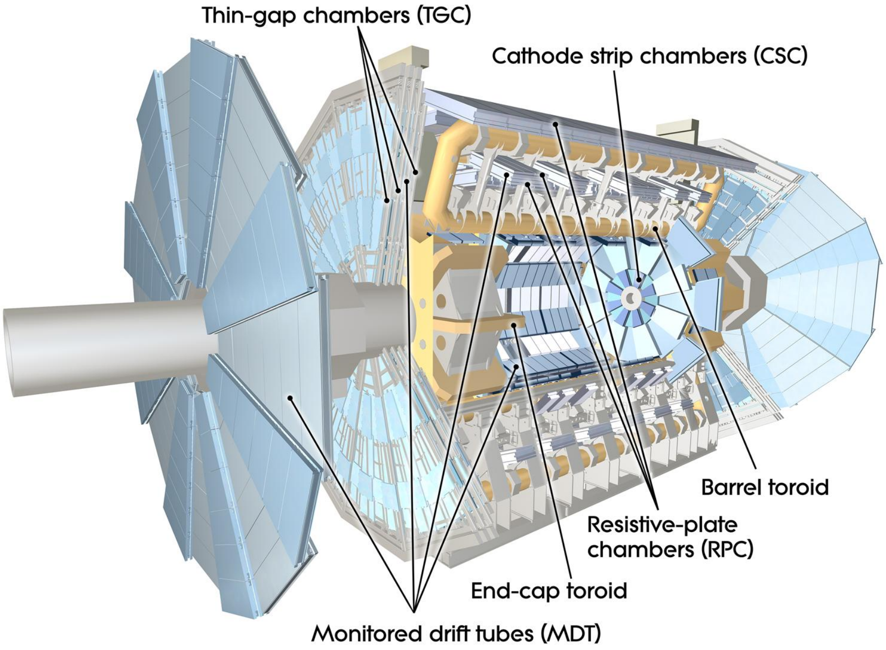


Figure 3.6.: The ATLAS muon spectrometer. Figure taken from [20].

The spectrometer is a composition of four different detector systems:

- Monitored drift tubes (MDT):
  The widely installed (Figure 3.6), 3 cm thick drift tubes are filled with Argon-CO_2 and comprise a centred wire at high voltage of 3 kV. Atoms, ionised by a crossing muon, drift to the outer tube, electrons to the central wire. Since a muon traverses a multiplicity of such tubes, its path can be reconstructed at high precision.
- Cathode strip chambers (CSC):
  The strip chambers are mainly located in the central area of the spectrometers' end-caps (Figure 3.6), since they show a good performance at high particle rates. A chamber consists of anode wires and cathode copper strips which are arranged in a crossed position. Similar to MDT, muons are causing an ionisation process which is then detected. Due to the cross-structure, a good extraction of the traversing position can be achieved, which leads to high precision in the muon reconstruction.
- Resistive-plate chambers (RPC):
  These chambers are basically identical to the above CSC. However, the wires and strips are replaced by two charged plates with a distance of 2 mm. This results in fast particle detection for the usage in the trigger system (see next section).
- Thin-gap chambers (TGC):
  These chambers can once more be compared to CSC, but instead of crossed electronics, the wires are here installed in parallel, with a distance of about 2 mm. Similar to RPC, the ionised particles are recorded fast, which again is used for the trigger system.

With 44 m in length and 22 m in diameter, the muon spectrometer is the largest subdetector in ATLAS. With its huge scale, measurements of curvatures from high energetic muons are possible. This allows a reconstruction of the particles' momenta, similar to inner detector tracks. The corresponding energy resolution is quite good, with e.g. 3 % for 100 GeV muons. [32]

3.2.6. Trigger system

With the small distance between circulating proton bunches of 25-50 ns, the crossing rate at the intersection point of ATLAS is quite high (up to 40 MHz). With a typical event size of 1.5 MB, this would lead to an unproccesable data volume of one petabyte per second. To deal with such big data, a reduction procedure with several hardware and software based selection mechanisms, called trigger, is executed (Figure 3.7).
First, the hardware-based trigger level L1/ LVL1 tags events based on inputs from the calorimeters and the muon spectrometer. The decision is taken within 2.5 μs, which leads to a reduction of the event rate by a factor of 500 from 40 MHz to 75 kHz. If accepted as of interest, the event is passed on to *Read Out Drivers* (ROD) for further analysis in the subsequent trigger levels. For those, *Region of Interests* (RoI) are marked,

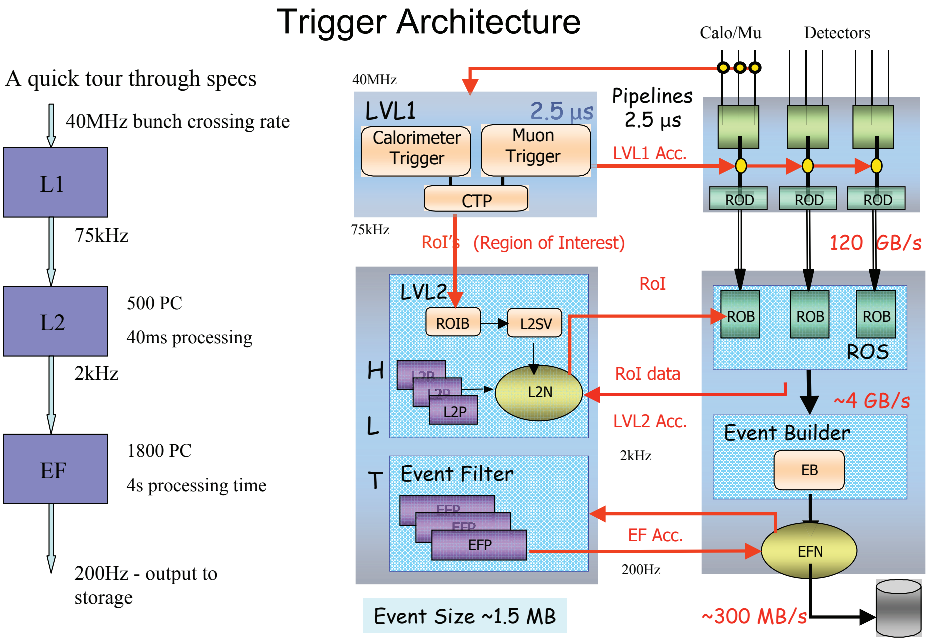


Figure 3.7.: Workflow of the ATLAS trigger system. Figure adopted from [33].

at which the important processes were found.
The second trigger level L2/ LVL2, which is software-based, then receives the events from the ROD. Its purpose is to examine solely the already marked RoI with higher precision and with additional information from other sub-detectors, which was not available for the L1 decision. This step reduces the event rate by a factor of 40, yielding 2 kHz.
Finally, the data are passed to the *Event Filter* (EF), at which a reconstruction of the event is performed. There about 90 % of the events are rejected, resulting in an event output of 200 Hz. This value equals a data production of 300 MB per second which is finally stored on disk. As depicted in Figure 3.7, the combination of the software-based levels L2 and EF is also known as the *High Level Trigger* (HLT).
The numbers stated here are based on the design values. During Run I, the data saving process performed better than predicted, which allowed lower trigger thresholds and more recorded events. [33]

3.3. Worldwide LHC Computing Grid

Even though the number of events is reduced dramatically by the trigger system, ATLAS as well as the other LHC experiments still deliver a huge amount of data which has to be analysed by scientists worldwide. Therefore, all experiments operate a delocalised, Worldwide LHC Computing Grid (WLCG) that stores experimental data as well as simulated data and provides computing power for consequent physics analysis, which can not be performed locally.

The infrastructure of the WLCG consists of three different layers (tiers). Tier 0 is the CERN main data centre, with an extension in Budapest (Hungary). It receives new LHC data directly from the experiments and consequently distributes them to 13 regional Tier 1 computing centres. From these, all data are provided to about 160 local Tier 2 centres in 40 countries where the actual analysis takes place.

Altogether, the WLCG represents a machine with ca. 350 000 cores and a total of 500 PB storage space. [34]

4. Datasets

The aim of this study is to measure the CKM matrix element $|V_{ts}|$ by examining the rate of events in which a top quark decays to a strange quark plus an associated W boson[1]: $t \rightarrow s + W$.

For the analyses of data events acquired with the ATLAS detector, Monte Carlo (MC) simulations are necessary for comparison. In the search of rare $t\bar{t}$-decays, the selection of a MC-sample, whose properties are well modelled, is crucial. Thus, the top quark specific MC generator `PROTOS` [35] is used for detailed studies of the $t \rightarrow s + W$ decay. Its behaviour and the chosen production options are summarised in section 4.1. In addition, the MC generator `MC@NLO` [36] (introduced in section 4.2) is used for further investigations and for the final analysis because it is able to describe higher order corrections. Apart from the signal, several background processes are considered as well (section 4.3). Data acquisition obligatory for the final data-to-simulation comparison is detailed in section 4.4.

4.1. Monte Carlo generator for enhanced top branching ratios

In particle physics, Monte Carlo simulations are required for a comparison of experimental data with approximated theoretical predictions. It is not possible to simply perform complete theoretical calculations of the events due to different reasons:

Problems of the quantum field theory are not generally solvable. To achieve nevertheless an adequate description, perturbative calculations are used which are an approximation of the solutions (see also chapter 2). This ansatz can only take lower order calculations of exclusively hard processes into account, and hence not all aspects of an event are fully considered. The missing higher orders, like next-to-leading order (NLO), usually influence the final results only within small correction ranges but can have a larger impact in some cases. Additional soft processes, like the Underlying Event [21], contribute as well, and thus are also not fully negligible.

Another reason why full calculations cannot to be used, are several missing constraints. For example, the “parton distribution function” (PDF) is only well known at high energies. For low energies, several studies have been performed but the uncertainties are

---

[1] Charge conjugation is always implied in this thesis and therefore the charge of particles might not be stated every time.

still limiting the results [37] for a fully theoretical calculation.
Due to these aspects, additional inputs have to be considered to accommodate the missing parts and to approximate the reality as close as possible. This can be achieved by combining different MC generators (e.g. for soft and hard processes) with additional models, describing the aspects of an event: The hard collision itself, pileup effects (section 5.5), hadronisation models (section 2.2), the detector geometry and the detector response (chapter 3). All of these packages still undergo continuous optimisation, performed by several theory groups worldwide. The quality is verified at ATLAS by different performance groups, where simulated descriptions are tested against the measured data. With this procedure one ends up with a well described and understood model.

4.1.1. Introduction to PROTOS

The measurement of $|V_{ts}|$ is a very specific analysis and so none of ATLAS officially tested MC generators is able to describe the decay $t \rightarrow s + W$ with enhanced branching ratios (BR). Therefore, one of the first crucial steps in this analysis was to find and test a MC generator to start with a good description of the data as well as a good consistency with the MC generators recommended by the ATLAS collaboration.
The final choice was the leading order (LO) MC generator `PROTOS` (**PRO**gram for **TO**p **S**imulations) [35]. Its development aimed for the description of several specific processes and dependencies related to the top quark sector, like anomalous Wtb couplings, flavour changing neutral currents, heavy vector-like quarks (e.g. $T\bar{T}$ and $B\bar{T}$) as well as different BR for top quarks decaying to d-, s-, b-quarks [35]. It is the only generator available which currently provides the option to set user-defined parameters for the branching fractions. The ATLAS officially validated MC generators `POWHEG` [38] and `ALPGEN` [39] do e.g. not support any option to set $V_{tb} \neq 1$, whereas for the also validated `MC@NLO` + `HERWIG` [36, 40] simulations, problems occur during the event generation when applying changes to the default CKM parameters.
Furthermore, `PROTOS` is also used by several analyses [41, 42] in the ATLAS collaboration and therefore has already been tested for the high energies and the physics environment of the LHC.

The `PROTOS` generator only simulates the proton-proton-collisions itself as well as the hard decays up to the beginning of hadronisation and showering. All other accompanying processes have to be simulated by additional MC generators as for example the widely used `PYTHIA` generator [43]. This generator generally takes the hard processes (i.e. daughter quark plus W boson as well as possible gluons) from the preceding generator as a starting point and builds a complete physics model for the event, while conserving energy and momentum. Also additional processes like initial-state-radiation (ISR) and final-state-radiation (FSR) as well as pileup are taken into account (detailed description in section 4.1.3).

For the 7 TeV simulation study which is presented in section 7.2, a `PROTOS` MC dataset was already available, having the BR $t \to b + W = 50$ %, $t \to s + W = 25$ % and $t \to d + W = 25$ % and a total of 150,000 events in the lepton+jets decay channel. Those parameters are a perfect starting point to understand the influence of the different branches on the event topology.

For the 8 TeV analysis and especially for the here presented $|V_{ts}|$ measurement, a dataset with higher statistic as well as new parameters like the higher centre-of-mass energy for 2012 data was needed. The original plan was to produce three independent samples, the first with $V_{ts} = 1$, the second with $V_{td} = 1$ and the last with $V_{tb} = 1$. However, it was refrained from this approach because in these combinations, no event would be available with top and anti-top decaying to different flavours, like e.g. $t \to bW^+, \bar{t} \to \bar{s}W^-$ or $t \to dW^+, \bar{t} \to \bar{b}W^-$, and so events which are most relevant for the event selection would not have been present.
Therefore only one sample was produced including a total of 1,000,000 events with enhanced BR, i.e. 50 % of $t \to b+W$, 25 % of $t \to s+W$ and 25 % of $t \to d+W$ decays. Similar to the 7 TeV dataset, all top pairs should decay in the lepton+jets $t\bar{t}$ channel with one of the associated W bosons decaying into a quark/anti-quark pair (hadronically) and the second W boson into a charged lepton and a neutrino (leptonically). Thus, it was expected to have around 250,000 events with the decay topology $t\bar{t} \to b\bar{s}W^+W^-$ and similar amount of events for the other interesting combinations $t\bar{t} \to b\bar{d}W^+W^-$ and $t\bar{t} \to b\bar{b}W^+W^-$ including all charge conjugation states.
Additionally, it was decided to go for another sample having the same statistics but with *particle data group* (PDG) parameters[2] resulting from measurements and theoretical constraints [44], namely with the BR$(t \to b + W) \approx 99.83$ %, BR$(t \to s + W) \approx 0.16$ % and BR$(t \to d + W) \approx 0.0074$ %. With these two choices, all possible cases were covered with a high number of events. The request with PDG, i.e. SM, CKM matrix elements gives the advantage to perform a comparison with other top generators.

4.1.2. PROTOS truth samples and validation

Before starting with the production itself, a validation of the `PROTOS` dataset was needed. A small number of events was created for a comparison with other generators. Specific job options and event parameters were selected which are crucial for the simulation. The following settings have been used, which are identical to the final production parameters: First, τ lepton decays were included and ISR as well as FSR have been turned on which differs from the settings of the 7 TeV dataset where only the ISR was fully turned on. Moreover, the ATLAS default `PYTHIA 6` parameters[3], including the tune “Perugia2011C” [45] based on CTEQ6L1 PDFs [46], were used. They comprise the following values:

---

[2] The PDG parameters for the CKM matrix are further labelled as SM parameters.

[3] ATLAS internally known as the job options “Pythia Perugia2011C Common”

- Weak mixing angle $sin^2\theta_W = 0.23113$
- Mass of the top quark $m_{top} = 172.5$ GeV
- Mass of the W boson $m_W = 80.399$ GeV
- Mass of the Z boson $m_Z = 91.1876$ GeV

The corresponding widths of the particles were calculated during the processing, based on theoretical predictions. Additional parameters describing the electroweak processes were adopted from the official ATLAS recommendations [47], which are compatible with the values from the PDG 2010 [44]. The most important ones were the default "SM BR" originating from the CKM matrix elements:

$$V_{td} = 0.00862, \quad V_{ts} = 0.04030, \quad V_{tb} = 0.999152$$

In addition, the two packages `Tauola` [48] and `Photos` [49] were included, describing the hadronic or leptonic tau decays and higher order QED radiative corrections, respectively. Furthermore, `PYTHIA` specific parameters have been applied as recommended by the ATLAS collaboration:

- Complete event records for the first event were dumped (*"pyinit pylistf 1"*)
- Particle data as well as decay process data were dumped after initialization (*"pyinit pylisti 12"*)

After the choice of the parameters, events without any detector interactions (so called truth n-tuples) were generated and reconstructed. Those can be used for a comparison with the ATLAS approved MC generators.
For such a validation, several information was still missing. More precisely, for the `POWHEG` + `PYTHIA` and `MC@NLO` + `HERWIG` datasets, truth n-tuples have not been available and so had to be produced similar as described above, starting with generated events (also called "EvGen" samples), which were available within the ATLAS collaboration. Both samples consist of a non-allhadronic sample (combination of lepton+jets and dileptonic samples) instead of the pure lepton+jets sample, as for `PROTOS`. Instead of producing two new lepton+jets samples (for `POWHEG` and `MC@NLO`), an additional dileptonic `PROTOS` sample was simulated for the comparison, to also have a non-allhadronic `PROTOS` sample on hand.

With these datasets, the most crucial distributions could be thoroughly validated. This is illustrated in Figure 4.1 and Figure 4.2 for electrons and jets, respectively. The `PROTOS` + `PYTHIA` sample with the SM BR (Dataset-ID 110608) is drawn in red while the dataset of `POWHEG` + `PYTHIA` (Dataset-ID 117050) is indicated in blue and the `MC@NLO` + `HERWIG` dataset (Dataset-ID 105200) has black lines. Additional validation plots for

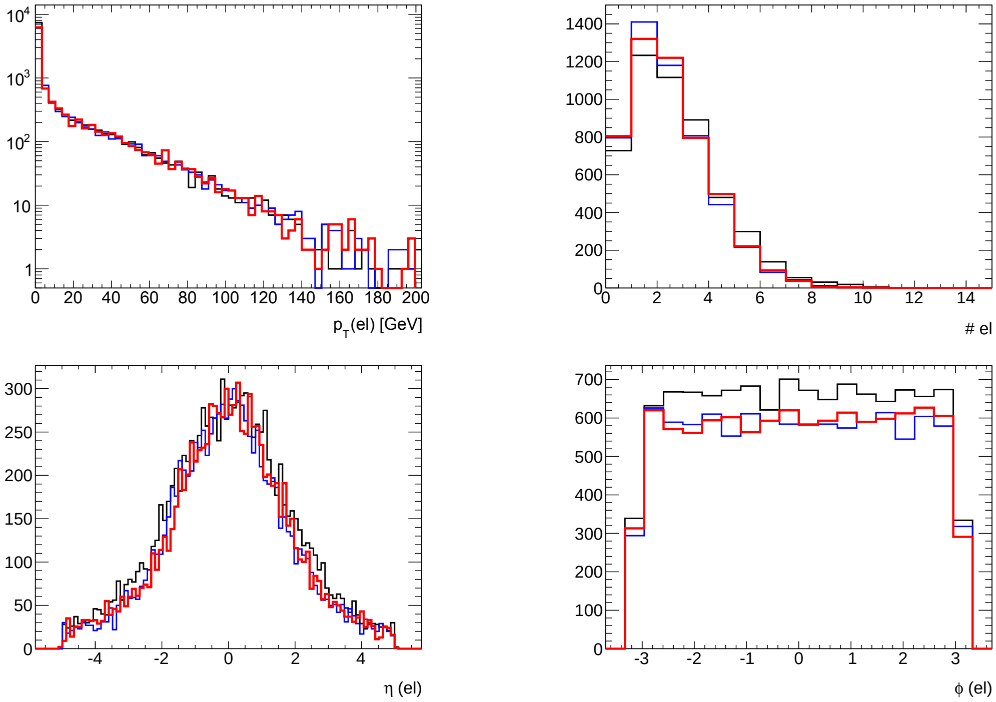


Figure 4.1.: Comparison of `PROTOS` generated events (red) with those generated by `POWHEG` (blue) and `MC@NLO` (black). In the four plots the distributions for p_T, number of electrons and their η and ϕ distributions are illustrated. The `PROTOS` distributions are compatible with the two approved generators.

muons (Figure A.1) and for the missing transverse energy distributions (Figure A.2) can be found in the Appendix A.
Overall, the shapes of the three different MC generators agree very well.[4] In most cases, the `PROTOS` + `PYTHIA` distribution lies in between `POWHEG` + `PYTHIA` and `MC@NLO` + `HERWIG`.

In addition, the `PROTOS` sample with the enhanced BR (non-SM BR) was cross-checked by comparing it to the `PROTOS` samples with SM BR. A selection of plots (red: non-SM BR, black: SM BR) can be found in Figure 4.3 (a,b) for jets and in Figure 4.3 (c-f) for electrons. All distributions look reasonable while emerging discrepancies can be explained in analogy to the $|V_{ts}|$ study presented in section 7.2 and are caused by the different physics processes in the two samples.
Additional plots can be found in the Appendix A (Jets: Figure A.3, Muons: Figure A.4, Missing transverse energy: Figure A.5).

---

[4] All distributions were validated by the responsible ATLAS working groups (details in [50])

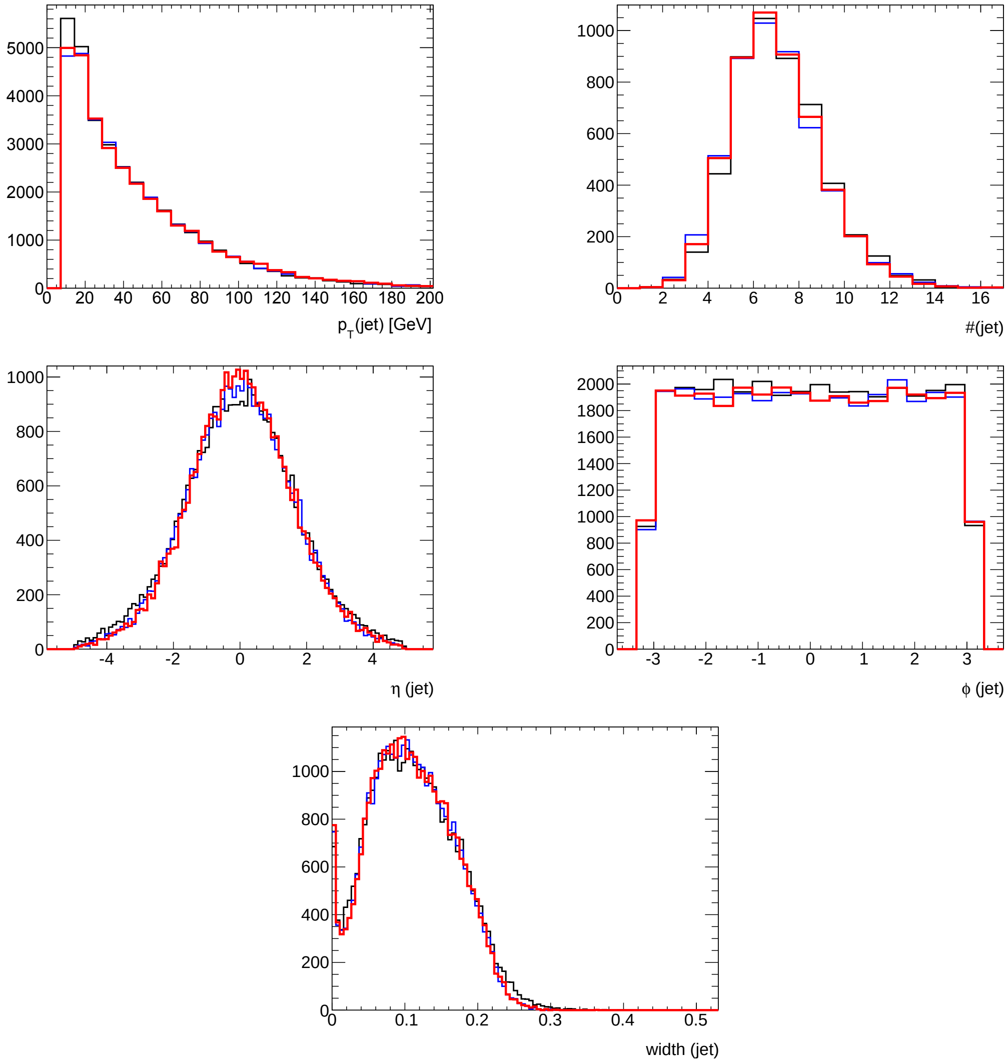


Figure 4.2.: Comparison of `PROTOS` generated events (red) with those from `POWHEG` (blue) and `MC@NLO` (black), similar to Figure 4.1 but here for jets. In addition the size of the jet cone is illustrated in the last plot. All `PROTOS` distributions are in agreement with the two approved generators.

4.1.3. Production of PROTOS MC samples

After a successful validation, the final workflow of the event generation was started. The resulting MC samples can then directly be used for physics analysis.

Since `PROTOS` is not able to produce the required 1,000,000 hard scattering events at once, they were split into 200 data files. Every file was produced separately with different initial generator seeds. To make sure that all data files contain at least 5000 events after reconstruction, the `PROTOS` authors recommend to use a 10 % safety margin to allow for necessary mapping of the phase space for integration. Thus 200 files, with a

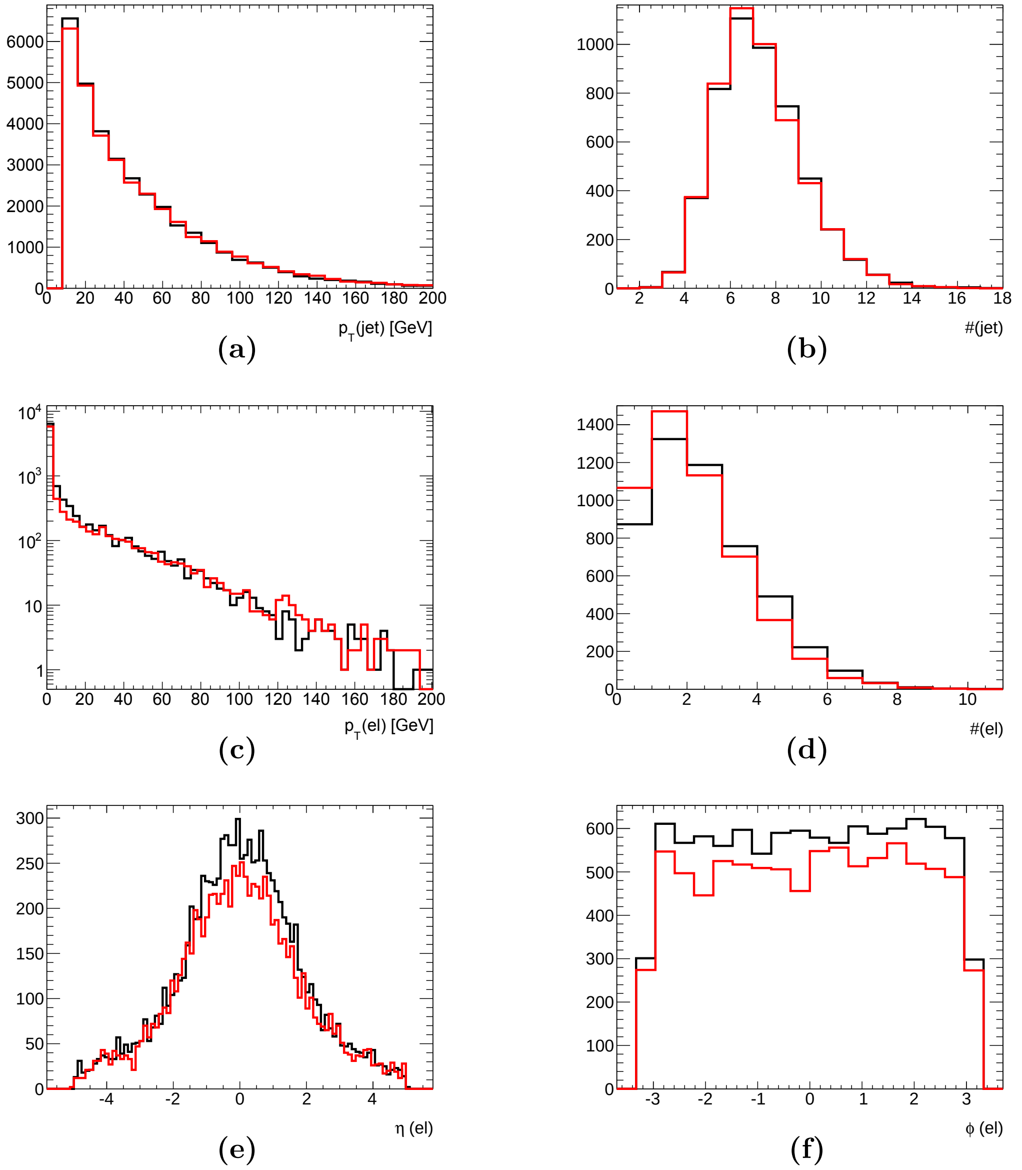


Figure 4.3.: (a,b) Comparison of the two `PROTOS` generated datasets with respect to the jet properties: SM BR (black) vs. non-SM BR (red). As explained in detail in section 7.2 the p_T^{jet} of $t \to s + W$ is slightly higher compared to $t \to b + W$ decays causing the emerging differences in (a). Also the number of jets per event (b) is affected. Taking those effects into account the jet distributions are in good agreement with the expectations.
(c-f) Comparison of the two `PROTOS` generated datasets with respect to the electron properties. The visible discrepancies for the electron p_T (c) and for the number of electrons (d) stem from semileptonic B hadron decays which are further discussed in section 7.5. The shapes of the η (e) and ϕ (f) distributions seem to be identical within statistical fluctuations, when taking the total number of electrons into account.

minimum of 5500 events each, were produced locally.
In the next step the events were “unweighted”. All `PROTOS` events are related to the phase space of the generated channel which is represented by the event weight. The “unweighting” procedure saves a lot of CPU time since there is no need for simulating millions of events to cover all phase space regions including those with a small probability density. [51]
Having these unweighted hard scattering events from `PROTOS` available, the general production could start with the help of the program package “D3PDMaker” [52]. For that, four separate steps [53] were performed centrally on the GRID:

- The event generation with `PYTHIA` is the first step after the simulation of the hard processes with `PROTOS`. `PYTHIA` receives the MC particles as input to simulate further decays of all particles as well as showering and hadronisation. Moreover pileup and similar processes are calculated for each event.

- The detector simulation with Geant4 [54] calculates the trajectories in ATLAS. Interactions with the detector are identified and recorded by taking the particle properties (e.g. momentum) and the different sensitivities of the detector material into account. There are in principle two ways to simulate these interactions:
  - Simulating the detector response in full, which is CPU intensive.
  - Using the package ATLFAST II (AFII) which approximates the energy profiles in the calorimeter. The inner detector simulation is identical to the first option. The resulting CPU time is reduced by one order of magnitude compared to the full simulation. [55]

  ATLFAST II was the final choice since the important tracking information are identical and so no large differences are expected for the $|V_{ts}|$ analysis. [56]

- The digitisation translates the simulated interactions between particles and detector into electronic signals (“digits”), equivalent to the detector output. At this, the simulated energy deposit as well as the detector material have the main impact on the resulting electronic signals.

- The reconstruction uses the digits from the precedent step and aims for a reconstruction of trajectories and properties of the particles traversing the detector. This procedure is identical for data reconstruction. The results are the quantities which are used in the analysis.

The above mentioned steps were performed twice to obtain the two different `PROTOS` samples for further physics analyses which are accessible as “`mc12_8TeV.110608.ProtosPythia_P2011CCTEQ6L1_ttbar_singlelepton`” and “`mc12_8TeV.110609.ProtosPythia_P2011CCTEQ6L1_ttbar_singlelepton_nonSM_BR`” in the ATLAS database.[5]

---

[5] Dataset description: “ttbar production (l+jets) with non-SM BR for top”. Keywords: “top”, “b-tagging”, “topproperties”, “V_ts”. Contact person: “Christopher.Schmitt@cern.ch”/ “Christopher.Unverdorben@cern.ch”.

Their generator properties are listed in Table 4.1, while the two dataset parameters can be summarised as: 1,000,000 events each, `MC12 8TeV,` $t\bar{t}$ `lepton+jets channel, PROTOS + PYTHIA P2011C, AFII` and the ATLAS job options `MC12job-options-00-08-47`. The two samples differ regarding their BR:

- Dataset ID 110608: SM branching ratios for top quarks
- Dataset ID 110609: BR(t→d) = 0.25 , BR(t→s) = 0.25 , BR(t→b)= 0.50

The corresponding cross sections were calculated automatically during the simulation of the hard processes with `PROTOS`.

Table 4.1.: Properties of the `PROTOS` generated $t\bar{t}$-events. The "Dataset-ID" assigns a unique number for each sample in the ATLAS database. The "cross section" is assigned by the generator for each channel and has to be multiplied by the "k-factor" to get the best approximations available for now. Since the orders of calculation are limited in different generators, the k-factor accounts for the higher order corrections on the cross section value. The last column lists the number of generated events per channel. Differences to round numbers are caused by the rejection of events due to generator related effects. [57]

Process	Generator	Dataset-ID	cross section	k-factor	#Events
$t\bar{t}$ with SM-BR	Protos+Pythia	110608	252.89 pb	0.43800	984,499
$t\bar{t}$ with enhanced BR	Protos+Pythia	110609	252.89 pb	0.43800	999,998

4.2. Monte Carlo generator `MC@NLO` for $t\bar{t}$ events

Besides the samples from the leading order `PROTOS` generator, the $t\bar{t}$-MC-sample produced by the `MC@NLO` generator is used in the final steps of the $|V_{ts}|$ analysis. This second generator is chosen since it is the only ATLAS approved next-to-leading order (NLO) generator which simulates the top BR as predicted by the CKM matrix (see also section 4.1.1). With in total about 15,000,000 events in the non-allhadronic sample, there should be around 40,000 events in the included lepton+jets channel with a top quark decaying to a strange quark plus associated W boson. Compared to 250,000 events from the `PROTOS` sample with enhanced BR this is still a reasonable number of events for the final $|V_{ts}|$ analysis steps (chapters 8,9). Nevertheless, the s-quark studies in chapter 7 are based on the `PROTOS` sample with enhanced BR due to the 6 times higher statistic.

The `MC@NLO` dataset was generated centrally by the ATLAS collaboration during the MC production campaign 2012, internal named as "mc12a"[58]. It was processed with Herwig 6.520.2, the "ATLAS Underlying Event Tune 2" (AUET2) [59] and the PDF set CT 10 [60]. A full detector simulation was performed with the package Geant4. Additional generator properties can be found in Table 4.2.

Table 4.2.: Properties of the `MC@NLO` generated $t\bar{t}$-events. A description of the parameters can be found in the caption of Table 4.1.

Process	Generator	Dataset-ID	cross section	k-factor	#Events
$t\bar{t}$	MC@NLO+Herwig	105200	112.94 pb	1.2158	14,997,103

4.3. Background estimation

For the reconstruction of the very rare decay $t \to s + W$, a clean $t\bar{t}$ sample is mandatory. This can be achieved with several different selection criteria as described in chapters 5 and 6. Nevertheless, there are some additional physics processes which have a similar topology and so can pass the selection cuts as well. For an optimisation of these cuts and in order to perform the $|V_{ts}|$ analysis itself, the important backgrounds have to be modelled. All in all, five main processes are taken into account. All b-quarks in the following diagrams could also be replaced by light-/c-quarks, which are misidentified as b-quarks. The signal topology of $t\bar{t}$ events follows

$$t\bar{t} \to WW\, b\bar{q} \to l\nu\, q\bar{q}\, b\bar{q}$$

which is illustrated in Figure 6.3 of section 6.3.

- **Single top processes**
  The single top background is one of the most dominant backgrounds in this analysis. For the reconstruction of a $t\bar{t}$ event in the lepton+jets channel at least one b-quark has to be identified since top quarks are decaying in 99.8 % of all cases in b-quarks. A minimum of 2 b-quarks is not requested due to the search for $t \to s + W$ decays.
  With requiring only one b-quark-jet in the signal events, single top events can imitate those. The additional jet(s) and, if required, charged leptons[6] for the selection criteria can arise from gluon radiation or bremsstrahlung. Therefore single top events (left-hand diagram of Figure 4.4) can be misidentified as $t\bar{t}$ events.
  For another single top process, as illustrated in the right-hand diagram of Figure 4.4, even the topology is identical to $t\bar{t}$ processes. Here, only the reconstructed invariant top mass of $W + b$ differs.

---

[6] The absent missing transverse energy of neutrinos can either arise from detector effects or incorrect particle reconstructions.

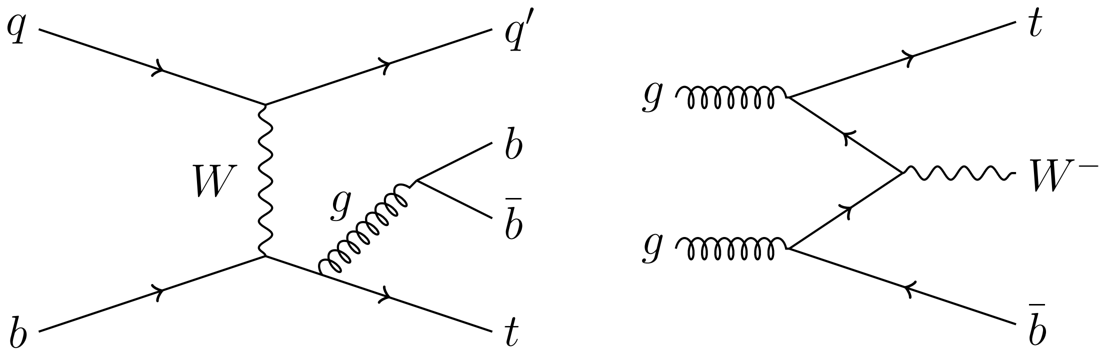


Figure 4.4.: Feynman diagrams illustrating two possible single top processes.

- **W + jets processes**
  For this channel, the exemplary $W + b\bar{b}$ process shows the similarity to $t\bar{t}$-events. It is already very close to the signal topology $t\bar{t} \rightarrow WWb\bar{q}$, whereas the difference is only a missing W boson, which can decay to $l\nu$ or $q\bar{q}$. This can be faked by either bremsstrahlung (including pair production) resulting in an additional charged lepton, or by gluons caused by higher order perturbative corrections, decaying into further quarks.

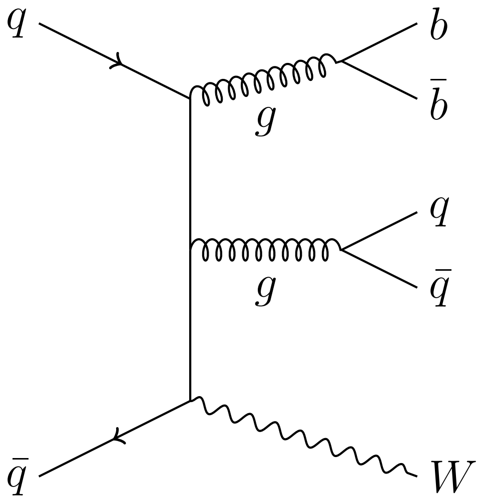


Figure 4.5.: Feynman diagram illustrating one possible W + jets process.

- **Z + jets processes**
  A Z boson can directly decay into two virtual W bosons. Gluon radiation results in the required jets whereof only one has to be (mis-)identified as a b-quark to end up with the required signal topology $t\bar{t} \to WWb\bar{q}$.

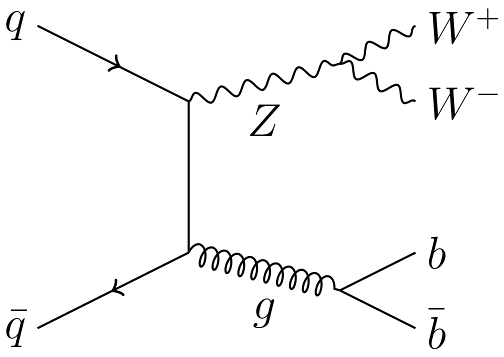


Figure 4.6.: Feynman diagram illustrating one possible Z + jets process.

- **Diboson processes**
  For diboson events the statements of “Z + jets” and “W + jets” can be combined to end up with an event shape similar to the lepton+jets channel of top pairs. For the exemplary process (Figure 4.7), further jets arising from higher orders identify the event topology of $WW + jets$ with $t\bar{t}$ events.

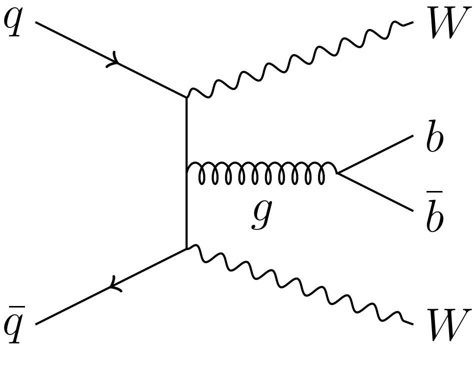


Figure 4.7.: Feynman diagram illustrating one possible Diboson process.

The MC simulation parameters of the four different background channels are summarised in Table 4.3. Some aspects are identical for all of them:
The detector response was modelled in full simulation with the package Geant4, similar

to the $t\bar{t}$-signal samples. The PDF was set to CTEQ6L1. For the single top, $W \to l\nu$ and $Z \to ll$ processes the `PYTHIA` version 6.426 was chosen, while for the $W + quark(s)$ events it was `PYTHIA` 6.427, all with the UE tune P2011C. In contrast, the diboson production used Herwig 6.520.2 with the UE tune AUET2. [61, 62, 63, 64]

Besides the MC simulated backgrounds, **QCD-multijet processes** may also have a similar topology like $t\bar{t}$ events. This background consists in principle of a multitude of jets. On the contrary, the top pair signal events have 4 jets (incl. 1 b-tagged jet), 1 charged lepton and missing transverse energy ($\not{E}_T$) stemming from a neutrino in the final state. In the QCD-multijet processes, the charged lepton could e.g. be faked by hadrons decaying in flight or by misidentified jets (non-prompt or fake leptons [65]) while the $\not{E}_T$ can arise from detector effects or incorrect particle reconstructions.
The estimation of this QCD background is performed with a data driven "matrix method". The data events utilised here are "tight" electron and muon events as well as "medium" muon and "medium++" electron events (cf. chapter 5). The main differences between these two categories is that "medium(++)" leptons have more relaxed identification criteria, compared to "tight" leptons, mainly based on isolation requirements, and thus can stem from in-flight decays of hadrons or from misidentified jets faking a charged lepton. The central aim of the method is the determination of the fake rate of medium charged leptons which still pass the tight criteria. The corresponding number of medium and tight charged lepton events can be expressed as follows:

$$N_{medium} = N_{medium}^{real} + N_{medium}^{fake} \tag{4.1}$$

$$N_{tight} = N_{tight}^{real} + N_{tight}^{fake} = \epsilon_{real} \cdot N_{medium}^{real} + \epsilon_{fake} \cdot N_{medium}^{fake} \tag{4.2}$$

Here ϵ_{fake} (ϵ_{real}) is the efficiency for a fake (real) charged lepton to pass the tight selection if it had already passed the medium selection cuts. These equations can also be rewritten in a matrix form which is the reason for the naming:

$$\begin{pmatrix} N_{medium} \\ N_{tight} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \epsilon_{real} & \epsilon_{fake} \end{pmatrix} \begin{pmatrix} N_{medium}^{real} \\ N_{medium}^{fake} \end{pmatrix} \tag{4.3}$$

The important number of "fake leptons reconstructed as real leptons", N_{tight}^{fake}, can also be expressed as $\epsilon_{fake} \cdot N_{medium}^{fake}$, resulting in

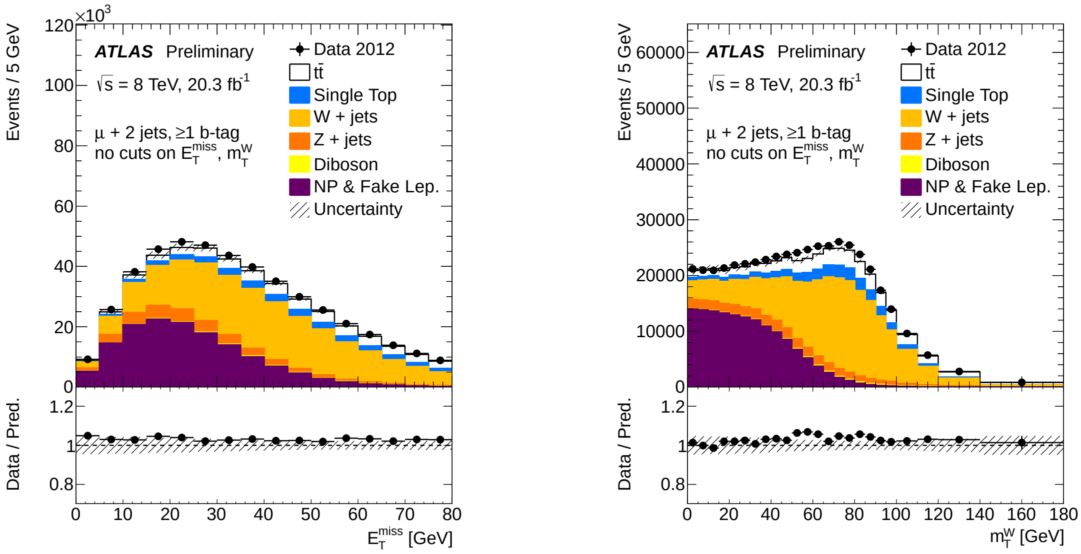
$$N_{tight}^{fake} = \epsilon_{fake} \cdot N_{medium}^{fake} = \frac{\epsilon_{fake}}{\epsilon_{fake} - \epsilon_{real}} (N_{tight} - \epsilon_{real} \cdot N_{medium}). \tag{4.4}$$

A corresponding matrix method weight w can now be derived which is applied to data events to obtain the background estimation.

$$w = \frac{\epsilon_{fake}}{\epsilon_{fake} - \epsilon_{real}} (b_{tight} - \epsilon_{real} \cdot b_{medium}) \tag{4.5}$$

with $b_x = 1$ for events passing the tight/ medium criteria, $b_x = 0$ otherwise. The corresponding efficiencies ϵ_{real} and ϵ_{fake} were extracted from data events with an highly

enriched number of either real or fake charged leptons as described in [65].
The results are in good agreement with the number and shape of the QCD multijet background processes as visible in Figure 4.8. [65, 66]



(a) **(b)**

Figure 4.8.: Distributions of (a) $\not\!\!E_T$ and (b) $m_{\mathrm{T}}^{W} = \sqrt{2\, p_{\mathrm{T}}^{lep} \not\!\!E_T (1 - \cos \Delta\phi_{lep,\not E_T})}$. The selection demands exactly two jets of which at least one is b-tagged, enhancing the fraction of QCD-multijet processes (labelled as "NP & Fake Lep"). A good description of the data is achieved. Figures taken from [65].

Table 4.3.: Overview of all background MC samples used for this analysis that have been produced by ATLAS. The "NpX" value in the process names represents the *additional* ***N**umber of **p**artons or jets in the event* which arises due to higher order perturbative calculations. A description of the parameters can be found in the caption of Table 4.1.

Process	Generator	Dataset-ID	cross section	k-factor	#Events
Single top (t)	AcerMC+Pythia	110101	25.750 pb	1.1042	8,997,672
Single top (s)	Powheg+Pythia	110119	1.6424 pb	1.1067	5,985,993
Single top (Wt)	Powheg+Pythia	110140	20.461 pb	1.0933	999,692
$W \to e\nu$ (Np0)	Alpgen+Pythia	147025	8127.3 pb	1.133	29,434,220
$W \to e\nu$ (Np1)	Alpgen+Pythia	147026	1792.7 pb	1.133	48,155,904
$W \to e\nu$ (Np2)	Alpgen+Pythia	147027	542.18 pb	1.133	17,554,347
$W \to e\nu$ (Np3)	Alpgen+Pythia	147028	147.65 pb	1.133	4,985,287
$W \to e\nu$ (Np4)	Alpgen+Pythia	147029	37.736 pb	1.133	2,548,292
$W \to e\nu$ (Np5)	Alpgen+Pythia	147030	11.962 pb	1.133	799,192
$W \to \mu\nu$ (Np0)	Alpgen+Pythia	147033	8127.1 pb	1.133	31.965.655

Process	**Generator**	**Dataset-ID**	**cross section**	**k-factor**	**#Events**
$W \rightarrow \mu\nu$ (Np1)	Alpgen+Pythia	147034	1792.9 pb	1.133	43.677.615
$W \rightarrow \mu\nu$ (Np2)	Alpgen+Pythia	147035	542.24 pb	1.133	17.611.454
$W \rightarrow \mu\nu$ (Np3)	Alpgen+Pythia	147036	147.66 pb	1.133	4.956.077
$W \rightarrow \mu\nu$ (Np4)	Alpgen+Pythia	147037	37.745 pb	1.133	2.546.595
$W \rightarrow \mu\nu$ (Np5)	Alpgen+Pythia	147038	11.970 pb	1.133	788.898
$W \rightarrow \tau\nu$ (Np0)	Alpgen+Pythia	147041	8127.1 pb	1.133	31,902,157
$W \rightarrow \tau\nu$ (Np1)	Alpgen+Pythia	147042	1792.2 pb	1.133	48,255,178
$W \rightarrow \tau\nu$ (Np2)	Alpgen+Pythia	147043	542.27 pb	1.133	17,581,943
$W \rightarrow \tau\nu$ (Np3)	Alpgen+Pythia	147044	147.64 pb	1.133	4,977,982
$W \rightarrow \tau\nu$ (Np4)	Alpgen+Pythia	147045	37.781 pb	1.133	2,548,295
$W \rightarrow \tau\nu$ (Np5)	Alpgen+Pythia	147046	11.959 pb	1.133	789,096
$W+c$ (Np0)	Alpgen+Pythia	200056	758.93 pb	1.52	22,999,046
$W+c$ (Np1)	Alpgen+Pythia	200057	274.47 pb	1.52	8,198,769
$W+c$ (Np2)	Alpgen+Pythia	200058	71.643 pb	1.52	2,090,290
$W+c$ (Np3)	Alpgen+Pythia	200059	16.482 pb	1.52	499,498
$W+c$ (Np4)	Alpgen+Pythia	200060	4.7824 pb	1.52	199,499
$W+bb$ (Np0)	Alpgen+Pythia	200256	52.237 pb	1.133	1,599,997
$W+bb$ (Np1)	Alpgen+Pythia	200257	45.628 pb	1.133	1,398,396
$W+bb$ (Np2)	Alpgen+Pythia	200258	23.955 pb	1.133	699,398
$W+bb$ (Np3)	Alpgen+Pythia	200259	13.633 pb	1.133	398,397
$W+cc$ (Np0)	Alpgen+Pythia	200156	149.39 pb	1.133	4,299,592
$W+cc$ (Np1)	Alpgen+Pythia	200157	143.90 pb	1.133	3,987,891
$W+cc$ (Np2)	Alpgen+Pythia	200158	84.227 pb	1.133	2,394,394
$W+cc$ (Np3)	Alpgen+Pythia	200159	44.277 pb	1.133	985,295
$Z \rightarrow ee$ (Np0)	Alpgen+Pythia	147105	718.97 pb	1.18	6,298,988
$Z \rightarrow ee$ (Np1)	Alpgen+Pythia	147106	175.70 pb	1.18	8,169,476
$Z \rightarrow ee$ (Np2)	Alpgen+Pythia	147107	58.875 pb	1.18	3,175,991
$Z \rightarrow ee$ (Np3)	Alpgen+Pythia	147108	15.636 pb	1.18	894,995
$Z \rightarrow ee$ (Np4)	Alpgen+Pythia	147109	4.0116 pb	1.18	398,597
$Z \rightarrow ee$ (Np5)	Alpgen+Pythia	147110	1.2592 pb	1.18	229,700
$Z \rightarrow \mu\mu$ (Np0)	Alpgen+Pythia	147113	719.16 pb	1.18	6,298,796
$Z \rightarrow \mu\mu$ (Np1)	Alpgen+Pythia	147114	175.74 pb	1.18	8,188,384
$Z \rightarrow \mu\mu$ (Np2)	Alpgen+Pythia	147115	58.882 pb	1.18	3,175,488
$Z \rightarrow \mu\mu$ (Np3)	Alpgen+Pythia	147116	15.673 pb	1.18	894,799
$Z \rightarrow \mu\mu$ (Np4)	Alpgen+Pythia	147117	4.0057 pb	1.18	388,200
$Z \rightarrow \mu\mu$ (Np5)	Alpgen+Pythia	147118	1.2544 pb	1.18	229,200
$Z \rightarrow \tau\tau$ (Np0)	Alpgen+Pythia	147121	718.87 pb	1.18	19,352,765
$Z \rightarrow \tau\tau$ (Np1)	Alpgen+Pythia	147122	175.76 pb	1.18	10,669,582
$Z \rightarrow \tau\tau$ (Np2)	Alpgen+Pythia	147123	58.856 pb	1.18	3,710,893
$Z \rightarrow \tau\tau$ (Np3)	Alpgen+Pythia	147124	15.667 pb	1.18	1,091,995
$Z \rightarrow \tau\tau$ (Np4)	Alpgen+Pythia	147125	4.0121 pb	1.18	398,798
$Z \rightarrow \tau\tau$ (Np5)	Alpgen+Pythia	147126	1.2560 pb	1.18	229,799
WW	Herwig	105985	12.416 pb	1.6833	2,499,890
ZZ	Herwig	105986	0.9908 pb	1.5496	245,000
WZ	Herwig	105987	3.6706 pb	1.9011	999,998

4.4. Data acquisition

Apart from the availability and understanding of MC datasets, the experimental input from the ATLAS detector is the crucial piece of the analysis itself since it displays the real data and potential new physics. The data used were collected in 2012 at a centre-of-mass energy $\sqrt{s} = 8$ TeV in proton-proton collisions and with a distance of 50 ns between the individual proton bunches. For the data acquisition and recording several steps had to be performed and additional information was taken into account. All reconstruction steps are in agreement with the recommendations of the ATLAS top working group [67].

4.4.1. Luminosity and run periods in 2012

The total integrated luminosity delivered by the LHC between the 4th of April and the 6th of December 2012 was about 22.8 fb^{-1} from which ATLAS could record about 93 %. Its rise during the 2012 Run I period is illustrated in Figure 4.9. Due to different issues with the detector, like downtimes of single sub-detector elements or problems with the reconstruction system, a so called "Good Run List" (GRL) was build for accepted runs. Summing those up, an integrated luminosity of 20.28 ± 0.57 fb^{-1} was found to be "good" for physics analysis. This value, including an error of 2.8 %, was determined at the end of the 2012 run period by the "ATLAS Luminosity Calculator" [68] combined with *"van der Meer* scans" [69]. For these scans, the two LHC proton beams were first separated, thus having no interaction point. Step-by-step they were then brought together until they had a maximum overlap. This procedure extracted the beam profiles. Combined with the external measured bunch-population, the luminosity could be determined [70, 71].
Furthermore, the detector performance and status had to be known. Therefore, the altering detector and accelerator conditions are taken into account which changed several times during data taking. For that purpose the whole dataset was split into 10 different data periods named A-E, G-J and L. The most significant differences between those runs were the different trigger conditions and hardware issues of the detector, e.g. small dead regions in sub-detectors. During each single period the machine conditions stayed mostly identical. The missing data periods F and K correspond to proton-lead runs (period F) and dedicated cosmic runs (period K). In addition there was also a stable proton-proton run called "period M" with a bunch spacing of only 25 ns. It was the first dedicated test for the LHC Run II and therefore not considered for this analysis. [72, 73]

4.4.2. Trigger selections

From the ATLAS data streams the interesting top pair events were selected by using "unprescaled" triggers[7]. The triggers in ATLAS are split into two different trigger levels (L1, L2) and one event filter (EF). The level 1 (L1) is implemented hardware-wise, the second level (L2) is based on software decisions whereas the event filter can draw on

---

[7] In comparison, a "prescaled" trigger selects only every x$^{\text{th}}$ event passing the trigger criteria for further processing. It is used for high event rates and relatively low trigger levels since it reduces the disk space consumption to a fraction 1/x compared to regular i.e. unprescaled triggers.

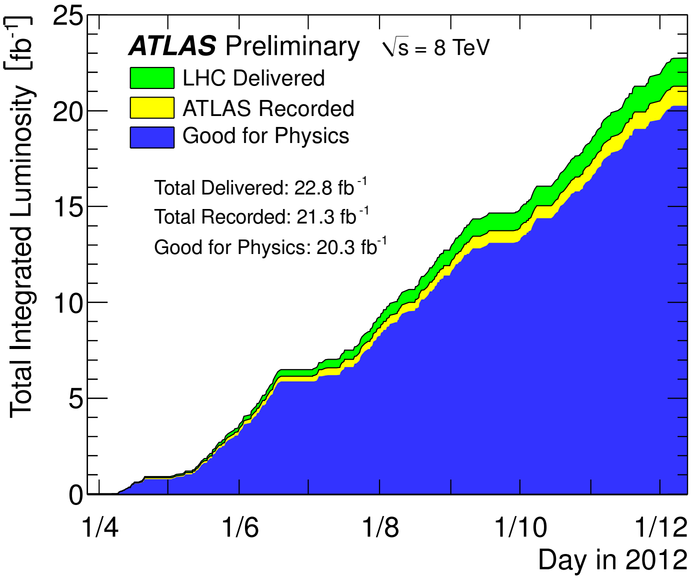


Figure 4.9.: Integrated luminosity of acquired data with the ATLAS detector in 2012 at a centre of mass energy of $\sqrt{s} = 8$ TeV. Figure taken from [74].

reconstructed events. L2 and the EF are also labelled as high level triggers (HLT) due to their software-based method. A more technical description can be found in section 3.2.6. This analysis made use of four different triggers, two for the electron channel and two for the muon channel. For the first trigger in the electron channel, the transverse energy deposit in calorimeter clusters $E_T = E/cosh(\eta_{cluster})$ had to reach a minimum of 18 GeV to pass the L1 decision. With the tracking and calorimeter information available the electron had to be reconstructed isolated[8] and as one single track with an E_T larger than 24 GeV in the HLT decision (`EF_e24vhi_medium1`). The second trigger however did not ask for any isolation criteria, but required $E_T > 30$ GeV on the L1 level, while at the HLT the E_T was set to be larger than 60 GeV (`EF_e60_medium1`).

For both muon triggers, the L1 criteria asked for a minimal transverse momentum of $p_T =$ 15 GeV for a muon track. The first trigger then requested that the reconstructed muon in the EF stage had to reach $p_{\mathrm{T}}(\mu) > 24$ GeV while also being isolated[9] (`EF_mu24i_tight`). The second muon trigger accepted also muons without isolation requirement if the muon p_{T} exceeded 36 GeV (`EF_mu36_tight`).

---

[8] Isolated e: Within a cone $R = 0.2$ in the $\eta - \phi$-plane about the e, there must be not more than 10 % of the electron-E_T stemming from other particles.

[9] Isolated μ: Within a cone $R = 0.2$ $\eta - \phi$-plane about the μ, there must be not more than 12 % of the muon-p_T stemming from other particles.

5. Physics Objects

Top anti-top pairs have a specific decay chain in the lepton+jets decay channel. The daughter particles are two quarks, ending in jets, and two W bosons, whereof one decays into 2 jets (hadronic decay), the other in a charged lepton + ν (leptonic decay). For a full reconstruction of these $t\bar{t}$ signal events and a simultaneous rejection of the background processes, all "physics objects" involved (final state particles plus possible additional jets) need to be identified very precisely. Hence, very specific object definitions are required, to end up with the demanded particles with high probability. To achieve a deep understanding, several ATLAS combined performance groups work on each physics object separately (e/γ, μ, τ, $\not{E}_T$, jets, flavour tagging) to determine the best requirements for each object.

The selections are based on various different information, including all sub-detectors and tracking information to reconstruct possible candidates. In sections 5.1 to 5.4 all objects used in this analysis (charged leptons, $\not{E}_T$, jets)[1] are presented, as well as the selection procedure itself that leads to a clean composition. All object definitions are identical to the analysis of [75] and in agreement with the proposals of the top working group as summarised in [67], aiming for the best combination of background rejection and high signal efficiency. Additionally a new method for cosmic muon rejection was developed during the research for this thesis which is presented in section 5.2.2.

After the particle and jet definitions, the reconstruction of vertices is discussed in section 5.5. This procedure is highly influenced by the number of simultaneous interactions per bunch crossing, called pileup, which is also presented here. Vertices and especially secondary vertices are of great interest for several b-tagging algorithms which are separately described in section 5.6.

5.1. Electrons

Possible electrons are first identified by energy deposits in the liquid argon (LAr) calorimeter (section 3.2.3) and have to be matched with reconstructed tracks from the inner detector.

In the offline selection, all candidates need a minimum of 30 GeV for the transverse energy E_T, and the very forward detector regions are excluded by the geometrical requirement $|\eta_{cluster}| < 2.47$. Also, all electrons in the transition region between barrel and endcaps of the calorimeter, $1.37 < |\eta_{cluster}| < 1.52$ (also known as "crack" region), are rejected due to a limited precision in that area.

---

[1] The ATLAS detector can also identify γ, but this is not prerequisite for the analysis in hand.

Additional selections based on the the shower shape, track quality and track-cluster matching are then performed, separating the remaining electrons in “loose++”, “medium++” and “tight++” categories [76]. The latter one is then used for further analysis, showing a good separation between isolated electrons and jets. This aspect is very important since electrons are also always reconstructed as jets since the jet reconstruction algorithm is based on calorimeter energy deposits, too. An additional cut on the perigee of the electron track is set to the longitudinal impact parameter $|z_0| < 2$ mm (cf. section 5.6).
To efficiently exclude electrons which are not arising from the relevant W boson decays $W \to e\nu$, isolation requirements are formulated to reject electron candidates with a misleading origin like e.g. hadrons and photon conversions. This procedure is performed by the “EisoTool2012” [77] with an background rejection efficiency of 90 %. It is sensitive to the calorimeter cells energy E around the electron within a cone of $R = 0.2$ in the $\eta - \phi$-plane, and the transverse momentum p_T of the sum of tracks within $\Delta R = 0.3$.
In contrast to the default ATLAS top working group recommendations [67] the E_T threshold a “tight++” electron has to pass, is raised from 25 GeV to 30 GeV due to imprecise E_T modelling in the signal region as stated in [75].
Inexact modelling of the trigger and identification efficiencies end up in discrepancies between simulation and measured data, that are corrected by E_T and η dependent scaling factors. These correction values were obtained by the performance groups from tag-and-probe studies with W/Z boson samples, using e.g. $Z \to e^+e^-$ decays [67].
In addition to the “tight++” electrons, also “medium++” electrons without any isolation requirements are part of this analysis. They are used for the derivation of the QCD-multijet background (cf. section 4.3).

5.2. Muons and cosmic muon rejection

Apart from the trigger selection of muons (section 4.4.2), the “muid” algorithm [78] of the ATLAS Muon Combined Performance group is applied using a so called “tight” selection. This implies a combination of inner detector tracks and hits in the muon spectrometer as well as a specific number of hits in sub-detectors as described in [79], in order to be accepted as a possible muon candidate. The cut on the transverse momentum is recommended to be at least 25 GeV, but to maintain consistency with the electron criteria, a cut of $p_\mathrm{T} > 30$ GeV is applied. The geometrical selection is set to $|\eta| < 2.5$ to take account for the loss of detector sensitivity at higher pseudorapidity. The longitudinal impact parameter is set to $|z_0| < 2$ mm.

Similar to the electron cuts, isolation criteria are chosen to reject muons which are part of jets. For that purpose the ATLAS Top Working Group developed so called "mini isolation" requirements [76]:

$$\sum_{tracks} p_T^{track} / p_T^{\mu} < 0.05. \tag{5.1}$$

Only tracks with $p_T^{track} > 1$ GeV and within a p_T dependent radius R around the relevant muon have to be included in Equation 5.1.
In addition, an overlap removal is applied, rejecting all muon candidates within $\Delta R(\mu, jet) < 0.4$ around a jet with $p_\mathrm{T}^{jet} > 25$ GeV and with a jet vertex fraction $|JVF| > 0.5$ as defined in section 5.4. Events are also rejected for muons and electrons sharing the same inner detector track.
Similar to electrons, MC simulations and data do not agree perfectly for the muon identification efficiency. Therefore, correction factors were applied which were also derived by the Muon Combined Performance group. Isolation efficiencies however have a good data-to-simulation agreement within an error of 0.5 % [76].
In the QCD multijet background estimation, "medium" muons [78] with no isolation criteria are used as well, similar to electrons.

The above muon reconstruction steps also reject most, but not all, cosmic muons traversing the ATLAS detector. This particular background is relatively high with about 70 muons per second and m^2 at the surface, each with $E_\mu > 1$ GeV and many above 20 GeV [1]. A lot of these particles can reach or even traverse the ATLAS detector and can therefore fake collision muons. Up to now, only a geometrical rejection method is formulated for cosmic particles that are misidentified as two muons diverging from the collision point, as detailed in section 5.2.1. The $|V_{ts}|$ analysis does not apply any cosmic cuts since it allows only one muon per event. A clean identification of single cosmic particles, as proposed in section 5.2.2, could further enhances the signal significance in the $|V_{ts}|$ and other analyses.

5.2.1. Established method to reject cosmic muons

The established and still recommended rejection method only works for di-muon events. It states that the event has to contain a Primary Vertex candidate (see section 5.5) with at least five associated tracks, each with $p_\mathrm{T} > 400$ MeV, in order to be accepted as a possible signal event. In addition it also has to met geometrical criteria:
Events with two muons per event are rejected if (i) the transverse impact parameter of the two muons $|d_0|$ (cf. section 5.6) is larger than 0.5 mm and (ii) the ϕ-angle between them is close to π ($|\Delta\phi| > 3.10$ which equals 177.6°). Several analyses also include a third cut, $\eta_1 + \eta_2 \approx 0$, to support the decision based on back-to-back topology. This selection reduces the cosmic background significantly as illustrated in Figure 5.1 for the geometrical properties and in Figure 5.2 for the impact parameters. Nevertheless, some cosmics still pass the cuts and are misidentified as collision muons.

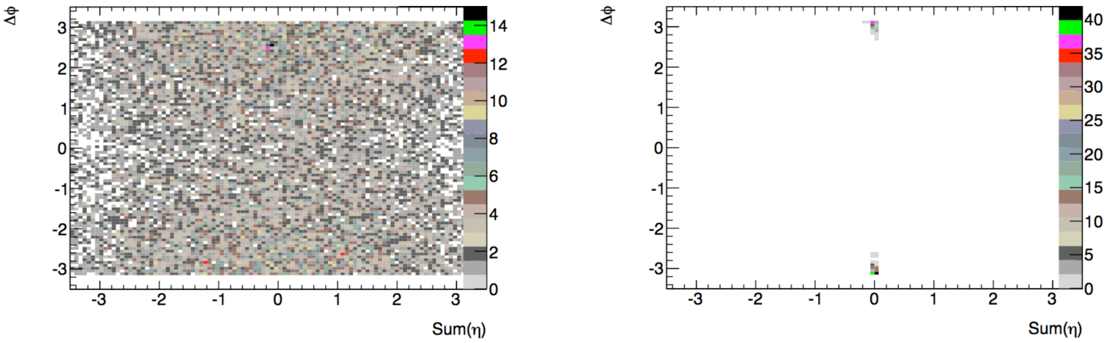


Figure 5.1.: The back-to-back topology of a cosmic muon can be used as a separation criterium if it was incorrectly reconstructed as a di-muon. The difference between a muon pair arising from collisions (left) and cosmic muons (right) is clearly visible. Figures adopted from [80].

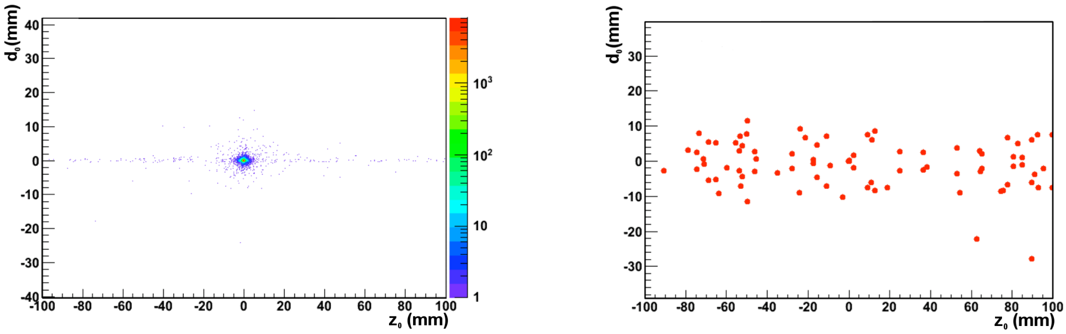


Figure 5.2.: Collision muons (left) generally have a transverse impact parameter d_0 close to zero. Cosmic muons (right) instead are randomly distributed. Figures adopted from [80].

5.2.2. Cosmic muon identification using timing information

The geometrical based decisions can be improved with additional timing information from sub-detectors. Compared to collision muons, cosmic muons only penetrate the surface in a top-down flight direction. By the usage of timing information at different sections in the ATLAS detector, three properties are in total available for the following cosmic study: Spatial information, flight direction and velocity of the particle in question.

For this study, cosmic muons have been recorded when empty bunches crossed each other during the 2012 data taking. This has the advantage that the detector conditions are identical for cosmic and collision events which also have been used. They stem from run period B with $\sqrt{s} = 8$ TeV. As a selection criterium, a p_{T} cut of $p_{\mathrm{T}}^{\mu} > 6$ GeV is applied on both datasets. For an additional cross-check the dedicated cosmic run in November 2012 (Period K) is available.

The ATLAS detector has a fixed time acceptance window in which particles can be associated to the actual bunch crossing. Cosmic muons, however, are entering the detector at random times. As a result, two categories of cosmic muons exist, each representing about 50 % of the cases:

- **Double-leg tracks:** These cosmic particles are reconstructed as di-muons with one particle-trajectory (leg) in the upper part of the detector, the other one in the bottom part. The time acceptance window is open during the whole flight of the muon through the detector and therefore two muon legs can be extracted.
- **Single-leg tracks:** In this case, only one half of the cosmic track lies in the time acceptance window of triggered collision events. Therefore only either the leg in the upper or in the bottom detector region is reconstructed. Those kind of cosmic muons could in principle pass the established method of cosmic muon rejection.

Collision particles are always flying outwards, from the interaction point (IP) to the outer region. The timing calibration takes this into account. For a particle emerging from a proton-proton collisions at $t = 0$ and flying outwards with $v = c$, all detector elements are calibrated to $t = 0$ at the time the particle is supposed to pass by. Exploiting this principle, a value for the “measured velocity” can be obtained.

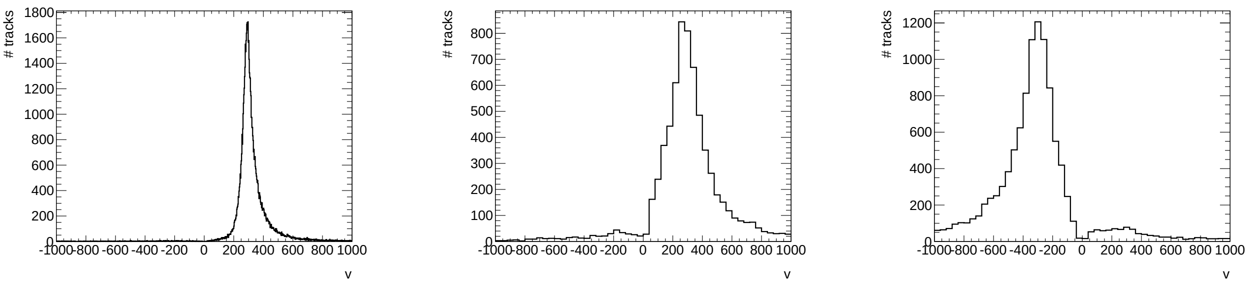


Figure 5.3.: Velocity values (mm/ns) of all collision muons (left), cosmic muons in the lower detector region (centre) and cosmic muons in the upper detector region (right). The measurement have been performed using timing information of sub-detectors.

In the bottom part, collision and cosmic muons are flying in the same direction: top-down. Due to the calibration, a direct velocity measurement yields $v_{\text{measured}} = \infty$ since those particles seem to fly a finite distance in no time. Collision muons in the upper detector part show the same behaviour. Cosmic particles instead are measured there with a velocity $v_{\text{measured}} = -2c$. With a recalibration from the ATLAS internal timing to the laboratory time measurement, cosmic muons in the lower detector part and all collision muons are supposed to fly with $v_{\text{real}} = c$. Cosmics in the upper detector are instead flying in the opposite direction with $v_{\text{real}} = -c$. The corresponding measurements are illustrated in Figure 5.3. The central peaks of the distributions have been fitted with a Gaussian distribution, resulting in a mean value of 295.2 ± 0.1 mm/ns for all collision muons, 296.3 ± 0.2 mm/ns for cosmic muons in the lower detector region and

-302.6 ± 0.4 mm/ns for cosmic muons in the upper detector region. In this way, selection cuts could in principle separate single cosmic muons in the upper detector region from other muons.

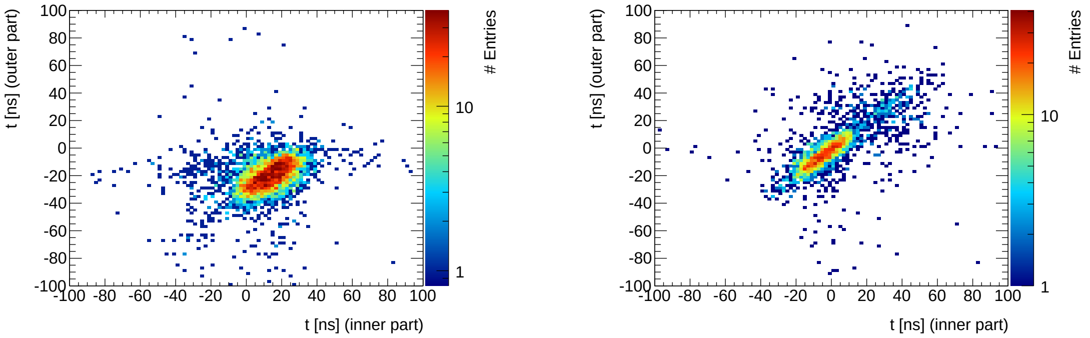


Figure 5.4.: Cosmic muons in the upper (left) and lower detector part (right) are shown in logarithmic z-scale. Both cosmic diagrams have been produced with loose selection criteria. The timing information was taken from the MDT layers.

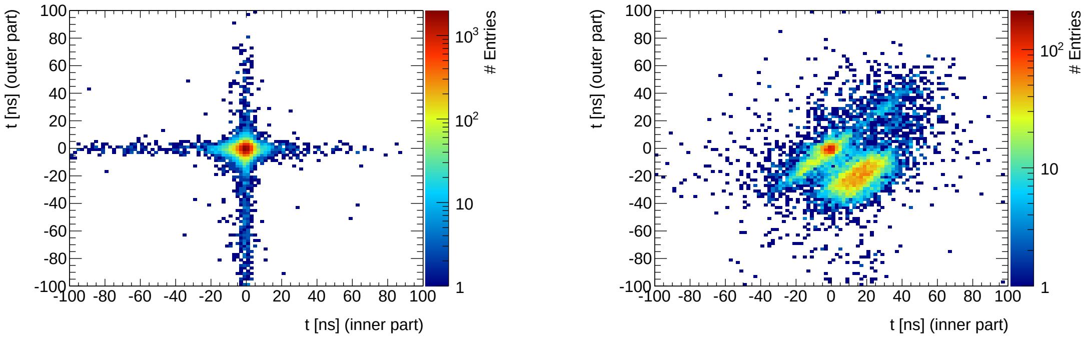


Figure 5.5.: Collision muons with a tight reconstruction algorithm (left) and with looser criteria (right) are shown in logarithmic z-scale. The timing information was taken from the inner and outer section of the MDT.

But the timing measurements contain even more information than the velocity, as illustrated in Figure 5.4 for cosmic and in Figure 5.5 for collision muons. Several aspects can be directly seen:

The distance between the depicted inner and outer MDT sections is about 5 m which corresponds to a flight time of 17 ns with $v = c$. This statement can be cross-checked by the cosmic muons in the upper detector (Figure 5.4 left). The time difference between the outer and the inner detector part is about 35 ns for most muons, corresponding to a velocity of $|v| = 2c$. Since cosmics do not enter the detector in-time with collisions, the time measurement is not clustered at one point but spread in an elliptical shape of ~40 ns in length, corresponding to the time acceptance window.

The same elliptical behaviour can also be clearly seen for cosmics in the lower part (Figure 5.4 right). In addition, muons arising from the subsequent timing window were also registered, visible in the region (40, 40). All cosmic muons have been reconstructed with relatively loose selection criteria to obtain an high enough number of events.
In contrast, collision muons which passed the tight selections are nicely centred (Figure 5.5 left). The particles which are forming a cross shape are negligible due to the logarithmic z-scale and can be explained with imperfect reconstruction algorithms. For comparison, collision muons with the identical loose selection criteria as the cosmic muons are illustrated in (Figure 5.5 right). Those combine all the above statements. First, a majority of reconstructed particles is nicely centred around (0, 0) stemming from in-time collisions. Moreover, two elliptical areas, contaminated by cosmics, are visible: The ellipse around the centre stems from the lower detector part, the one around (17, -17) from the upper detector. Additionally, muons from the subsequent bunch crossing were also registered in the area (40, 40).

The utilisation of all components of the ATLAS muon reconstruction system (cf. chapter 3 for details), but not only the MDT elements, improves the timing measurements a lot, which also enhances the discrimination of cosmic vs. collision muons. For this, the time determinations for all components are calibrated with respect to in-time collision events. The corresponding distributions including the uncertainties can be found in section A.2. The resulting timing information is then combined into two hypotheses, one for collision tracks, the other for cosmic tracks.

Collision hypothesis: Collision muons are expected to pass all components at the time $\mu = 0$ ns using the ATLAS internal clock. The uncertainty for the time measurement of the detector components considered is $\sigma_0 \approx 4$ ns. In simplified terms, this can be formulated as:

$$\chi^2_{collision} = \frac{\sum (t_i - \mu)^2}{\sigma_0^2} \qquad \text{with } t_i = t_{in},\ t_{centre},\ t_{out} \tag{5.2}$$
$$\mu = 0 \text{ ns}, \qquad \sigma_0 = 4 \text{ ns}$$

Cosmic hypothesis: On the contrary, the measured times for cosmics, when passing the detector elements, are not fixed but correlated. As an example the difference between the outer and the centred MDT components in the upper detector part is about 2.5 m corresponding to a measured time of 17 ns for cosmic particles with $v_{\text{measured}} = -2c$. For the bottom region the measured time is generally not equal to zero but the difference between the single detector sections should be approximately 0 ns. The corresponding uncertainties are identical to those from the collision hypothesis, with $\sigma_0 = 4$ ns.

$$\chi^2_{cosmic} = \frac{\sum (\Delta t_k - \mu_k)^2}{\sigma_0^2} \qquad \text{with } \Delta t_k = |t_{out} - t_{centre}|, ... \tag{5.3}$$

$$\text{Upper detector region: } \mu_k = 17 \text{ ns}, \quad \text{Lower detector region: } \mu_k = 0 \text{ ns}$$

The corresponding likelihood functions L and probability functions $P_n(\chi^2)$ can then be formulated with respect to the χ^2 value and to the number of degrees of freedom, which is connected to the number of detector sections used:

$$L \sim exp(-\chi^2/2) \tag{5.4}$$

$$P_n(\chi^2) = \frac{\chi^{n-2} e^{-\chi^2/2}}{\Gamma(\frac{1}{2}n) 2^{n/2}} \qquad \text{with n: degrees of freedom} \tag{5.5}$$

The “Cosmic Hypothesis and the “Collision Hypothesis” described by χ^2_{cosmic} and $\chi^2_{collision}$ are expected to be true if the corresponding probability distributions are flat, and expected to be false if a peak occurs around zero. For a better discrimination, the two distributions are compared in several 2-dimensional plots.

For the upper detector region, Figure 5.6 illustrates a distinct difference between collision muons (left) and cosmic muons (right) and implies that the hypothesis is clearly defined.

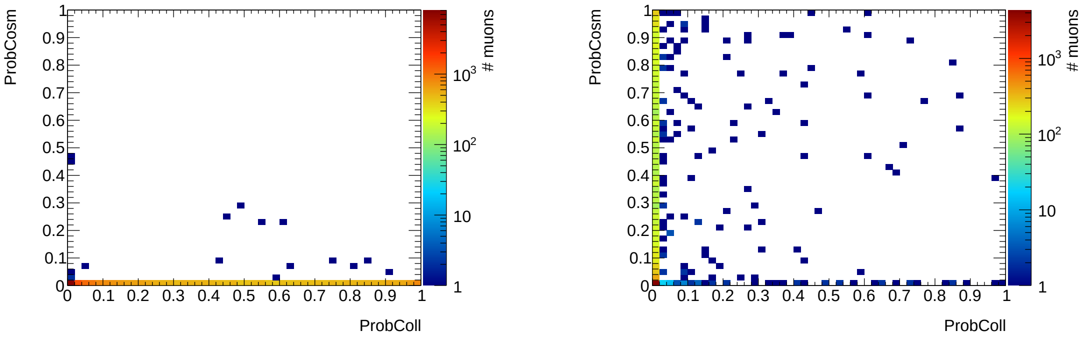


Figure 5.6.: Probability distributions to separate cosmic and collision muons in the upper detector part, with collision muons on the left side, cosmics on the right.

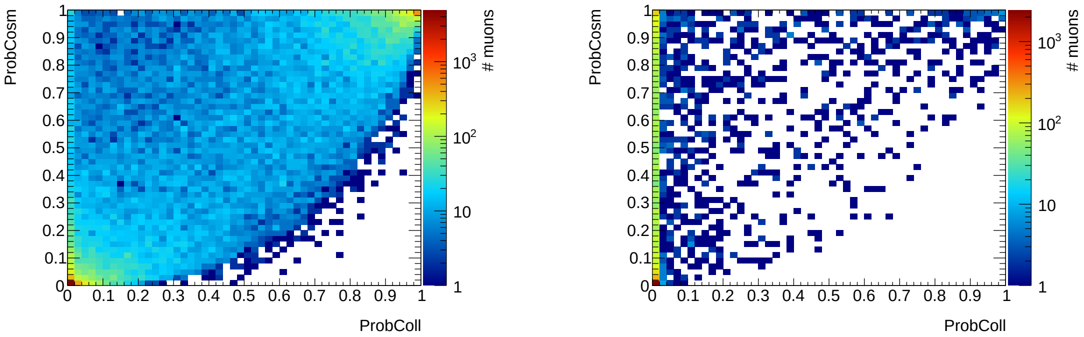


Figure 5.7.: Probability distributions to separate cosmic and collision muons in the lower detector part, with collision muons on the left side, cosmics on the right.

For the lower detector, Figure 5.7 indicates that the collision muons (left) are also compatible with the cosmic hypothesis, which was expected since it only takes time differences into account. Cosmic particles instead (right) are mostly incompatible with the collision hypothesis, emphasised by the logarithmic z-scale. Only cosmic muons which are traversing the detector in-time are compatible with the collision hypothesis.

By setting different cuts on the distributions, a high number of cosmic muons can be rejected:

- For the upper detector part, a cosmics rejection of 72.7 % ± 0.8 % can be achieved with losing only 0.016 % ± 0.007 % of collision muons.
- For the lower detector part, the cosmic rejection is found to be 37.0 % ± 0.7 % with rejecting 1.34 % ± 0.06 % of collision muons at the same time. Here difficulties with the MDT subsystem detector have been observed which could not be resolved yet. By leaving the MDT out of the hypotheses an improvement was achieved with 55.4 % ± 0.9 % rejected cosmic muons while rejecting 1.10 % ± 0.06 % of collision muons.

Further improvements of this new method could be achieved, first by combining the established geometrical rejection method with it, and second by assigning muons from the upper detector part to those of the lower part where a 50 % match is expected.
Apart from the application in cosmic particle rejections, this tool could also be used for analyses with low-β-particles like the search for R-hadrons [81, 82]. But initially the development intended to reject cosmic muons with the advantage of also reconstructing single-leg muon tracks, which is not possible with the established method presented in section 5.2.1. Nevertheless, the implementation in the ATLAS software package still requires some final validations. For a better comparison with collision muons, ATLAS performed an additional special cosmic run to record a reasonable number of cosmic muons passing the tight reconstruction criteria. After this validation, which is still underway, the cosmic rejection tool will be available inside the ATLAS collaboration.

5.3. Missing transverse energy

The "Missing Transverse Energy" ($\not\!\!E_T$, MET) is the only way in ATLAS to estimate the energy of possible neutrinos in an event. Since there is no considerable transverse momentum at the initial state due to protons flying only parallel to the beam line, the sum of the transverse energies of all physics objects per event has to equal zero. This measurement can be performed since the ATLAS detector covers almost 4π of the solid angle. A significant deviation from zero indicates the presence of a particle, that is not detectable. Within the SM, only neutrinos can pass the detector without leaving a trace, thus the missing transverse energy is accounted for by a neutrino. The

z-component of the neutrino can not be reconstructed with that method since the momentum of the two colliding partons ("Bjorken-x" in protons) is unknown.
In simplified terms, the $\not{E}_T$ is calculated as a vector sum of all calorimeter clusters and specified by additional tracking information from all reconstructed physics objects. The detailed calculation and algorithms can be found in [83]. In the $|V_{ts}|$ analysis, a $\not{E}_T$ larger than 30 GeV per event is requested due to the neutrino in the leptonic W boson decay.

5.4. Jets

Quarks do not occur separately as explained in section 2.2. They hadronise and form parton showers, called particle jets, in short jets. A deep understanding of these jets is crucial due to the presence of 4 to 6 jets per event in the $|V_{ts}|$ analysis. They are reconstructed with the infra-red and collinear safe anti-k_t algorithm [84] using the parameter $R = 0.4$ which is proportional to the jet size whereas the jet shape is in principle similar to a cone. The energies deposited in calorimeter cells are combined into topological clusters which are then used as inputs for the algorithm. Those clusters are calibrated by the "local cluster weighting" method [85, 86] that corrects for different calorimeter responses and fluctuations. Additionally, the following five adjustments are applied [67]:

(i) To bring the jet p_{T} virtually back to the parton level p_{T}, a jet energy scale factor (JES) is applied. The detailed procedure is explained in [86]. The calibration itself is evaluated using MC simulations where the resulting scale is p_{T} and η dependent. The corresponding uncertainties are extracted bin-wise and dependent mostly on p_{T}, η, light- and gluon-flavour composition as well as on in-time pileup (see section 5.5).

(ii) For the correction of the global effect of in-time pileup on a jet's momentum, the individual jet transverse momentum is diminished by an adaptable p_{T} value, which is dependent on the jet size and the event energy density.

(iii) To suppress extra jets due to the pileup effects, a cut on low-p_{T}-jets ($p_{\mathrm{T}}^{jet} < 50$ GeV) with a Jet Vertex Fraction $|JVF| < 0.5$ is applied. The latter variable is defined as the fraction of tracks in a jet which have the Primary Vertex (PV) as origin compared to all tracks of that jet, meaning that at most 50 % of the tracks have the PV as their origin. The related systematic uncertainties, resulting from discrepancies between data and simulations, are obtained by increasing and decreasing the cut value.

(iv) The jet energy resolution (JER) is generally in agreement with the energy resolution of MC simulations. Small differences are taken into account as an individual systematic uncertainty.

(v) Furthermore, so called “loose” jets [67], with a $p_{\mathrm{T}} > 10$ GeV, cause the whole event to be rejected. Those jets correspond to out-of-time energy deposits in the calorimeter caused by either beam-gas interactions, cosmic particles or from noisy calorimeter cells[2].

For the final selection, a pseudorapidity of $|\eta| < 2.5$ and a minimum transverse momentum of $p_{\mathrm{T}} = 30$ GeV are set, in order to exclude inefficiencies of the jet reconstruction algorithm. As for the isolation, jets overlapping within $\Delta R = 0.2$ with a possible electron are rejected. However, if a second jet exists with $p_{\mathrm{T}} > 25$ GeV and in a maximum distance of $R = 0.4$, the electron candidate is removed since it likely is a real jet. [67]

5.5. Vertices and Pileup

Vertices are the anchor points in an event. They are reconstructed by at least three tracks from the inner detector, crossing each other in one point within the uncertainties of the track reconstructions. Each track must have a minimum p_{T} of 400 MeV. The primary proton-proton collision point in the reconstruction is defined as the vertex with the largest sum of p_{T}^2 of associated tracks, while having a minimum of five tracks. This collision point is also called the Primary Vertex (PV). Events without an identified PV are removed. Any d_0 and z_0 cuts which are used for object quality criteria are performed with respect to that point.
All remaining vertices could in principle be identified as Secondary Vertices[3] (SV). These are crucial for b-tagging algorithms since they represent the decay point of relatively long-living particles like b-quarks. The distance between the Primary Vertex and the Secondary Vertex is then interpreted as the decay length as illustrated in Figure 5.8. Vertices which are in agreement with the decay point of already identified particles get rejected.

The vertex reconstruction is also closely related to the number of additional reconstructed *pp* collisions per event, called *pileup*. For the 2012 run period, an average number of pileup interactions $<\mu> \approx 20.7$ was reached. The corresponding intersection points are in general also reconstructed as vertices and have to be tested for the hard-scatter Primary Vertex [88]. Two origins are causing pileup:

- **Out-of-time pileup:** The first effect occurs due to the small distance between the proton bunches in the LHC. During the recording of the 2012 data, every 50 ns a bunch crossing took place. In comparison, subsystems of the ATLAS detector have much wider time windows of e.g. 600 ns, for parts of the LAr calorimeter

---

[2] The corresponding energy deposits can not be associated to particle sources. Those effects mainly arise from hardware problems.

[3] The naming convection of these vertices is analysis dependent. For a decay chain, e.g. $t \to (W+)b \to (W+)c$, secondary vertices can also be labelled as tertiary vertices, etc.

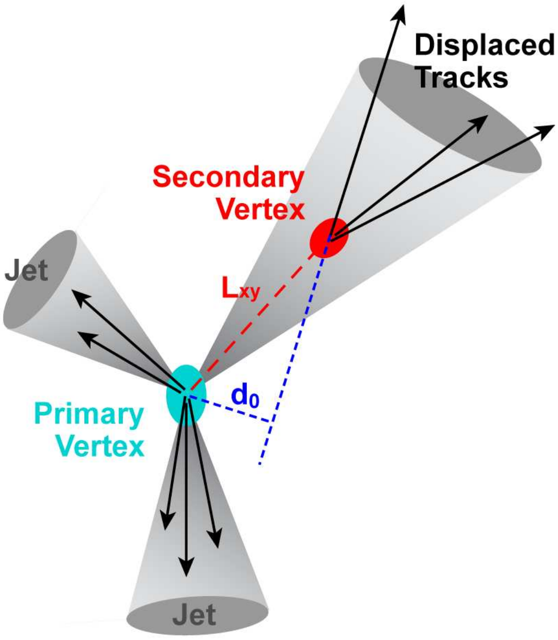


Figure 5.8.: Illustration of a possible hadronical decaying top quark. Due to the very short life-time of the top quark and the W boson (each $\sim 10^{-25}$ s) they seem to decay directly at the Primary Vertex. The b-quark with its longer lifetime decays instead at the Secondary Vertex as illustrated by the three displaced tracks. The distance between PV and SV yields the decay length L_{xy} of the b-quark which can be used for its identification by b-tagging algorithms as presented in section 5.6. Figure taken from [87].

[89]. Thus, more than 10 additional bunch crossings occur until the first one is fully registered. With sophisticated reconstruction algorithms, it is nevertheless possible to assign most particles to the correct bunch crossing.

- **In-time pileup:** The second pileup origin is due to the large number of protons per bunch. Several *pp* interactions can therefore take place simultaneously during one bunch crossing. More than 20 inelastic *pp* collisions can occur at the same time. In-time pileup is also often just named “pileup” since it is the most common problem faced by analysis groups.

To incorporate these experimental effects into MC simulations, events get reweighted by the average number of inelastic interactions per bunch crossing $<\mu>$. Furthermore a calibration adjusts the ratio “number of interactions per event” vs. “number of reconstructed vertices per event” in MC simulations to that in data. [88]

5.6. b-tagging algorithms

One of the most important aspects in the $|V_{ts}|$ analysis is the identification of jets arising from b-quarks. For this purpose, the ATLAS collaboration developed tagging algorithms which are based on different properties of B hadrons and the corresponding b-jets to discriminate those from gluon jets and jets arising from light quarks (u-, d-, s-quarks) or c-quarks. The dominant differences are the relatively large decay length due to the long lifetime, the large branching ratio to charged leptons, and the high B hadron mass which is connected to the invariant mass of the secondary vertex [67]. The inputs for all tagging algorithms are originating from calorimeter jets, from tracks of the inner detector, from the reconstruction of secondary and primary vertices as well as from the associated impact parameters d_0, z_0.

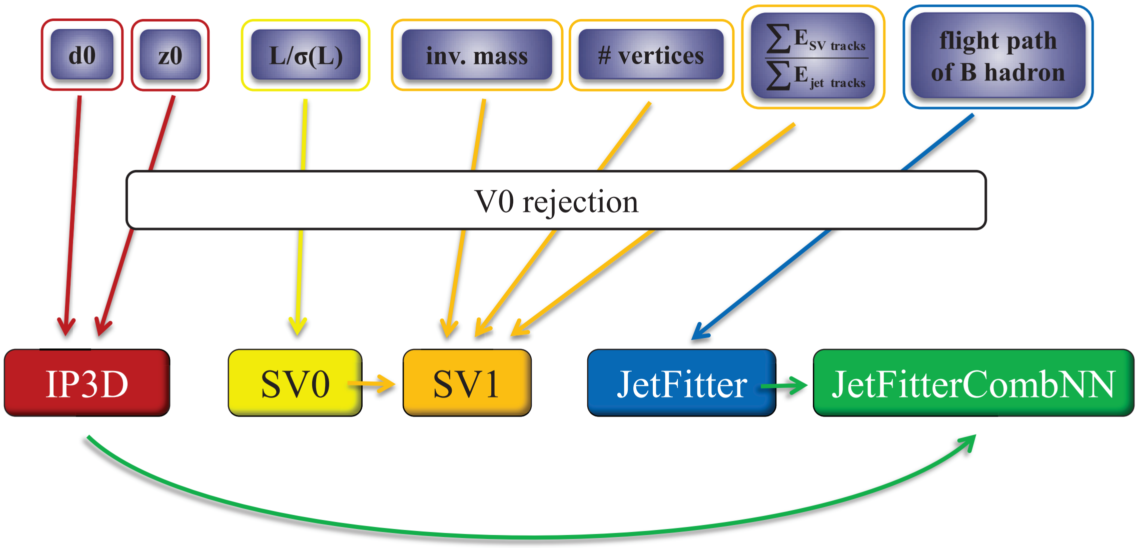


Figure 5.9.: Overview of most b-tagging algorithms and their input variables which are used in the ATLAS top quark group.

The most basic b-tagging algorithm is the **SV0** [90] which refers to a reconstructed secondary vertex (SV). The distance between PV and SV is measured in three dimensions resulting in a finite decay length L for the possible b-quark, as illustrated in Figure 5.8. When taking into account the uncertainty, the final SV0 variable corresponds to the decay length significance L/σ. Moreover, vertices that might result from detector material interactions are rejected.

The refined algorithm **SV1** [90] uses a “likelihood ratio technique” to combine the SV0 outputs with the invariant mass of all vertex tracks, the number of vertices with at least two tracks and the energy ratio of the associated tracks

$$\frac{\sum E_{\text{tracks of SV}}}{\sum E_{\text{tracks of corresponding jet}}}, \tag{5.6}$$

in order to achieve a better b-quark identification efficiency than SV0 alone.

Another b-tagger, the **IP3D** algorithm [90] uses a similar likelihood ratio technique as SV1, to combine the impact parameter significances d_0/σ_{d_0} and z_0/σ_{z_0} of the jet tracks. The transverse impact parameter d_0 is defined as the distance between the PV and the track's point of closest approach to that vertex, calculated in the r-ϕ-plane i.e. perpendicular to the beam axis, as displayed in Figure 5.8. The corresponding, longitudinal impact parameter z_0 is the distance between these two points in the z-projection i.e. in the direction parallel to the beam. The significances describe the quality of the two variables.
Thus, the resulting IP3D parameter represents the minimum distance of the jet to the primary vertex which can be used as a discriminator in the b-jet identification.

The developers of **JetFitter** [90] instead chose a totally different approach. Their algorithm tries to reconstruct the whole decay chain as well as the flight path of the B hadron and its daughter particles. For this purpose, a Kalman filter [91] identifies all decay steps

$$t \quad \rightarrow \quad (W+)\ b \quad \rightarrow \quad (W+)\ c \quad \rightarrow \quad ... \tag{5.7}$$

as well as the corresponding B and D hadrons in the jet.

A combination of the JetFitter and the IP3D taggers, based on artificial neural network decisions, yields the so called **JetFitterCombNN** b-tagging algorithm [90]. The neural network was trained with MC simulated events and the resulting b-jet identification is much more refined compared to the individual inputs.

All tagging algorithms presented above also have the purpose to reject V0 particles[4]. With that, contributions from misleading secondary vertices, originating from strange quarks or photon conversions, are reduced.

Three of the well established, high performance b-tagging algorithms can then be combined into an artificial neural network, called **MV1** (**1$^{\text{st}}$ M**ulti**V**ariate b-tagger) [92, 93]: IP3D, SV1 and JetFitterCombNN. The resulting output distribution is a likelihood function, distinguishing b-jets from other jets. In particular the discrimination power of b-quarks vs. s-quarks is the major reason for the utilisation of b-tagging algorithms in this analysis. All three inputs of the MV1 tagger already have a reasonable selection

---

[4] The term V0 stems from the decay topology ending with two charged tracks where the tracks form a shape similar to the letter V. The index 0 was introduced, due to the neutral charge of the mother particle. Typical V0 particles are K_s^0, Λ and γ, even though the latter one is sometimes not labelled as such a particle since it is not decaying via the weak interaction but undergoing pair production.

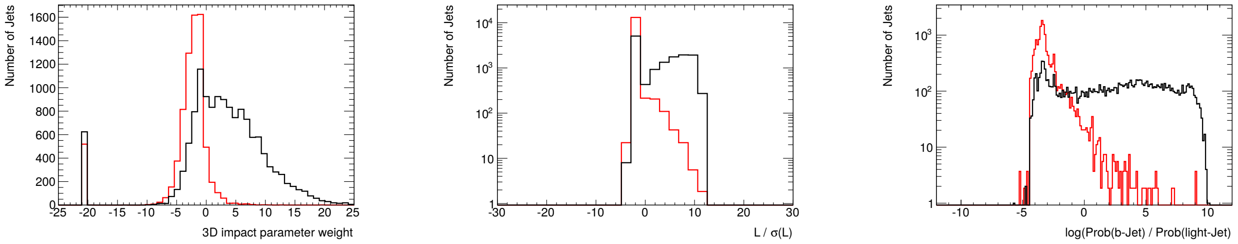


Figure 5.10.: The discrimination power of b-quarks (black line) vs. s-quarks (red line) is shown here, for the b-tagging algorithms IP3D (left), SV1 (centre) and JetFitterCombNN (right) on truth level, based on `PROTOS` generated events. Their weights are used as inputs for the MV1 algorithm.

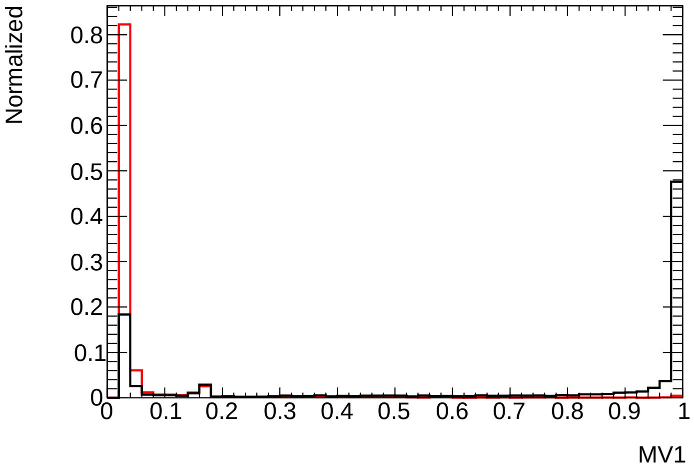


Figure 5.11.: Discrimination power of b-quarks (black line) vs. s-quarks (red line) illustrated for the probability weights of the MV1 b-tagging algorithm on truth level, based on `PROTOS` generated events.

efficiency of s-quarks, as shown in Figure 5.10, but by using the combination MV1, an even better discrimination can be achieved as illustrated in Figure 5.11.

The MC simulations of the three inputs are in agreement with data at the level of 10 % [90]. For a better data-to-MC modelling, the ATLAS top working group chose different MV1 working points. These correspond to b-tagging efficiencies of 60 %, 70 % and 80 % which can be translated to probability weight cuts of 0.9827, 0.7892 and 0.3511, respectively. For the $|V_{ts}|$ analysis, b-quark jets are reconstructed with the MV1 70 % operating point. With this 70 % b-tagging efficiency, the light-jet rejection factor is about 140, meaning that just one of 140 light-jets passes the tagger. For c-jet rejection, it results in a factor of about 5, while for τ-jets the rejection factor is close to 14. All these values have been determined by MC simulations.

To test and adjust the MV1 simulations as well as to obtain the underlying uncertainties, several procedures are available. The data-to-MC b-tagging calibration, chosen for this purpose, is a combination of the “System8”, the “Kinematic Selection” and the “Tag & Probe” methods:

- The **System8** calibration method, which is explained in detail in [94], was designed to be independent of MC simulations. The input data use a selection of jets with associated muons, stemming from leptonically decaying B hadrons, which leads to an enrichment of b-quarks. The whole method is based on a combination of eight equations describing the ratio between the expected and the observed events.
  As a result, 5-10 % more MC than data events are found for the MV1 tagger. However, the ratio is still compatible with unity when taking statistical and systematical uncertainties into account. [67]

- The **Kinematic Selection** based calibration method considers all possible b-decays, resulting in higher number of events but also in higher background contributions. The inputs are di-leptonic $t\bar{t}$ data events in which the b-tagging procedure was already performed, yielding a fixed number of b-tagged jets per event N_{tagged}^{data}. This number can be expressed as a combination of true b-jets with a correct tag (tagging efficiency ϵ_b) and non-b-jets with incorrect tags (mistagging efficiencies ϵ_x with $x =$ *light, c, τ, fakes*):

  $$N_{tagged}^{data} = N_b^{MC}\epsilon_b + N_{light}^{MC}\epsilon_{light}^{MC} + N_c^{MC}\epsilon_c^{MC} + N_\tau^{MC}\epsilon_\tau^{MC} + N_{fakes}^{MC}\epsilon_{fakes} \tag{5.8}$$

  With this formula the tagging efficiency ϵ_b can be extracted. The numbers of mistagged jets (N_x) as well as all mistagging efficiencies (ϵ_x) arise from MC simulations, except for the mistagging efficiency of fakes, ϵ_{fake}, which is derived from data events with same-sign charged leptons.[5]
  In this method, the resulting data-to-MC discrepancies for the b-tagging efficiency ϵ_b is nearly identical to unity, especially when all uncertainties are considered. [67]

- The **Tag and Probe** method’s principle idea is to tag one b-jet in a di-leptonic $t\bar{t}$ event (“tagged jet”) and use the kinematic properties of the event topology to choose an unbiased “probe jet”. As in the method before, all b-decays are considered. The resulting probe jets have a probability of up to 90 % to be a b-jet, depending on the p_{T} value and $\Delta\phi$ between tag and probe jet. With that sample the MV1 tagger can again be tested. Once more, this method claims consistency between data and MC simulation, taking all uncertainties into account. [67]

The final combination of these three calibration methods (including the uncertainties) is clearly consistent with unity, i.e. simulations and data coincide. In an η dependent

---

[5] In di-leptonic $t\bar{t}$ events, the two charged leptons are oppositely charged (opposite-sign) since they stem from the W^+ and W^- bosons. If the two leptons have identical charge (same-sign) they can not arise from $t\bar{t}$ events.

test, no discrepancy from unity is seen at all. The p_{T}-wise test instead indicates small discrepancies. Thus, a p_{T} dependent scaling factor is applied to rescale b-tagged MC events with a factor of about 0.98. An additional uncertainty of 2 % has to be applied. [92]

Further calibrations in the range 0.9 - 1.0 (depending on p_{T} and η) are applied for c-jets with a total uncertainty of 8-15 %. The light-jet calibration for the mistag rate uses a scaling factor of 1.1 - 1.5 with a total uncertainty of 15-40 %. [93]

No correlation between the calibration and the analysis is expected since the calibration method uses di-leptonic $t\bar{t}$ events, while the $|V_{ts}|$ analysis is based on lepton+jets $t\bar{t}$ events.

Another tagging algorithm used in the analysis is the **MV3** b-tagger [95]. Unlike the MV1 tagger, it is not officially calibrated by the ATLAS top working group, but still shows a good data-to-MC agreement, which is illustrated in Figure 7.32 of section 7.5. The tagger is split up into three sub-taggers with the basic purposes to discriminate b- from c-quarks, c- from light-quarks and b- from light-quarks where the light quarks are u-, d- or s-quarks. Nevertheless, this tagger is also not optimised for s-quark reconstruction arising from $t\bar{t}$ events due to the high momentum of these quarks. A direct application as a main discriminator in the $|V_{ts}|$ analysis is therefore not possible, but nonetheless it was tested as an input variable for the multivariate analysis presented in chapter 8.

Similar to the MV1 algorithm, the three MV3 sub-taggers are all based on multivariate methods themselves, here with a total of 27 inputs from the established b-taggers. 10 individual training runs for different p_{T} regions were performed for each sub-tagger, yielding an approximately 20 % better performance, in terms of the rejection factor, compared to the MV1 b-tagger. [95]

6. Event selection

Besides the general definitions, the selection cuts for the relevant physics objects have to be formulated. The aim is to maximise the signal significance value $S/\sqrt{S+B}$, which is accomplished in two steps. The pre-selections can be found in section 6.1, which are following the ATLAS recommendations for the top pair selection. Additional cuts for a better signal to background separation are described in section 6.2. Furthermore, the reconstruction procedure of top quarks and the associated W bosons is explained in section 6.3. All those analysis steps result in an improved signal-to-background ratio, as summarised in section 6.4.

6.1. Pre-selection cuts

The pre-selection cuts of the $|V_{ts}|$ analysis are identical to the ATLAS top working group recommendations of [67]. All objects used here are in agreement with the "Physics Object" definitions of chapter 5. To be accepted in the pre-selection the events have to pass all following cuts, one after the other. Some of those could easily be combined but for a better comparison with other analyses they are separated.

1. All events have to be listed in the **GRL** as detailed in section 4.4.
2. Events are rejected when arising from **noise bursts** in the liquid argon calorimeter **(LAr)** or dead cells. Events recorded within a fixed time window around these processes are also removed. Incomplete events due to the unavailability of detector information (LAr data integrity errors) and events affected by errors in the tile calorimeter cell system are removed as well.
   In addition, this cut includes a "Heavy Flavour Overlap Removal" (**HFOR**) procedure, rejecting heavy flavour MC events which appear multiple times in different samples. This prevents redundant MC simulations of the same event.
3. Events with electrons have to pass either the **trigger** "`EF_e24vhi_medium1`" or "`EF_e60_medium1`", events with muons either the trigger "`EF_mu24i_tight`" or "`EF_mu36_tight`". Details can be found in section 4.4.2.
4. A **Primary Vertex** (defined in section 5.5) per event is required to reject "non-collision background" processes which do not arise from collisions.
5. **At least one charged lepton** which passes the "tight++" (electron) or "tight" criteria (muon) has to be present, as described in sections 5.1 and 5.2. With this cut the dataset is split up in electron and muon flavour channels.

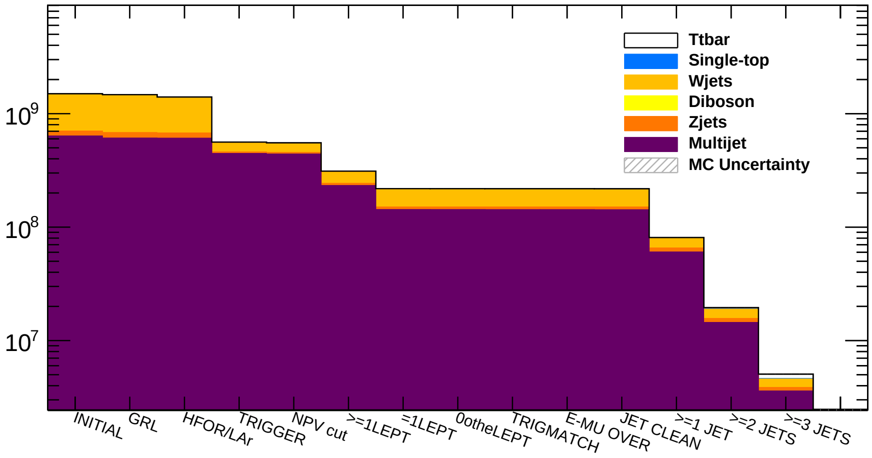


Figure 6.1.: Semi-logarithmic cutflow of all MC simulated samples and the data driven QCD multijet sample. These pre-selection cuts are applied to enhance the signal significance $S/\sqrt{S+B}$ and to reduce the background contributions.

6. **At the maximum one charged lepton** of the appropriate flavour channel is allowed, which passes the "tight++" (electron) or "tight" criteria (muon).
7. **No additional charged lepton** of a different flavour channel than from the appropriate one is allowed, which passes the "tight++" (electron) or "tight" criteria (muon).
8. The identified **charged lepton must match the triggered lepton**.
9. The event gets **rejected if any electron and muon overlap** in the inner detector i.e. sharing the same inner detector track.
10. If a **loose jet** (details in section 5.4) is present, the whole event gets rejected.
11. **At least 1 jet**, as defined in section 5.4, has to be reconstructed in the event.
12. **At least 2 jets**, as defined in section 5.4, have to be reconstructed in the event.
13. **At least 3 jets**, as defined in section 5.4, have to be reconstructed in the event.

Tau leptons are not explicitly included in this selection but those decaying into either electrons or muons are considered in the respective flavour channel.
All pre-selection cuts are summarised for the signal and all background samples in Figure 6.1. For comparison, the $t\bar{t}$ signal events only are illustrated in Figure 6.2. While these cuts reduce the quantity of all events by two orders of magnitude, the number of $t\bar{t}$ events decreases by only one order of magnitude.

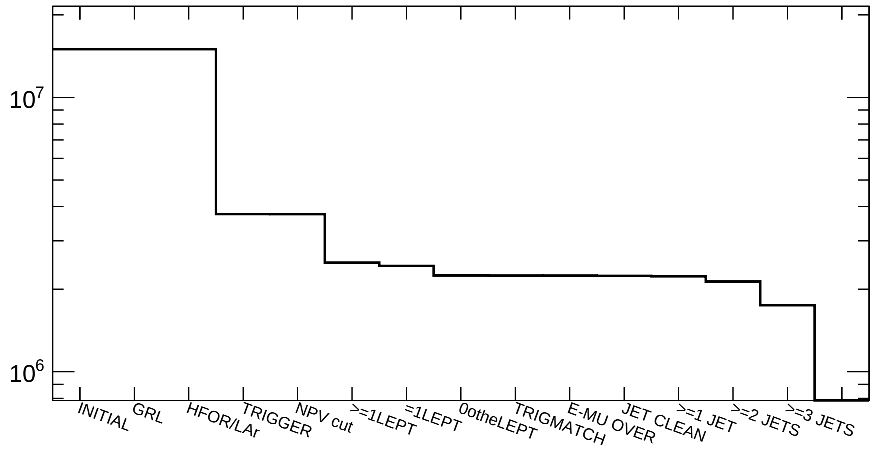


Figure 6.2.: This semi-logarithmic plot is similar to Figure 6.1, but illustrates the MC simulated $t\bar{t}$ events only. It shows the effects of the individual pre-selection cuts on the signal events.

6.2. Selection cuts

For a better differentiation between the $t\bar{t}$ events and the relevant background, further analysis specific cuts are applied. These are adjusted to the signal event topology $t\bar{t} \to WWq\bar{q}$ where the quarks are high energetic and the WW are decaying semileptonically into $qq + l\nu$. After the hadronisation, the entire final state at the lowest order consists of 1 charged lepton and a neutrino (stemming from one W boson) and 4 jets. Higher order perturbation calculations can then contribute more jets, while two additional jets are expected as a maximum, resulting in 6 jets in total. This composition is used to formulate the final selection:

- **4 to 6 jets**, as defined in section 5.4, must be reconstructed in the event to be accepted.

- **At least 30 GeV for the $\not{E}_T$** is required. This variable corresponds to the transverse neutrino energy as explained in section 5.3. The neutrino has to be relatively highly energetic since it carries a certain energy fraction from its massive mother particle, the W boson.

- **At least one b-jet** has to be identified by the tagging algorithm MV1 by using a working point corresponding to the b-tagging efficiency of 70 %.
  Theory predicts that about 99.8 % of all top quarks decay into a bottom quark and an associated W boson [44], while experimental results state a value larger than 95.0 % at 95 % confidence level [96]. These high percentages allow to ask for at least one b-tagged jet per event in the $|V_{ts}|$ analysis. The second t-daughter quark can then be probed for s-quark properties.

6.3. W boson and top pair reconstruction

After the selection cuts, the reconstruction of the W bosons and the top-quark pair is performed. A multipurpose kinematic fitting package, based on the χ^2 method is therefore implemented. The following section is based on [97, 98] in which a more detailed overview of the procedure can be found.

General method

The basic idea of the reconstruction method is to test an hypothesis, using the four-momenta and the covariances of the decay products. Additional assumptions are made for the invariant masses of the final state particles and for the particles which are part of the decay chain. All possible combinations of particle assignments to the mother particles are tested. The final choice of the correct assignment is then based on the corresponding χ^2 values. This procedure has three main advantages:

(i) It can reduce the physical and the combinatorial background due to the comparison of all possible combinations and the choice of the best χ^2.

(ii) It individually corrects the measured momenta of the particles in a defined range, resulting in recalibrated events with less detector and reconstruction effects.

(iii) It can reconstruct the missing properties of involved particles with high accuracy, like the z-component of the neutrino momentum.

The fundamental fitting procedure is based on a method that is explicitly described in [99]. It starts with several measured variables y_i, e.g. the four-momenta of all particles. Next, the algorithm tries to fit these variables to a stated hypothesis by introducing correction values Δy_i. The combination $y + \Delta y$ should then match the given constraints. The covariances of the variables specify the range, the fit is allowed to vary the corrections Δy_i. Minimising their summed value χ^2 is the aim of the reconstruction algorithm:

$$\chi^2 = \sum_{i=1}^{n} \frac{(\Delta y_i)^2}{\sigma_i^2}. \tag{6.1}$$

This sum is then used as the benchmark for the quality of the fit. To make sure that it reached its minimum, additional iterations are performed at the end of the fit as a cross check, that should only show small variations of χ^2. All described steps are part of the `KinFitter` library [100, 101], which is available in the `A++ framework` [102], used in the $|V_{ts}|$ analysis.

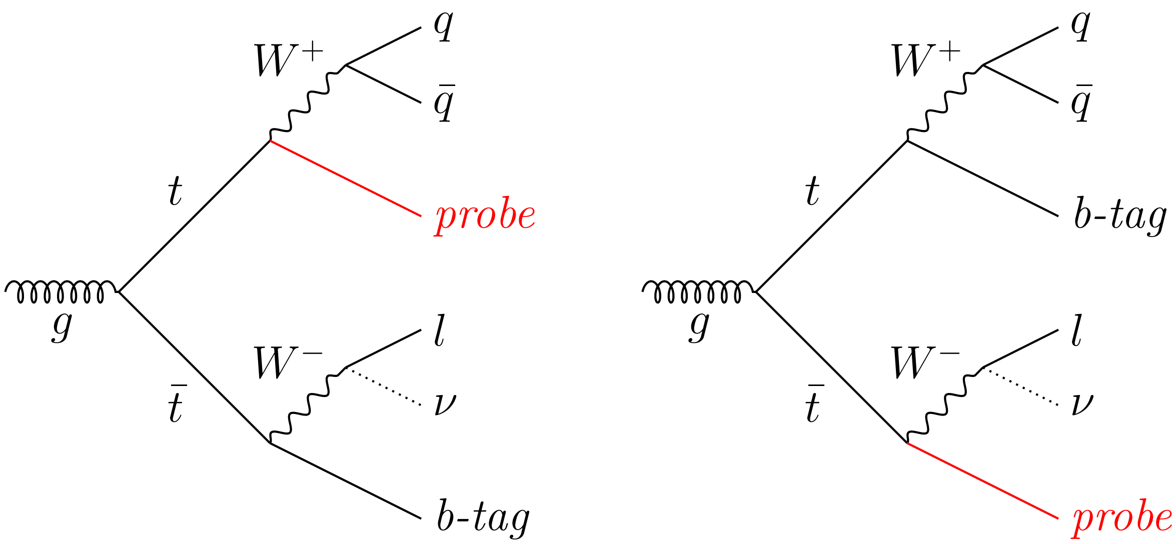


Figure 6.3.: Hypotheses used by the χ^2 fitter. Charge conjugation is implied.

Applied χ^2 fit

The reconstruction of the top pair and the associated W bosons is based on a kinematic χ^2 fit. The corresponding hypothesis states that $l+\nu$ as well as two light quarks should each stem from one W boson. Each W boson itself combined with an additional jet should form a top quark as illustrated in Figure 6.3. For a standard top analysis, these additional jets are expected to be b-jets, whereas the $|V_{ts}|$ analysis uses those for a tag and probe method: One of those top quark daughter jets has to be tagged as a b-quark jet while the other one is used as a probe, to perform the search for the process $t \rightarrow s + W$.
For the fitter itself, mass constraints and width constraints for the top quarks and the W bosons are set, where the mass distributions are supposed to be Gaussian [103]:

- Top quark mass: $m_{top} = 173.5$ GeV
- W boson mass: $m_W = 80.385$ GeV
- Top quark full width: $\Gamma_{top} = 2.0$ GeV
- W boson full width: $\Gamma_W = 2.085$ GeV

Furthermore, if the *probe jet* is not b-tagged, the mass of the s-quark, which is assumed to be close to zero, is used as additional constraint for that jet, whereas for a b-tagged *probe jet* the b-quark mass $m_b = 4.5$ GeV is set. All those constraints are treated as further measurements and their fit deviations are added to the χ^2 function as well. The invariant masses used are equivalent to the pole masses. The fitting procedure itself is then performed, using the available four-momenta of the final state particles.

Consequently, an individual χ^2 value is calculated for each combination. With four to six jets per event a huge number of combinations is possible, each representing a top pair candidate. Four jets yield $4! = 24$ permutations, for six jets the number rises to $6! = 720$ possibilities. This huge number of combinations needs to be reduced due to the CPU-intensive fitting algorithm. First, the above mentioned b-tagging information is taken into account. Second, an additional reduction is achieved by ignoring the succession of the two jets, associated to the W boson, which has no influence on the final χ^2 value.

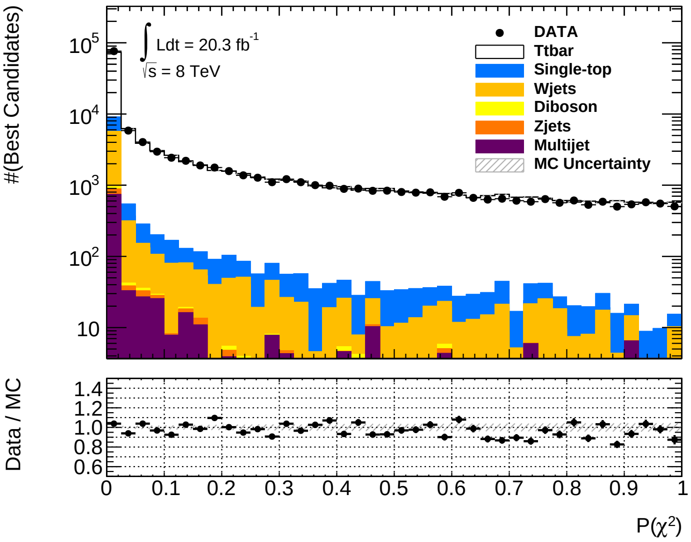


Figure 6.4.: Probability distribution of the $t\bar{t}$ reconstruction based on the χ^2 fitter, using `MC@NLO` generated events. Data events (black points) and Monte Carlo simulated events (coloured histograms) agree well.

The χ^2 values are then individually assigned to the χ^2 probability function $\mathcal{P}(\chi^2)$, also known as the p-value, which takes the number of degrees of freedom (d.f.) into account:

$$\mathcal{P}(\chi^2) = \int_{\chi^2}^{\infty} f_{\text{d.f.}}(x)dx \tag{6.2}$$

The corresponding distribution is illustrated in Figure 6.4. The combination with the highest probability $\mathcal{P}(\chi^2)$ is chosen to be the real $t\bar{t}$ process in the event. A non restrictive cut of $\mathcal{P}(\chi^2) = 10^{-10}$ is set, since this variable is integrated in the *boosted decision trees* (BDT) presented in chapter 8, which can find the best cut for a discrimination of b- from s-quarks more easily. Additionally, the maximum number of iterations is set to 100 which reduces the background significantly.

For a correct hypothesis, the resulting probability distribution is generally flat. Nevertheless, for the top pair assignment, it peaks at a low value (cf. Figure 6.4). This

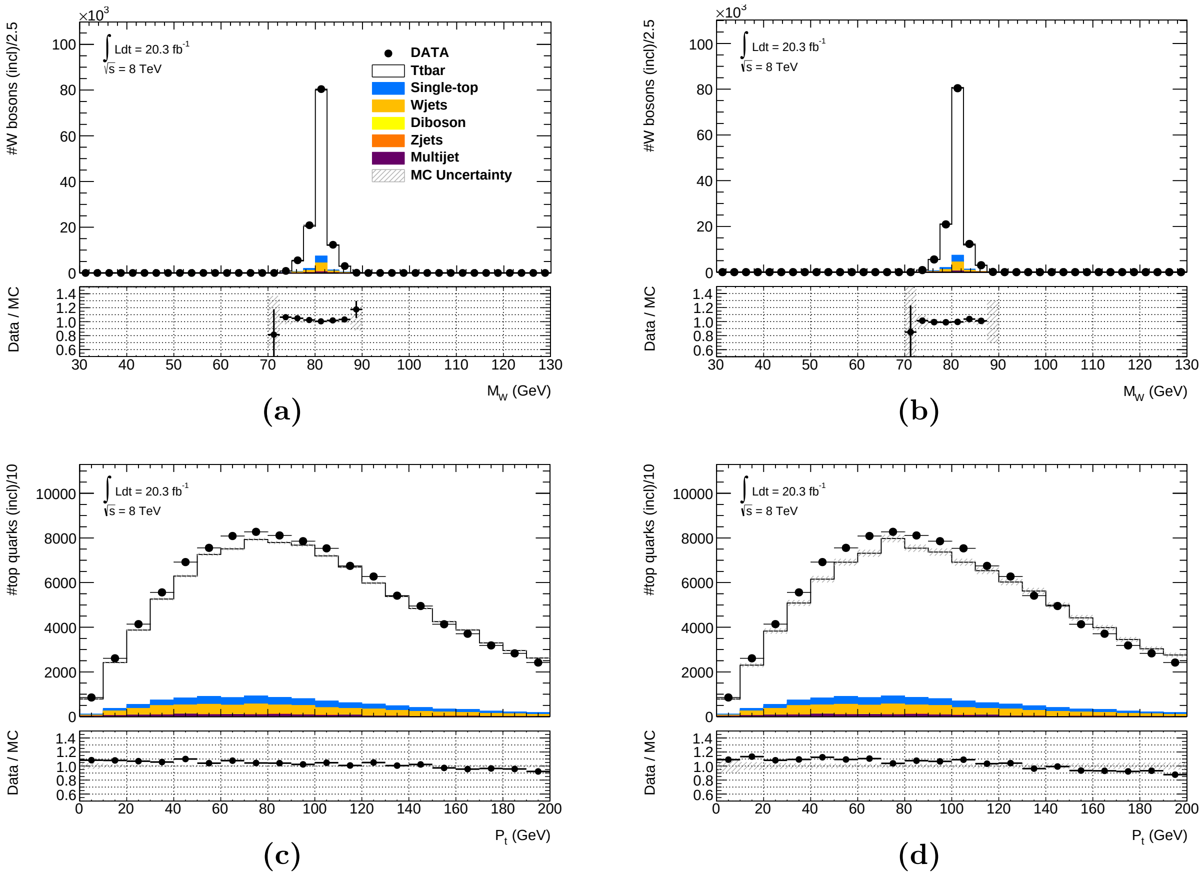
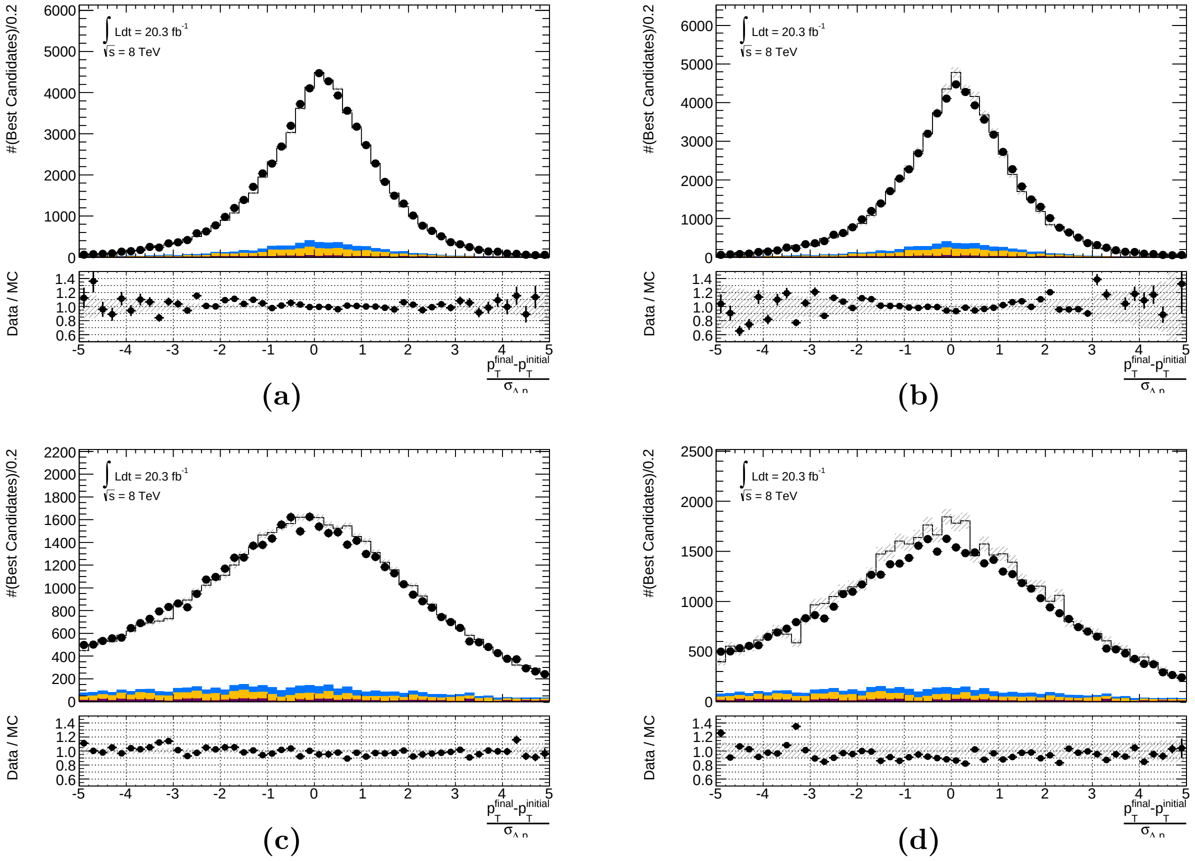


Figure 6.5.: (a,b): Mass distributions of the W bosons which are reconstructed with the χ^2 fitter. A good agreement between data and MC is observable. (c,d): Transverse momentum distributions of the top quarks, which are also reconstructed with the χ^2 fitter. Data and MC agree within 10 %, which is mostly compatible with the uncertainties (cf. section 9.1). The $t\bar{t}$ events were generated with `MC@NLO` (a,c) and `PROTOS` (b,d).

behaviour is caused by an improper assumption made for the difference *"measurement values − fitted values"*: The resulting distribution is described by a Gaussian, even though this is only true for its central part and not for the tails. Nonetheless, the top pair events can be reconstructed by the fitting method, as stated in [104], and as shown in the data-to-MC comparison plots of Figure 6.5.

Additional consistency checks with respect to the input variables are performed by considering the related pulls $p_i = \Delta y_i/\sigma_i$. For a perfect fitting procedure, the central peak would be at zero combined with a width of $\sigma_i = 1$. The results of this test are illustrated in Figure 6.6 (a,b) for the p_{T} of the b-quark on the leptonic side, and in Figure 6.6 (c,d) for the p_{T} of one jet arising from a W boson.

In summary, multiple tests indicate a high reliability of the whole kinematic fitter method [100, 105] and thus has been used in several ATLAS studies, e.g. [66, 104]. Moreover, it



(a) (b) (c) (d)

Figure 6.6.: (a,b): Pull distributions of the b-jets p_{T} which are associated to the leptonically decaying W boson. (c,d): Pull distributions of the p_{T} from the jets which are arising from the hadronically decaying W boson. The wide distribution is due to a lower p_{T} resolution of the W daughter particles, compared to leptonically decaying W bosons. The $t\bar{t}$ events were generated with `MC@NLO` (a,c) and `PROTOS` (b,d). The colour scheme is identical to Figure 6.5.

can also be applied for other purposes, like for the reconstruction of K_S^0 particles, which is presented in section 7.3.

6.4. Selection results and kinematic distributions

Following the reconstruction of the $t\bar{t}$ events, combined with the selection cuts of sections 6.1 and 6.2, an overview of the final signal event yields is given here. All uncertainties listed are due to MC and data statistics only, systematic uncertainties are discussed separately in section 9.1.

The numbers of signal and background events which pass all selections are summarised in Table 6.1, comprising the $t\bar{t}$ events simulated with the `MC@NLO` generator (cf. section 4.2). The number of expected events (MC simulations) is in good agreement with the data events. The resulting signal-to-background ratio (S/B) for the $t\bar{t}$ events is 89.98 %, while

Table 6.1.: Events passing the selection criteria and the reconstruction, including $t\bar{t}$ events generated by `MC@NLO`. The values are scaled to $\int \mathcal{L}\, dt = 20.3\, fb^{-1}$ and the uncertainties are statistically only.

Process	**Events**	
Top pairs: $t\bar{t} \to WWb\bar{s}$	183	± 10
Top pairs: $t\bar{t} \to WWb\bar{b}$	100464	± 238
Single Top	4471	± 50
W + Jets	5337	± 86
Diboson	83	± 5
Z + Jets	250	± 12
QCD multijet	1063	± 54
Theoretical expected (MC)	111851	± 185
DATA	112171	± 335

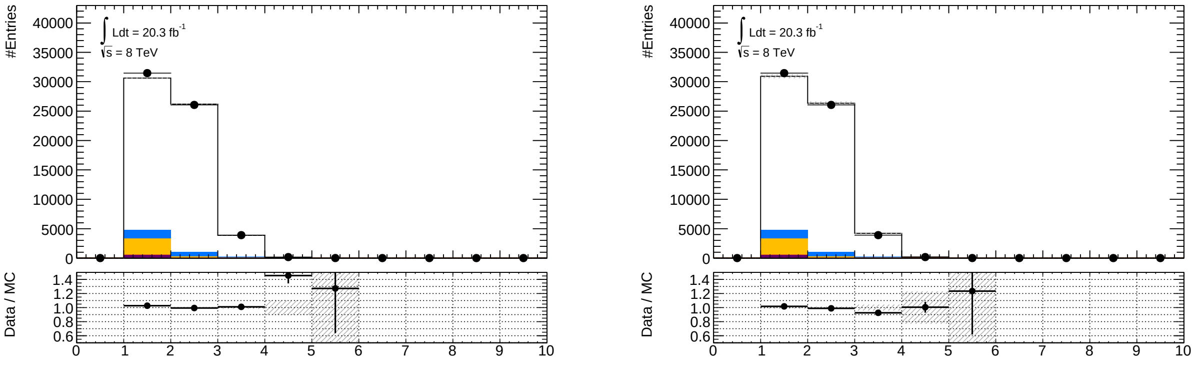
for the $t\bar{t}$ events with one top decaying to a s-quark, it is found to be 0.16 %. The signal significance ($S/\sqrt{S+B}$) for $t\bar{t} \to WWb\bar{s}$ is 0.55, with charge conjugations implied, but ignoring all uncertainties.

The $t\bar{t}$ events which were simulated with the `PROTOS` generator (details in section 4.1.1) have about 3.6 % more events than `MC@NLO`, as stated in Table 6.2. No calibration was applied at this stage, since the template fit in section 9.2 can correct for this. The total number of MC simulated events is also about 3 % higher compared to data. The signal-to-background ratio for all $t\bar{t}$ events is 90.30 %, while for $t\bar{t}$ events with one top decaying to a s-quark it is 0.19 %. The corresponding signal significance for $t\bar{t} \to WWb\bar{s}$ is found to be $S/\sqrt{S+B} = 0.66$ (ignoring all uncertainties).

Table 6.2.: Events passing the selection criteria and the reconstruction, including $t\bar{t}$ events generated by `PROTOS`. The values are scaled to $\int \mathcal{L}\, dt = 20.3\, fb^{-1}$ and the uncertainties are statistically only.

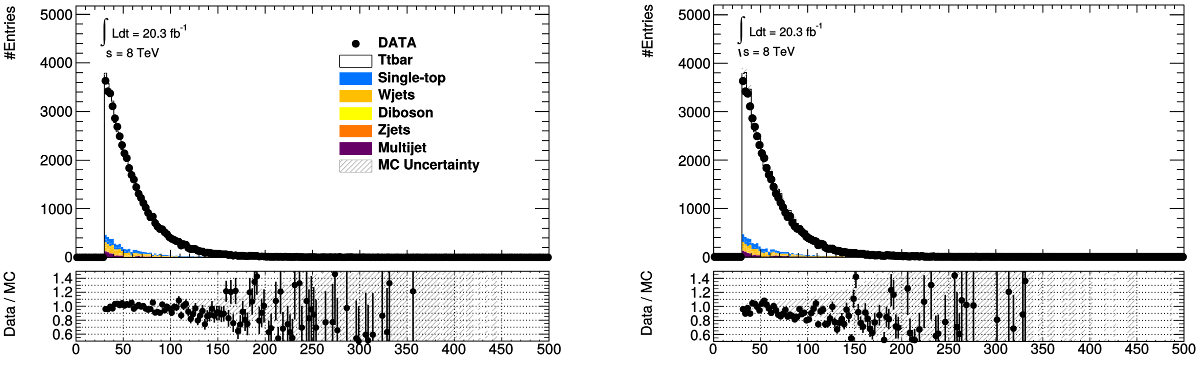
Process	**Events**	
Top pairs: $t\bar{t} \to WWb\bar{s}$	224	± 29
Top pairs: $t\bar{t} \to WWb\bar{b}$	104161	± 626
Single Top	4471	± 50
W + Jets	5337	± 86
Diboson	83	± 5
Z + Jets	250	± 12
QCD multijet	1063	± 54
Theoretical expected (MC)	115589	± 466
DATA	112171	± 335

The composition of the background with respect to `MC@NLO` $t\bar{t}$ events is as follows:
The remaining background events after all selections are clearly dominated by W+jets (4.8 %) and single top events (4.0 %). Further background processes contribute only with 1.0 % (QCD multijet), 0.2 % (Z + jets) and 0.1 % (Diboson).
To examine if also the distributions and shapes of important variables agree, plots comparing data and MC are shown in Figures 6.7 to 6.10. Diagrams on the left-hand side show `MC@NLO` $t\bar{t}$ events, plots on the right-hand side show `PROTOS` generated events. The variables displayed are the number of b-tagged jets, the electron p_{T}, the muon η and the $\not{E}_T$ distribution.



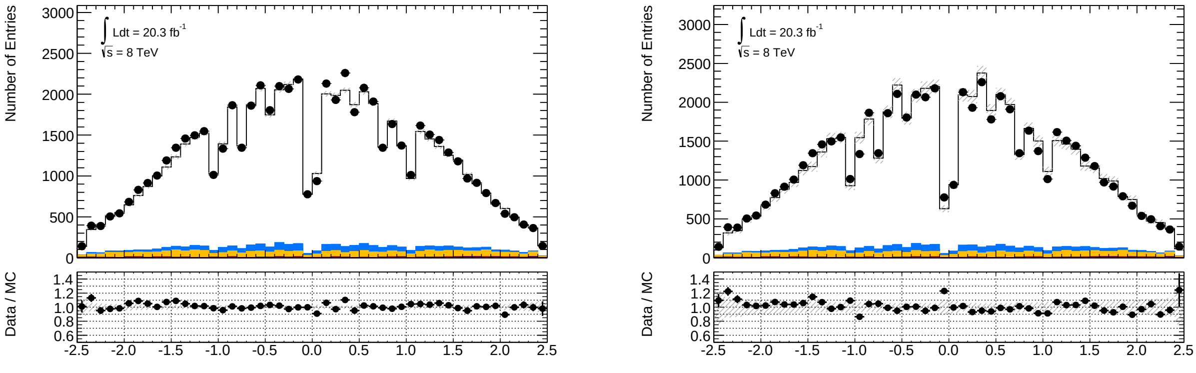
(a) (b)

Figure 6.7.: Data-to-MC comparison for the number of b-tagged jets ((a): `MC@NLO`, (b): `PROTOS`). A good agreement can be seen, especially in the important bins with one and two b-jets. The maximum number of all jets is 6 due to the event selection (section 6.2). The colour scheme is identical to Figure 6.8.



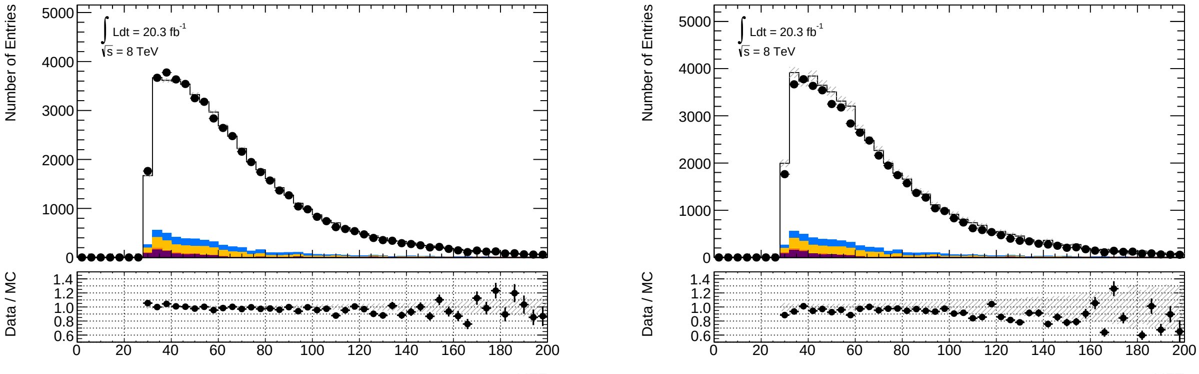
(a) (b)

Figure 6.8.: Data-to-MC comparison for the transverse momentum in GeV of the electrons which passed the selection and reconstruction steps. (a): `MC@NLO`, (b): `PROTOS`



(a) **(b)**

Figure 6.9.: Data-to-MC comparison for the η angle of the muons which passed the selection and reconstruction steps. The colour scheme is identical to Figure 6.8. (a): `MC@NLO`, (b): `PROTOS`



(a) **(b)**

Figure 6.10.: Data-to-MC comparison for the missing transverse energy $\not{E}_T$ in GeV indicated as MET in the plots. The colour scheme is identical to Figure 6.8. (a): `MC@NLO`, (b): `PROTOS`

7. Strange quark identification

The reconstruction of strange quarks originating from top quarks is the basis of this $|V_{ts}|$ analysis. Consequently, strange quarks need to be separated from the main background, i.e. from b-quarks. The major differentiating properties are summarised in the following:

- The invariant mass of b-quarks is much higher ($m_b = 4.2$ GeV) compared to s-quarks ($m_s = 0.1$ GeV) [1]. Thus, the invariant jet mass is influenced correspondingly.
- Both, b- and s-quarks hadronise before they decay. The resulting hadrons can directly be used for the discrimination.
- The mean lifetime of B hadrons is about 2 - 4 orders of magnitude shorter than of s-quark hadrons [1].
- The B hadron decay point can be identified as a Secondary Vertex close to the Primary Vertex and used as an additional discriminator. The decay point of s-quark hadrons is farer away and often outside the sensitive detector parts.
- The further decays of the associated hadrons are different. For B hadrons the decay topology follows $b \rightarrow (W+)\ c \rightarrow (W+)\ s$, whereas for s-quark hadrons the decay chain mainly ends after one cycle and is therefore much shorter.
- Massive hadrons, as those from b-quarks, transfer their total energy to a large number of daughter particles with a relatively small energy fraction. In comparison, daughter particles embodying the s-quark from $t \rightarrow s + W$ decays, like K_S^0 particles, carry a higher energy fraction. [106]
- The different decay chains influence the resulting jets and their cone size. Jets from b-quarks are expected to be wider compared to s-quark jets.
- The previous items combined, lead to a more spherical event topology for $t \rightarrow b + W$ decays than for $t \rightarrow s + W$ decays. [106]
- Some B hadrons decay in a characteristic semi-leptonic way, resulting in an high p_{T} electron or muon. [107] The identification of those charged leptons can be used for an improvement of the analysis.
- The s-quark hadrons from $t \rightarrow s + W$ decays, like K_S^0 particles, have the tendency to carry the quantum numbers of the mother quarks. [106]

Several of these discrimination aspects can strongly influence the properties of the *probe jets*, introduced in chapter 6. For a clean jet-to-quark assignment, a jet-parton-matching algorithm is implemented, described in section 7.1. The subsequent distinction of $t \to b + W$ and $t \to s + W$ decays is then split into different parts:
First, the 7 TeV `PROTOS` sample, which carries detector independent truth information (section 4.1.2), is used for a feasibility analysis, evaluating the prospects for a $|V_{ts}|$ determination (section 7.2). Second, the 8 TeV `PROTOS` dataset, which includes the simulated ATLAS detector and which is based on higher number of events, considers further possible discriminating variables. Thereby, K_S^0 reconstruction, based on jet tracks, is performed in section 7.3. Other variables, that are available in the 2012 dataset, are discussed in section 7.4. Finally, an overview of additional discrimination variables, which could not be integrated in the final $|V_{ts}|$ analysis, is given in section 7.5. All the following presented and available variables are considered for a multivariate analysis, namely a “boosted decision trees” method (BDT), which aims for an optimised separation of $t \to s + W$ and $t \to b + W$ decays in chapter 8.
Each data-to-MC comparison of this chapter is performed with `MC@NLO` generated $t\bar{t}$ events. For a better clarity, no MC uncertainties are drawn on discrimination diagrams when distributions are clearly separated and influences of the uncertainties can be neglected, i.e. the error bars are small. The below discussed $t \to s + W$ vs. $t \to b + W$ discrimination studies are based on MC models only, since experimental datasets enriched by high energetic s-quarks, which could be used for a description of $t \to s + W$ decays, are not available.

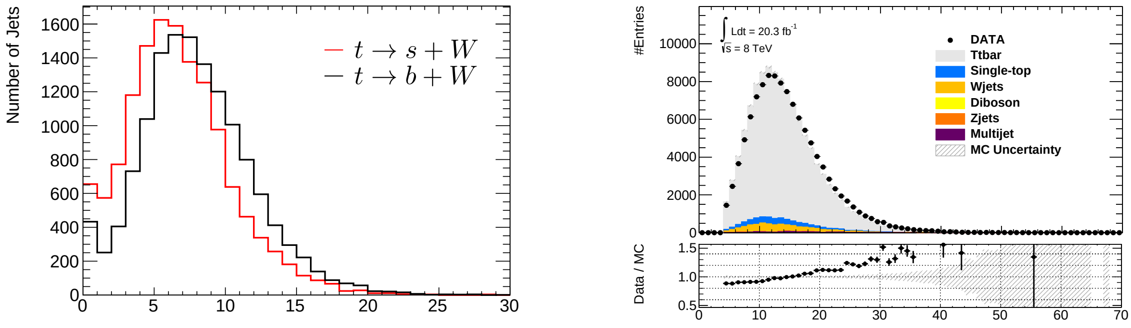
7.1. Jet parton matching procedure

Hadronisation is not fully described by perturbative theory and has to be modelled by different phenomenological approaches (cf. section 2.2). In these models, the clear assignment between jets and the corresponding mother quarks gets lost. To still have these information available for further studies, an algorithm performs a jet-quark matching afterwards, based on geometrical and kinematic decisions.
In the $|V_{ts}|$ analysis, this algorithm uses truth information to only select quarks which are arising from top quarks. Jets and quarks are tagged as “matched” if the kinematic condition satisfies $0.5 < p_{\mathrm{T,\,jet}}/p_{\mathrm{T,\,quark}} < 1.5$ as well as the spatial condition $\Delta R < 0.4$, describing the spherical distance in (ϕ, η) between quarks and jets. A study to optimise the cut on the latter parameter compared results for $\Delta R = 0.25$, 0.4 and 0.5. Here, $\Delta R = 0.4$ yields the best performance with a relatively low number of multiple assignments and a relatively low number of quarks with no assignments at all. The corresponding plots are depicted in section A.3.
Due to additional processes, like ISR/FSR, a quark can be associated to multiple jets, which comply with the matching requirements. To still have a unique assignment, the jet with the smallest ΔR is chosen as the final one. This simple procedure can be applied since at high energies, the flight path of quarks and the corresponding jet axis are relatively close to each other due to the following facts: When the energy of the jet

rises, the opening angle shrinks proportional to $1/\gamma = \sqrt{1-\beta^2}$ [108]. Simultaneously, the corresponding energy per particle increases since the number of particles within the jet increases only logarithmically, with $<N> \sim \log(s)$ [109]. Thus, the hadronisation processes, which could lead in a divergence between the jet axis and the quark's flight path, become more and more irrelevant.

7.2. Fundamental quark differences based on truth study

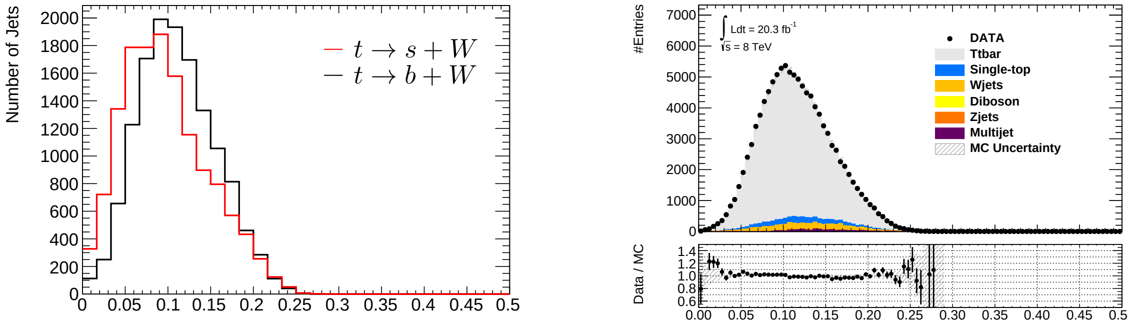
The first study, which is based on the $\sqrt{s} = 7$ TeV `PROTOS` dataset (cf. chapter 4), uses detector independent truth parameters for a feasibility analysis in order to discriminate $t \to b + W$ decays from $t \to s + W$ decays by considering fundamental jet properties, i.e. the jet transverse momentum, the jet size and the number of tracks per jet.



(a) **(b)**

Figure 7.1.: (a) Track multiplicity in jets stemming from $t \to s + W$ decays in red and for jets stemming from $t \to b + W$ in black are shown for `PROTOS` generated $\sqrt{s} = 7$ TeV events. The mean value of the track multiplicity for s- and b-jets is 6.21 and 7.52 with a corresponding RMS of 3.68 and 3.89, respectively, revealing a difference of about 20 %. (b) The corresponding data-to-MC comparison of the *probe jet* (cf. section 6.3) is based on the 8 TeV dataset using the `MC@NLO` generator. The visible discrepancy between data and MC events indicates to omit this variable in the final BDT analysis of chapter 8. The related `PROTOS` based 8 TeV discrimination plot for s- and b-quarks can be found in the supplemental Figure A.8.

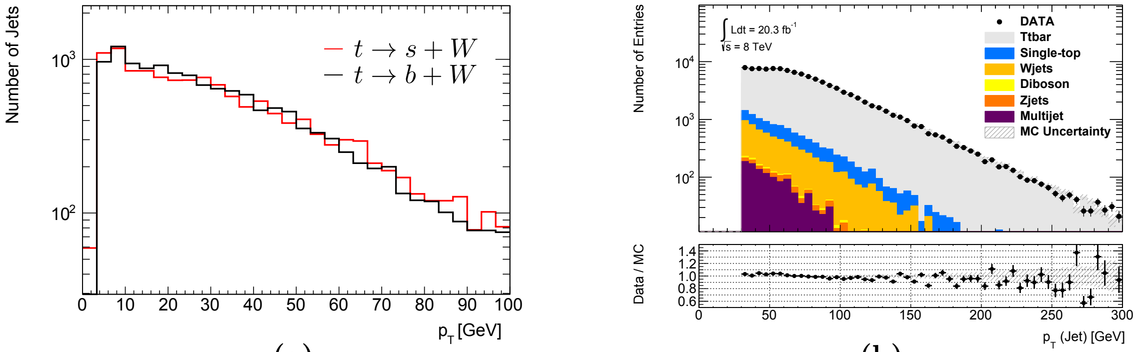
The large invariant mass of b-quarks strongly influences these parameters since high masses affect additional high energetic gluon radiations, which is explicitly discussed for 1st order calculations in [110] and for 2nd order calculations in [111]. Consequently, the track multiplicity of b-jets clearly differs from those in s-jets: The number of tracks in b-jets has a mean value of 7.52 with a RMS of 3.89, while the number of tracks in s-jets results in a mean value of 6.21 with a RMS of 3.68 (Figure 7.1 (a)). These numbers reveal a high discrimination power. However, a data-to-MC comparison, performed for the final $\sqrt{s} = 8$ TeV analysis, shows no satisfying agreement (Figure 7.1 (b)), which indicates to omit this variable in the final BDT method of chapter 8.



(a) **(b)**

Figure 7.2.: (a) The jet cone size of b-jets (black) is found to have a mean value of 0.111 and a corresponding RMS of 0.044. The s-jets (red) instead have a mean value of 0.097 and a RMS of 0.049. The corresponding difference is about 15 %. (b) The data-to-MC comparison of the *probe jet*, based on the 8 TeV `MC@NLO` dataset, correlates well. The corresponding 8 TeV s-/b-quark discrimination plot can be found in Figure A.9.

In contrast, the jets' cone width r_{cone} agrees quite nicely between data and MC events (Figure 7.2 (b)). It is also affected by the additional gluons, thus, the b-jets have a larger mean width of 0.111 and a RMS of 0.044, while the s-jets have a mean width of 0.097 and a RMS of 0.049 (Figure 7.2 (a)).



(a) **(b)**

Figure 7.3.: (a) Jet p_{T} for jets stemming from $t \to s + W$ in red and for jets stemming from $t \to b + W$ in black are shown, based on `PROTOS` generated 7 TeV events. By an exponential fit on the different distributions, a difference of about 6 % is found for the slopes. b) The data-to-MC comparison of the *probe jet*, based on the 8 TeV `MC@NLO` dataset, indicates a good agreement. Here, the x-axis has a p_{T} range of 0 - 300 GeV. The corresponding 8 TeV s-/b-quark discrimination plot can be found in Figure A.10.

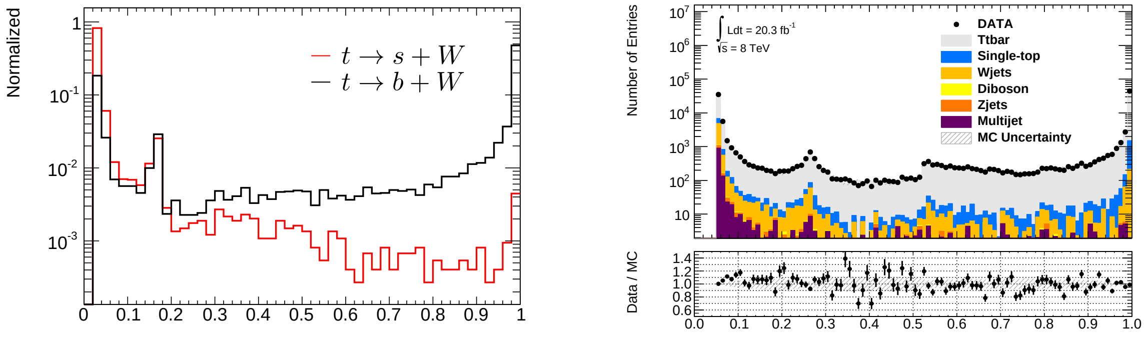
The transverse momenta of the jets also indicate an inherent difference between $t \to s + W$ and $t \to b + W$ decays. In the semi-logarithmic plot of Figure 7.3 (a) an excess of s-quark jets at high p_{T} values seems to be present. This impression is supported by an exponential fit, resulting in the fitting parameters $(-2.85 \pm 0.03) \cdot 10^{-5}$

for the b-jets and $(-2.68 \pm 0.05) \cdot 10^{-5}$ for the s-jets, which is equivalent to the slopes in the semi-logarithmic plot. The corresponding data-to-MC comparison on the right side (Figure 7.3 (b)) illustrates a good agreement.

Apart from these general jet properties, all variables introduced for b-tagging algorithms in section 5.6 can improve the s-/b-quark differentiation for the $|V_{ts}|$ analysis. This includes the impact parameter significances d_0/σ_{d_0} and z_0/σ_{z_0} of the jet tracks, the decay length significance of the B hadrons L/σ as well as the properties of the reconstructed Secondary Vertex, like its invariant mass, the number of vertex candidates and the energy ratio of the associated tracks. Furthermore, the reconstructed flight path of the B hadron, as used in the JetFitter algorithm, can be used. The corresponding b-tagging algorithms and their s- vs. b-quark separation efficiency can be found in Figure 5.10 of section 5.6.

A combination of three basic b-tagging algorithms, including V0 particle rejection, results in the multivariate tagger MV1 for which a detailed explanation is given in section 5.6. Its high discrimination power, which is illustrated in Figure 7.4 (a), can be used for the analysis in hand, especially since a nice data-MC agreement is achieved as depicted in Figure 7.4 (b).

To support the statement of a high discrimination power, truth based calculations are performed in the following, which show the theoretical ability to reconstruct the relative BR $t \to s + W$ and $t \to b + W$ using the MV1 variable only. This relies on the 7 TeV `PROTOS` dataset with enhanced BR.



(a) **(b)**

Figure 7.4.: (a) MV1 distribution for *probe jets* stemming from $t \to s + W$ decays (red) and $t \to b + W$ decays (black). A strong separation power is indicated which can be used for the calculation below (Equation 7.1 ff.) and for the BDT analysis of chapter 8. (b) The corresponding data-to-MC distribution, based on the 8 TeV dataset using the `MC@NLO` generator, indicates a nice agreement.

For these calculations, a variable x is introduced, representing the fraction of events in which both top quarks decay into b-quarks. Combined with the assumption that

the decay process $t \to d + W$ happens equally often as the $t \to s + W$ process[1], the individual BR in question can be determined.
For the corresponding computation of the variable x, the b-tagging outputs listed in Table 7.1 are used which stem from a pure $t\bar{t}$ MC sample. They are reconstructed with a b-tagging efficiency of $\epsilon = 70$ % which equals a MV1 weight of 0.6017. This value also implies a light-jet rejection of 130 and a charm-jet rejection of 5, which means that one out of 130 light-jets (u-, d- and s-quarks) and one out of 5 charm-jets is misidentified as a b-jet.

Table 7.1.: Number of events with different b-jets multiplicity used in Equation 7.1 ff.

0 b-tags	**1 b-tag**	**2 b-tags**	**multiple b-tags**
6490	7004	2499	309

With the b-jet multiplicities from Table 7.1, several independent formulas (Equation 7.1 ff.) can be formulated, each representing a different number of b-jets per event. Here, the following conventions are used:

$\epsilon = 70\%$	b-tagging efficiency of the MV1 tagger
$R_l = 1/130$	Mistagging rate for light-quark jets
$R_c = 1/5$	Mistagging rate for charm-quark jets
l	**Unknown** number of light-quark jets per event
c	**Unknown** number of c-quark jets per event
x	**Unknown** BR: $t \to b + W$

$$\frac{N_{\text{Events with 0 b-tags}}}{N_{\text{All events}}} = ((1-\epsilon)x + (1-x))^2 \cdot (1-R_l)^l \cdot (1-R_c)^c \tag{7.1}$$

---

[1] This assumption for the `PROTOS` dataset with enhanced BR is equivalent to the assumption of a negligible amount of $t \to d + W$ decays in datasets with SM BR, as it is supposed in the final $|V_{ts}|$ analysis.

$$\begin{aligned}\frac{N_{\text{Events with 1 b-tag}}}{N_{\text{All events}}} &= 2\left(\epsilon\cdot x\right)\left(\left(1-\epsilon\right)x+\left(1-x\right)\right)\cdot\left(1-R_l\right)^l\cdot\left(1-R_c\right)^c \\ &+\left(\left(1-\epsilon\right)x+\left(1-x\right)\right)^2\Big(l\cdot R_l\left(1-R_l\right)^{l-1} \\ &\cdot\left(1-R_c\right)^c+\left(1-R_l\right)^l\cdot c\cdot R_c\left(1-R_c\right)^{c-1}\Big)\end{aligned}\tag{7.2}$$

$$\begin{aligned}\frac{N_{\text{Events with 2 b-tags}}}{N_{\text{All events}}} &= \left(\epsilon\cdot x\right)^2\cdot l\left(1-R_l\right)\cdot c\left(1-R_c\right) \\ &+2\left(\epsilon\cdot x\right)\left(\left(1-\epsilon\right)x+\left(1-x\right)\right)\cdot\left(\left(R_l l\left(1-R_c\right)c\right)+\left(1-R_l\right)lR_c c\right) \\ &+\left(\left(1-\epsilon\right)\cdot x+\left(1-x\right)\right)^2\cdot\Big(R_l^2\left(1-R_l\right)^{l-2}\cdot\left(1-R_c\right)^c+R_l\left(1-R_l\right)^{l-1} \\ &\cdot R_c\left(1-R_c\right)^{c-1}+\left(1-R_l\right)^l\cdot R_c^2\left(1-R_c\right)^{c-2}\Big)\end{aligned}\tag{7.3}$$

The equations for multiple b-tags per event follow along the same line. These terms only arise due to additional misidentified light- or c-quark jets.

Combining all independent formulas gives rise to an overdetermined system which yields $x = 49.71\ \% \pm 0.66\ \%$, representing the BR $t \to b + W$. The corresponding BR $t \to s + W$ and $t \to d + W$ result each in $25.14\ \% \pm 0.33\ \%$. Those three values coincide with the BR of the input MC sample as summarised in Table 7.2.

Table 7.2.: With the help of the MV1 b-tagging algorithm a reconstruction of the $t\bar{t}$ BR can be achieved. As input, the `PROTOS` MC sample with enhanced BR (explained in section 4.1.1) is used.

	$t \to b + W$	$t \to s + W$	$t \to d + W$
Expected from MC	50 %	25 %	25 %
Calculated result	$49.71\ \% \pm 0.66\ \%$	$25.14\ \% \pm 0.33\ \%$	$25.14\ \% \pm 0.33\ \%$

With this simple approach, an adequate reconstruction of the BR can be performed. Consequently, the MV1 b-tagger seems to be a good choice for the main discriminator in the following chapters. Next to the use as a classical b-tagger, it is also applicable as an anti-b-tagger, rejecting b-quarks while s-quarks stemming from top quarks are allowed to pass. For that, the cut values are inverted compared to the regular b-tagging method.

Other, more exotic tagging algorithms, like e.g. the SV2 algorithm, could theoretically improve the total discrimination power but due to the high efficiency of MV1 and the similarity to its three input taggers, all other b-taggers are neglected in the $|V_{ts}|$ analysis at hand.

7.3. K_S^0 reconstruction

Most b-taggers comprise a V0 particle rejection method (cf. section 5.6) in order to reject s-jets which could fake b-jets due to similar jet kinematics. This principle can be adopted and used the other way round to identify $t \to s + W$ decays. A V0 reconstruction algorithm, based on Secondary Vertices, is basically available in the ATLAS framework, but large data-MC discrepancies of about 30 - 50 %, as depicted in Figure A.11, forbid the utilisation in this analysis. Therefore, an explicit reconstruction of V0 particles - more precisely of K_S^0 particles[2] - is conducted, based on the identification of two pion tracks in the *probe jet*.

For the reconstruction itself, two main problematic aspects have to be kept in mind to achieve a clean identification of s-quarks stemming from top quarks:
First, s-quarks and thus K_S^0 particles can in principle also arise from sea quarks. About 30 % of all sea quarks are s-quarks [112], the rest is dominated by down- and up-quarks. Second, K_S^0 might also result from b-quarks, following the decay chain $b \to (W+)\ c \to W + s$. An isolated identification of the additional W bosons is in general not possible since the angular distance to the mother jet is expected to be small. Hence, the W bosons are reconstructed as part of the final jet.

On the positive side, K_S^0 particles are expected to be high energetic if they arise from $t \to s + W$ decays. This gives the advantage that they clearly differ from K_S^0 which either arise from the sea quarks or from the b-quarks' decay chain. The latter ones are expected to have much lower momenta due to the additional emergences of W bosons. Thus, the utilisation of reconstructed K_S^0 particles can be a good way to support the identification of $t \to s + W$ decays.

Due to the high momenta of K_S^0 particles, the decay length is extended by the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{E}{E_0}, \tag{7.4}$$

resulting in an energy dependent decay length of $\gamma \cdot \tau c = \gamma \cdot 2.7$ cm. Owing to the high transverse momenta of the s-quarks in $t \to s + W$ decays, the p_T fraction X of the K_S^0, with respect to the s-quark p_T, must not be too large. Otherwise the two daughter pions are not reconstructable within the Inner Detector because of the relative small radius of 115 cm (cf. section 3.2.2).
Fortunately, with increasing p_T fraction X, the number of hadrons decreases (more

---

[2] With small contributions from $\Lambda \to p\pi$

than) exponentially [113, 114]. Thus, the majority of the K_S^0 particles is expected to carry less than 25 % of the mothers' p_T. This value combined with an assumed transverse momentum of $p_\mathrm{T} \approx 80$ GeV for s-quarks would result in 20 GeV K_S^0 particles with $\gamma \approx 40$ and a mean decay length of 107.4 cm. Consequently, most K_S^0 decay points and thus the daughter particles are still reconstructable. This is true especially when taking into account that the lifetime is determined by an exponential decay, meaning that even high energetic particles can decay much earlier. Nevertheless, very high energetic K_S^0 might be missed by that reconstruction algorithm. [1, 106]

Those very high energetic K_S^0 particles could be identified by energy loss and shower shape measurements in the hadronic calorimeter. However, the corresponding measurement steps have not been integrated in the $|V_{ts}|$ analysis since this method demands additional studies in the high energy region, like surveys of misidentified neutrons and of extra pileup contributions.

If that implementation would be available, K_L^0 particles could also be identified, which would be a further advantage, since a direct reconstruction of the K_L^0 decay process is not possible in the ATLAS detector. This is due to the fact, that the mean lifetime of K_L^0 is about 500 times longer than K_S^0, resulting in a mean decay length of about 10 times the diameter of the whole ATLAS detector, using the exemplary p_T values from above. Charged kaons could improve the $|V_{ts}|$ analysis as well but their long lifetime (same order of magnitude as K_L^0) also prohibits an identification of the full decay process. Thus, the only way would again be a particle identification based on information from the calorimeter and - because of the charge - from the ATLAS tracking system. But due to the high kaon momenta occurring in $t \to s + W$ decays, the discrimination power between charged kaons and protons is very small which makes a separation difficult [115].

To summarise, only K_S^0 particles with a p_T smaller than 20 GeV and two charged pions as daughter particles, can be considered in the $|V_{ts}|$ analysis. Since only about half of all s-quarks hadronise into neutral kaons, whereof only about 50 % are K_S^0 particles, whereof about 70 % decay into charged pions, a maximum fraction of $\sim$17.5 % of all s-quarks are reconstructable with that method [103]. Λ hadrons are not explicitly taken into account since only a small fraction of s-quarks hadronise into these particles. Nonetheless, the predominant decay $\Lambda \to p\pi^-$ has a similar topology like K_S^0 decays if the proton is misidentified as another pion. Therefore, they are partly included in the K_S^0 reconstruction method since no explicit rejection cut is set.

The expected differences between K_S^0 stemming from $t \to s + W$ decays and those from $t \to b + W$ decays are depicted in Figures 7.5 and 7.6 for truth K_S^0 particles, with a maximum distance of $\Delta R(K_S^0, \text{jet axis}) = 0.3$ to the jet. Therein, it is clearly visible that K_S^0 from s-quarks have a higher p_T than those from b-quarks. For the ratio $p_\mathrm{T}(K_S^0)/p_\mathrm{T}(\text{jet})$ an even cleaner discrimination can be achieved which already emerges at small values where there are less entries from $t \to s + W$. With raising values the s-quark related K_S^0 particles become dominant. In both illustrations (7.5, 7.6), the left diagram is plotted with a linear scale whereas the right diagram is semi-logarithmic. The corresponding p_T and p_T-ratio distributions for Λ truth particles are depicted in

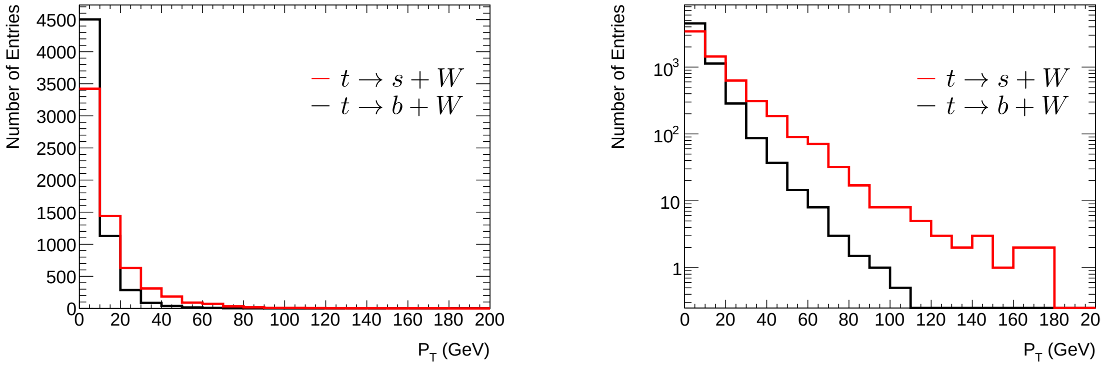


Figure 7.5.: Transverse momentum distribution for K_S^0 truth particles stemming from $t \to s + W$ decays (red) and from $t \to b + W$ decays (black), based on the 8 TeV `PROTOS` dataset. The right-hand diagram shows the same distribution as the left but in semi-logarithmic scale.

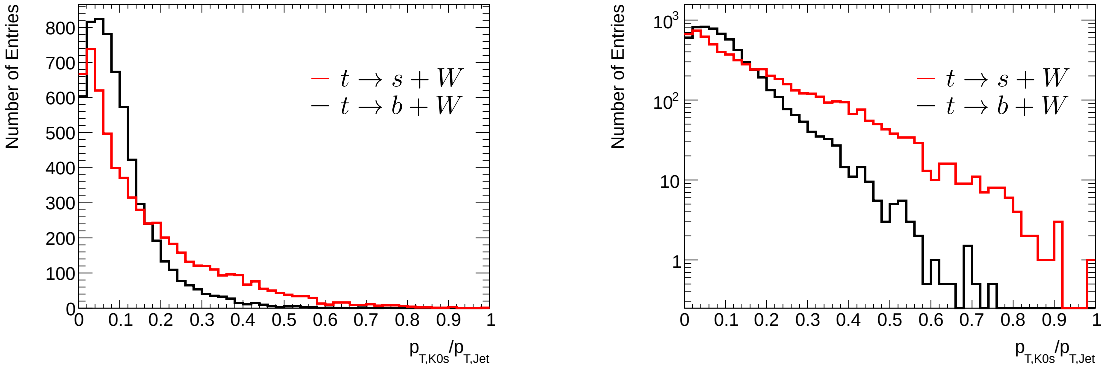
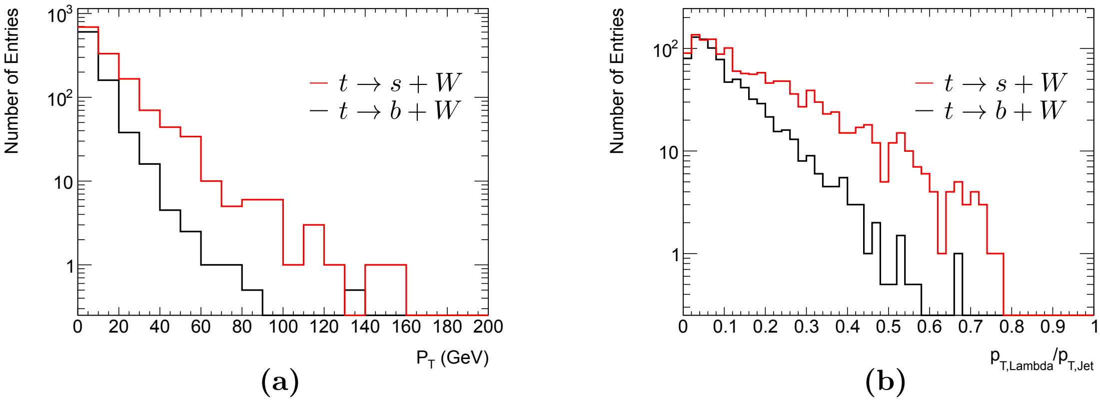


Figure 7.6.: The ratio between the transverse momenta of truth K_S^0 particles and the regarding jets is illustrated. K_S^0 truth particles stemming from $t \to s + W$ decays are drawn in red, those from $t \to b + W$ decays in black. The right-hand diagram is identically to the left-hand diagram apart from a semi-logarithmical scale. K_S^0 particles from s-quarks carry in general a higher p_{T} ratio of the corresponding jet, i.e. from the mother particle.

Figure 7.7, both in semi-logarithmic scale. The K_S^0 properties stated for the Figures 7.5 and 7.6 also apply for the Λ distributions. No data-MC comparison plots are drawn since these plots are based on truth information only.



(a) **(b)**

Figure 7.7.: (a) The transverse momenta of Λ truth particles and (b) the ratio of the Λ p_{T} vs. the jet p_{T} is shown, both based on the 8 TeV `PROTOS` dataset. Λ truth particles stemming from $t \to s + W$ decays are drawn in red, those from $t \to b + W$ in black. Both diagrams are drawn in semi-logarithmic scale. The Λ distributions show a similar behaviour as the K^0_S distributions from Figures 7.5 and 7.6.

7.3.1. Reconstruction procedure of K^0_S

General method

The reconstruction of a neutral particle, decaying into two daughter tracks, are usually based on algorithms similar to the `KinFitter` algorithm [100] of the top pair finder in section 6.3. In general the kinematic fit requires an hypothesis and results in several selection steps, which are summarised below, following the procedure of [116]:

- The **tracks** selected must have a fixed **minimum p_{T}** which depends on the mother particle's properties.
- All **tracks** have to be within in a specific **acceptance window of η**.
- The **tracks' reconstruction** algorithm must match a minimum χ^2 value.
- A central **fit is performed** which has to satisfy the constraints (e.g. invariant mass of the mother particle) and has to converge within a maximum number of iterations.
- The consequent **pointing angle** (Figure 7.8 (a)) has to pass a specified cut.
- The consequent **decay angle** (Figure 7.8 (b)) has to pass a specified cut.
- The consequent **opening angle** (Figure 7.8 (c)) has to pass a specified cut.
- The **p-value** of the kinematic fit has to be larger than a specified value.
- Since the mother particle is neutral, the **charge of the two tracks** must be oppositely signed.

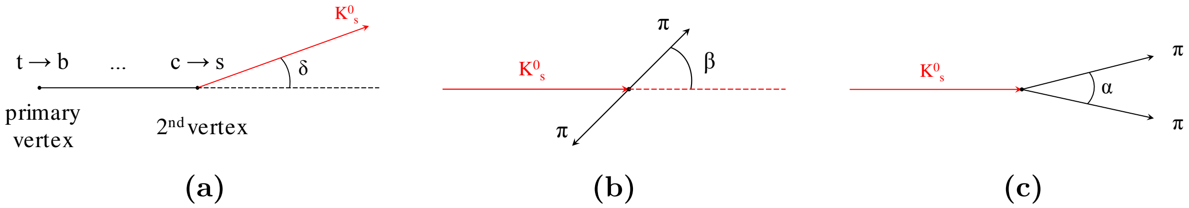


Figure 7.8.: Three different angles which are used in the reconstruction procedure of a neutral particle, e.g. K^0_S, decaying into two particles: (a) δ describes the pointing angle between the original flight path and the path of the particle in question (laboratory system), (b) β is the decay angle in the rest frame of the mother particle, (c) α is the opening angle between the two daughter particles (laboratory system).

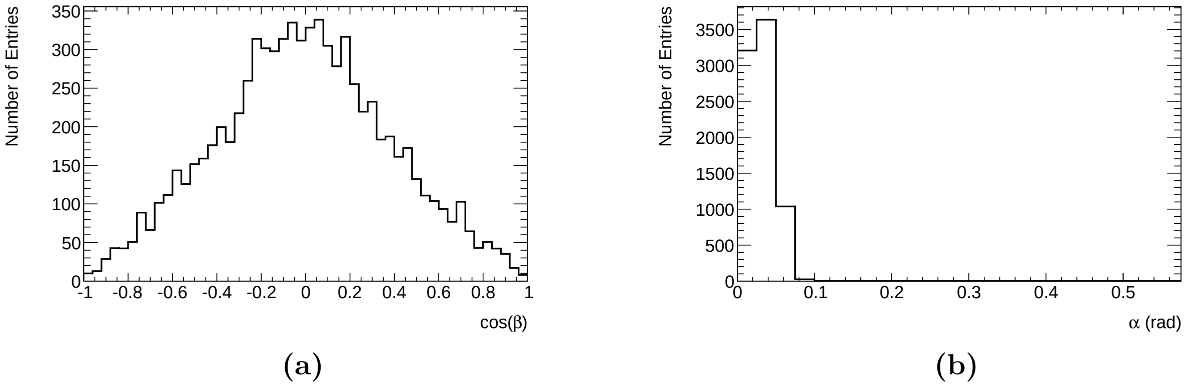


Figure 7.9.: (a) Cosine of decay angle β for the two pions in the K^0_S rest frame. (b) Opening angle α in the laboratory frame. The distributions are based on $t\bar{t}$ MC samples of `PROTOS` where the K^0_S is reconstructed from tracks in *probe jet*. A detailed explanation of the angles can be found in Figure 7.8.

As an output of these steps, the reconstructed particle candidates are obtained, each associated with a χ^2 value describing the quality of the fit which can be translated into the corresponding p-values (cf. Equation 6.2).

Chosen method for K^0_S reconstruction

In the specific reconstruction of $K^0_S \rightarrow \pi^+\pi^-$ decays, with K^0_S originating from $t \rightarrow s + W$ or $t \rightarrow b + W$ processes, most steps are identical to the general method described above, including the `KinFitter` algorithm from section 6.3. Solely the three cuts on the angles of Figure 7.8 can not be accomplished due to the following reasons: First, the pointing angle δ can not be determined since no Secondary Vertex is expected. Second, for the decay angle β an isotropic distribution is anticipated since there is no preferred decay direction in the rest frame of the K^0_S, thus no cut is applicable.

Third, the opening angle α would allow a selection but a cut on the K_S^0 mass window afterwards is equivalent and can be integrated in the BDT analysis of chapter 8. The corresponding MC distributions for the angles β and α are depicted in Figure 7.9.

For an optimisation of the remaining reconstruction steps in the K_S^0 identification, the track p_{T}, the track quality values (χ^2) and the p-values of the final K_S^0 are varied. In total, 15 different combinations are tested and optimised in order to find the best agreement with truth K_S^0 and in order to achieve the best discrimination between K_S^0 from $t \to s + W$ and $t \to b + W$ decays. The minimum and maximum values of the surveyed parameters are summarised in Table 7.3. A selection of these combinations are depicted in the figures of section A.4.

Table 7.3.: Summary of the 3 parameters which were varied for an optimisation of the K_S^0 reconstruction. In total 15 different combinations were tested to find the best composition.

	Minimum value	Maximum value
Minimum p_{T} of the tracks (GeV)	5	20
Maximum $\chi^2/n_{d.f.}$ of the tracks	0	2.5
Minimum `KinFitter` p-value	0	0.05

As a result, the p_{T} of the input tracks have to be larger than 10 GeV. Furthermore, no cut on the quality value χ^2 of the tracks is set because all tracks seemed to be highly reliable since quality criteria are already implemented implicitly in the reconstruction procedure. Finally, no clear indication of an improvement of the s-quark vs. b-quark discrimination is visible when adjusting the final p-value. Thus, no fixed cut is set on this parameter, even though, it can be used as an additional input for the BDT.

All consequent reconstruction steps for the K_S^0 identification are summarised in Figure 7.10 and below:

- **All possible track combinations** within $\Delta R = 0.6$ around the *probe jet* are considered[3].
- The **track p_{T}** must be larger than **10 GeV**.
- The tracks must lie in the sensitive detector region $|\boldsymbol{\eta_{trk}}| < \mathbf{2.5}$.

---

[3] In general, particle reconstructions are based on tracks which are associated to the SV.

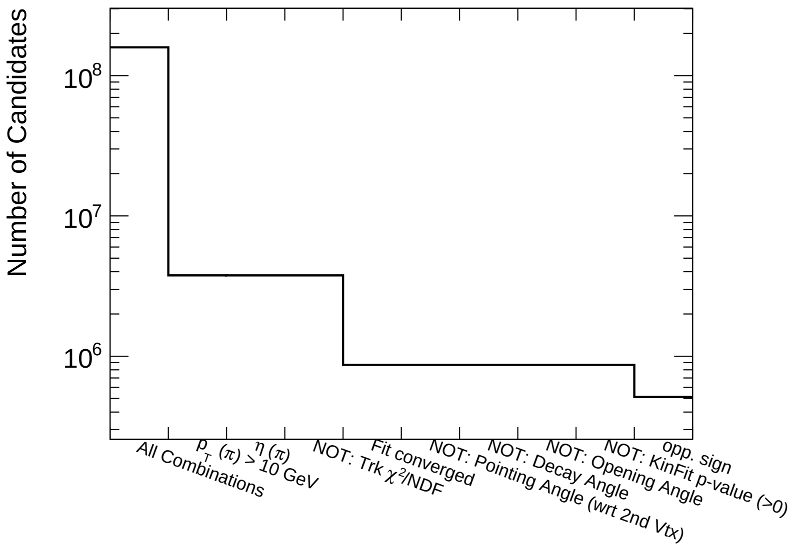


Figure 7.10.: Cutflow of the K_S^0 reconstruction based on the `KinFitter` algorithm in semi-logarithmic scale.

- **No fixed cut** is applied on the **quality of the tracks** since the tracks in the datasets are already preselected and tagged with regard to some quality criteria. These tags are then directly integrated as fit constraints for the K_S^0 reconstruction.
- **Execution of the `KinFitter` algorithm using several constraints:**
  - 4-momenta of the input tracks are taken as input.
  - Instead of a fixed quality cut on the tracks, the corresponding covariances of the reconstruction algorithm are used (cf. section 6.3).
  - The invariant track (i.e. pion) mass is set to $m(\pi) = 139.57$ MeV.
  - The reconstructed invariant K_S^0 mass should be close to $m(K_S^0) = 497.6$ MeV within a Gaussian distribution.
  - The constraints have to be fulfilled with a precision of at least $1 \cdot 10^{-6}$ (cf. equation 19 in [100]).
  - For the final fit, $\chi^2/n_{d.f.}$ has to be stable within $5 \cdot 10^{-5}$ when performing two further iterations of the fit.
  - The fit itself has to converge within 50 iterations.

- **No cut** is applied on the **pointing angle $\boldsymbol{\delta}$** as explained above.
- **No cut** is applied on the **decay angle $\boldsymbol{\beta}$** as explained above.
- **No cut** is applied on the **opening angle $\boldsymbol{\alpha}$** as explained above.

- **No cut** is applied on the **p-value of the fit** since this can be used as an input parameter of the BDT in chapter 8.
- The **two tracks** used have to be **opposite in charge**. Since K_S^0 is a neutral particle the charges of the daughter particles have to cancel out. Same sign candidates can be used for background studies.

The subsequent 4-momentum of the resulting K_S^0 is derived from the fitted tracks

$$p_{K_S^0} = p_{trk\,1}\,(fit) \;+\; p_{trk\,2}\,(fit). \tag{7.5}$$

Here, no cut on the transverse impact parameter d_0 of the tracks is set. Such a cut was suggested in several former K_S^0 analysis at lower energy $\sqrt{s}$, e.g. in [106] using $d_0 > 0.3$ cm. For the $\sqrt{s} = 8$ TeV analysis in hand, that does not apply owing to the much higher energies of all particles involved. Furthermore, the K_S^0 particle is expected to have an additional boost due to the top quark mother particle, which results in small opening angles $\measuredangle(K_S^0, \text{jet})$, $\measuredangle(\pi, \pi)$ and thus in small d_0 values.

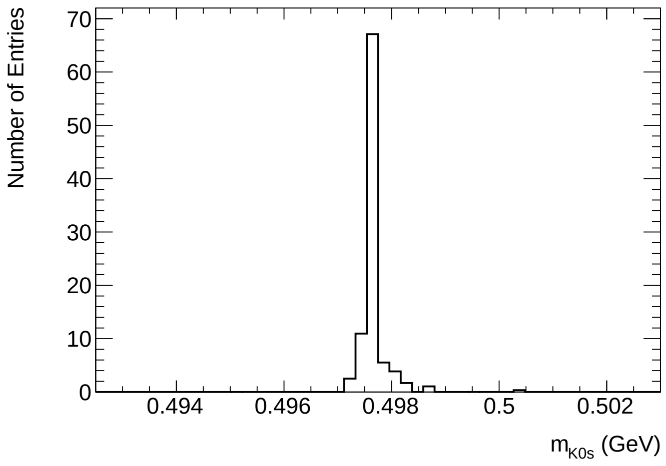


Figure 7.11.: Invariant mass of several K_S^0 candidates stemming from the `KinFitter` algorithm. The input constraint is set to $m(K_S^0) = 497.6$ MeV which is in exact agreement with the peak.

Among others, the effects of the constraints on the kinematic fit is important. Exemplary, the invariant mass distribution of the reconstructed K_S^0 particles is drawn in Figure 7.11, which justifies the mass constraint nicely, since no candidate lies outside the depicted mass window. A Gaussian fit leads to $m(K_S^0) = (497.63 \pm 0.02)$ MeV, which is identical to the input mass of $m(K_S^0) = 497.6$ MeV. Another typical output of the reconstruction are the p-values of the K_S^0 candidates, which describe the probability

for a correct hypothesis. Those are depicted in Figure 7.12.
Finally, the important truth comparison of resulting number of K_S^0 candidates in a jet is drawn in Figure 7.13 by overlying the distributions of truth and reconstructed particles, each divided according to their origin, $t \to s + W$ and $t \to b + W$ decays, respectievly. With respect to the truth study, the efficiency to reconstruct at least one K_S^0 candidate per jet is about 70 % for $t \to s + W$ decays and about 55 % for $t \to b + W$ decays. For the reconstruction there are more jets with multiple K_S^0 candidates than predicted. This is due to the high number of track combinations as well as due to additional Λ particles which are mis-identified as K_S^0, as described above. This behaviour can be neglected in the main $|V_{ts}|$ analysis, since only the particle with the highest p_{T} is considered per *probe jet*, called the "leading" or "highest" K_S^0, which is expected to arise directly from the high energetic s-quark.

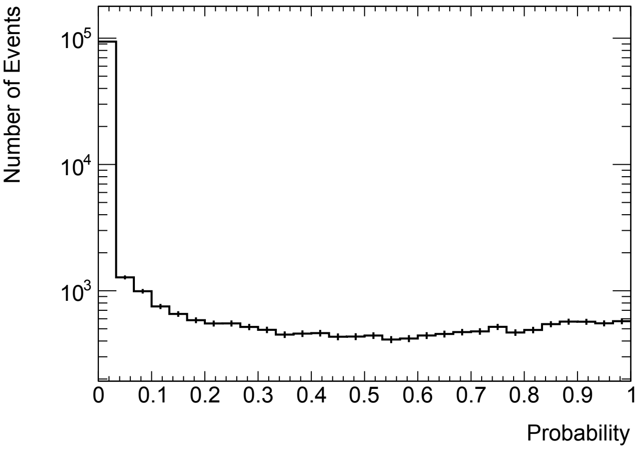
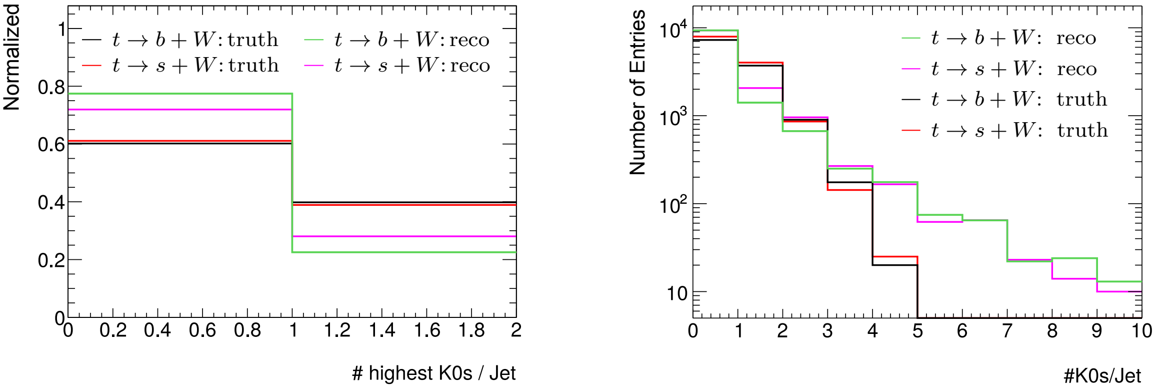


Figure 7.12.: Distribution of the K_S^0 candidates' p-values which describe the agreement with the hypothesis used. The diagram is illustrated in semi-logarithmic scale.

The reconstructed K_S^0 candidates can now be compared to the corresponding *probe jets* by analysing the ratio $p_{\mathrm{T}}(K_S^0)/p_{\mathrm{T}}(jet)$ (Figure 7.14). The value is higher for s-quark *probe jets*, thus implying a possible discriminability of s- and b-quarks.
An even better discrimination power in some bins is shown in Figure 7.15 in which the transverse momentum of the jet is subtracted by the $p_{\mathrm{T}}(K_S^0)$ resulting in the "remain. jet p_{T}" which is then compared to the $p_{\mathrm{T}}(K_S^0)$.

To sum up, a data-to-MC comparison is drawn in Figure 7.16 which shows the number of K_S^0 candidates per jet, identified with the `KinFitter` reconstruction method. Simulation and data agree well in this semi-logarithmic plot which indicates a thoroughly understood method. Each distribution from above includes all reconstructed K_S^0 candidates in order to present the full outcome of the method. However, in the following section only one K_S^0 particle per *probe jet* is regarded.



(a) **(b)**

Figure 7.13.: The number of reconstructed K_S^0 candidates is compared to the truth K_S^0 particles, combined with a division of K_S^0 stemming from $t \to s + W$ decays and those from $t \to b + W$ decays. (a) Ratio of jets with zero or at least one K_S^0 candidate. (b) Number of K_S^0 candidates per jet on a semi-logarithmical scale. All events are based on 8 TeV `PROTOS` events. Truth K_S^0 particles from $t \to s + W$ decays are red, those from $t \to b + W$ decays black. K_S^0 candidates reconstructed from s-quark *probe jets* are pink , those from b-quarks green.

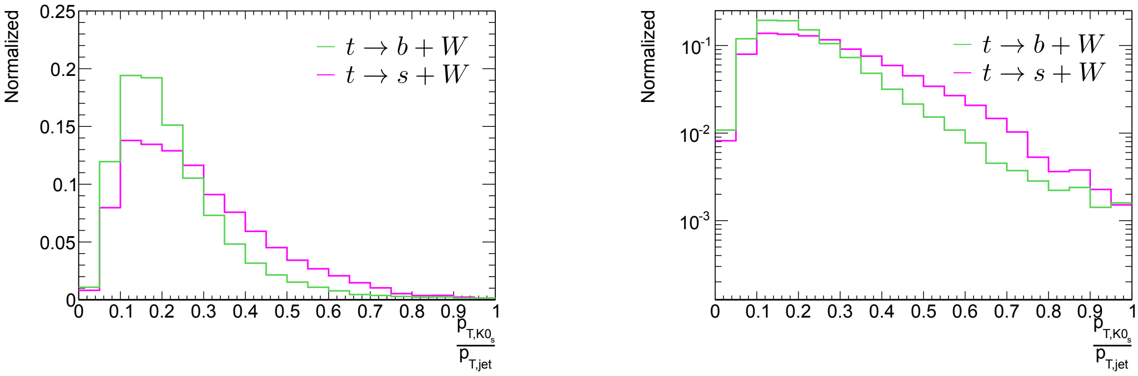


Figure 7.14.: Ratio $p_{\mathrm{T}}(K_S^0)/p_{\mathrm{T}}(\mathrm{jet})$ for K_S^0 candidates reconstructed from s-quark *probe jets* (pink) and b-quark *probe jets* (green). The right-hand plots are plotted semi-logarithmically.

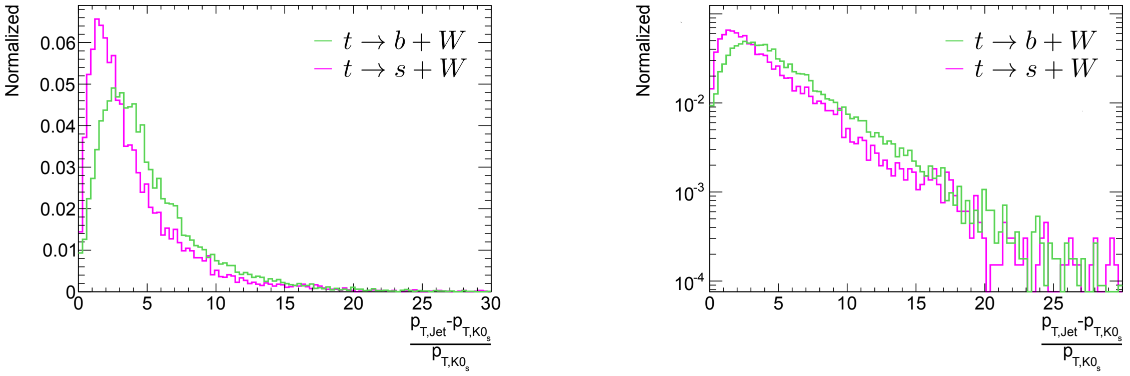


Figure 7.15.: Ratio "remain. jet p_T" vs. $p_T(K_S^0)$ for K_S^0 candidates reconstructed from s-quark *probe jets* (pink) and b-quark *probe jets* (green). The right-hand plots are drawn semi-logarithmically.

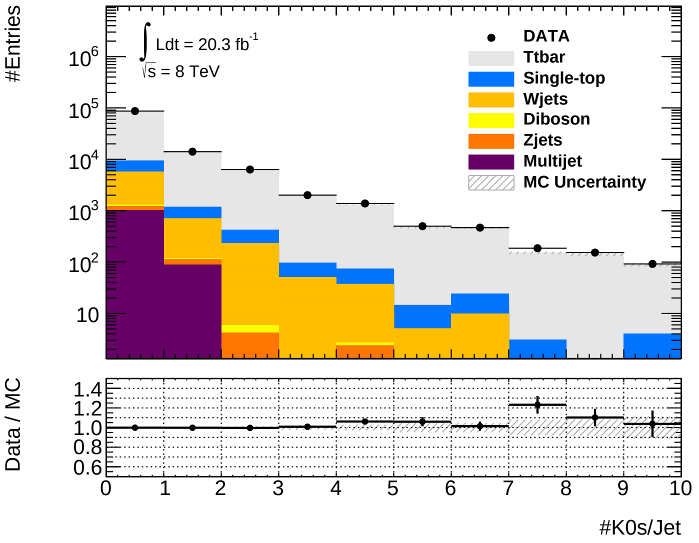


Figure 7.16.: The number of reconstructed K_S^0 candidates is in good data-to-MC agreement as shown in this semi-logarithmic plot, based on `MC@NLO` generated $t\bar{t}$ events. Thus, the reconstructed particles are usable for further analysis.

7.3.2. Available K_S^0 variables

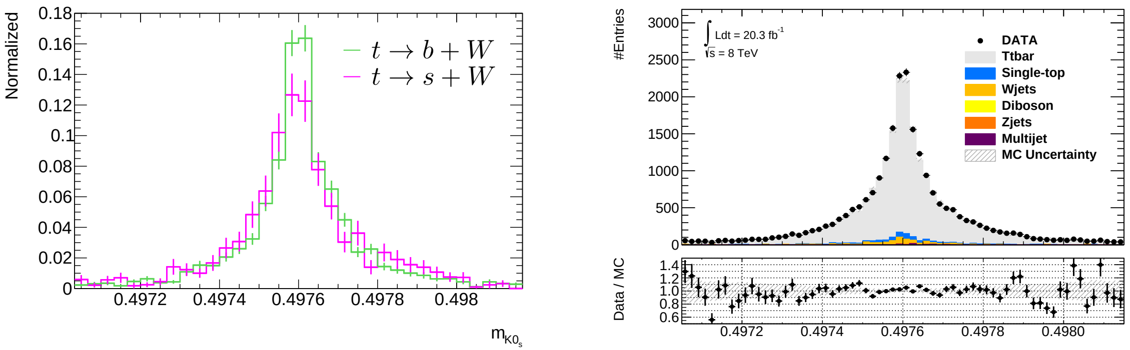
With the full K_S^0 reconstruction available and a good data-MC agreement for the number of reconstructed K_S^0, the next step consists of the separation between K_S^0 stemming from $t \to s + W$ and $t \to b + W$ decays. If there are multiple kaon candidates in a jet, it is assumed that the one with the highest transverse momentum is stemming from the top quark decay. Thus, only one K_S^0 per *probe jet*, the "leading" or "highest" K_S^0, is considered.

To find the best variables to identify $t \to s + W$ decays, all K_S^0 variables available are checked for their discrimination power and their data-MC agreement. In all following plots the discrimination power is drawn normalised on the left-hand side, with s-quark *probe jets* in pink and b-quark *probe jets* in green, based on 8 TeV `PROTOS` generated events due to the enhanced BR of $t \to s + W$ decays. On the right-hand side, a data-to-MC comparison is depicted, using `MC@NLO` generated $t\bar{t}$ events. This combination is chosen since `MC@NLO` does not comprise enough $t \to s + W$ events to perform the discrimination studies, whereas for the LO generator `PROTOS` the data-MC description is not as exact as for `MC@NLO`.

All explanations, describing the b-/s-quark differences and data-MC agreement, can be found in the corresponding captions. Only a short overview of the subsequent variables is given here:

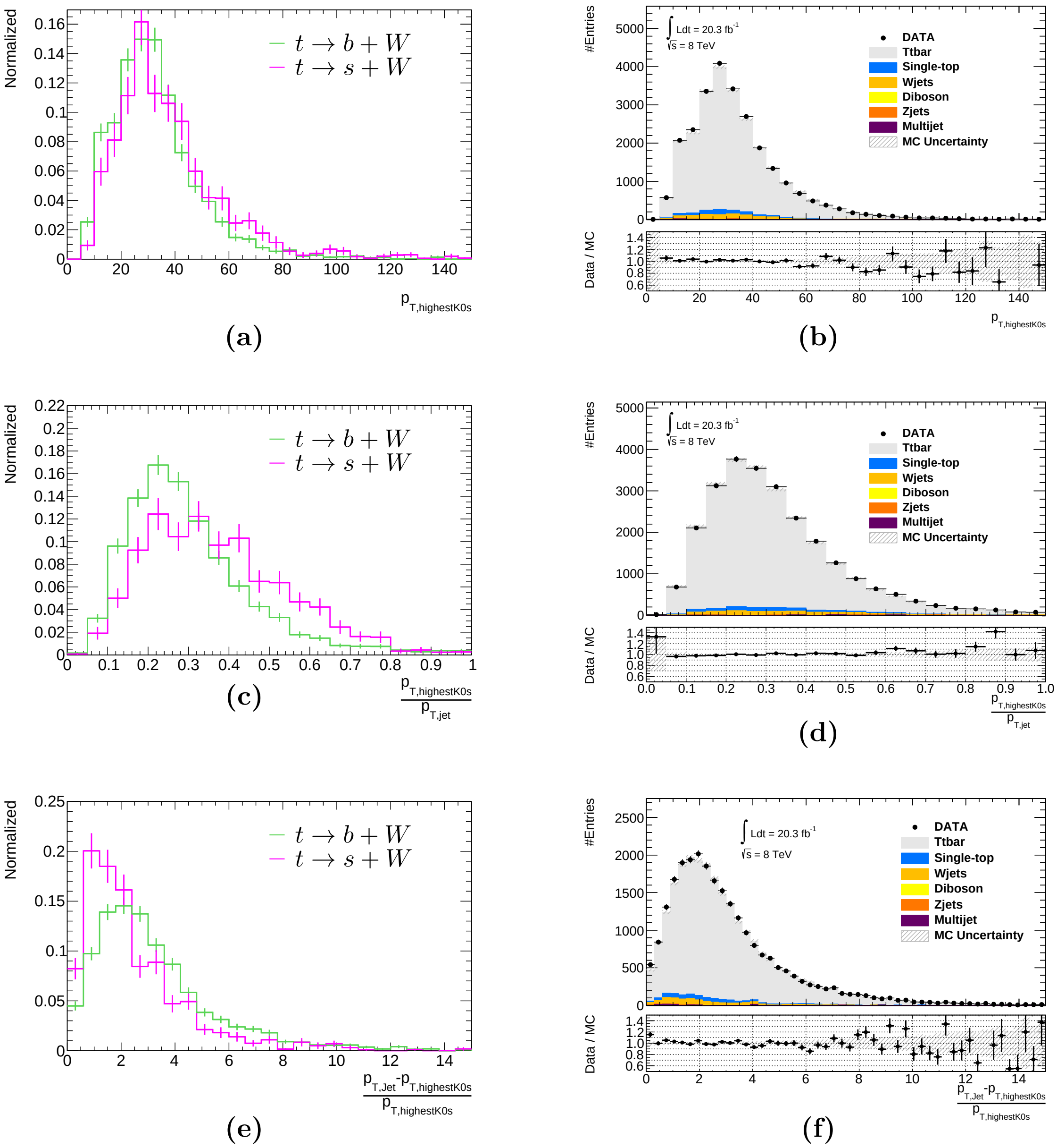
First, the mass and the kinematic properties of the reconstructed K_S^0 particles are illustrated in Figure 7.17 and in Figure 7.18, respectively. Within the latter, several combinations of the K_S^0 and jet transverse momenta are displayed, aiming for the best discrimination power. Second, the variables associated with the K_S^0 daughter tracks follow. They include the impact parameters d_0, z_0 and the opening angle between the two pion tracks in the rest frame of the mother top quark (Figure 7.19). Third, the reconstructed flight path of the K_S^0 is illustrated in Figure 7.20 which is derived from the tracks intersection point. The corresponding calculation is detailed below, starting from Equation 7.6. Finally, additional K_S^0 properties are considered, which give a relation to the corresponding *probe jet*, using the variable $p_{\mathrm{T,rel}}$ in Figure 7.24, and to the corresponding mother top quark, using the angle $\measuredangle(K_S^0$, top quark) in Figure 7.25.

In most of these plots, clear differences can be seen between $t \to s + W$ and $t \to b + W$ decays which are considered for the BDT algorithm in chapter 8.



(a) **(b)**

Figure 7.17.: The mass of the leading K_S^0 candidate is shown. The peak in (a) for b-quarks (green) is steeper than for s-quarks (pink). The `MC@NLO` data-MC comparison (b) is in good agreement.

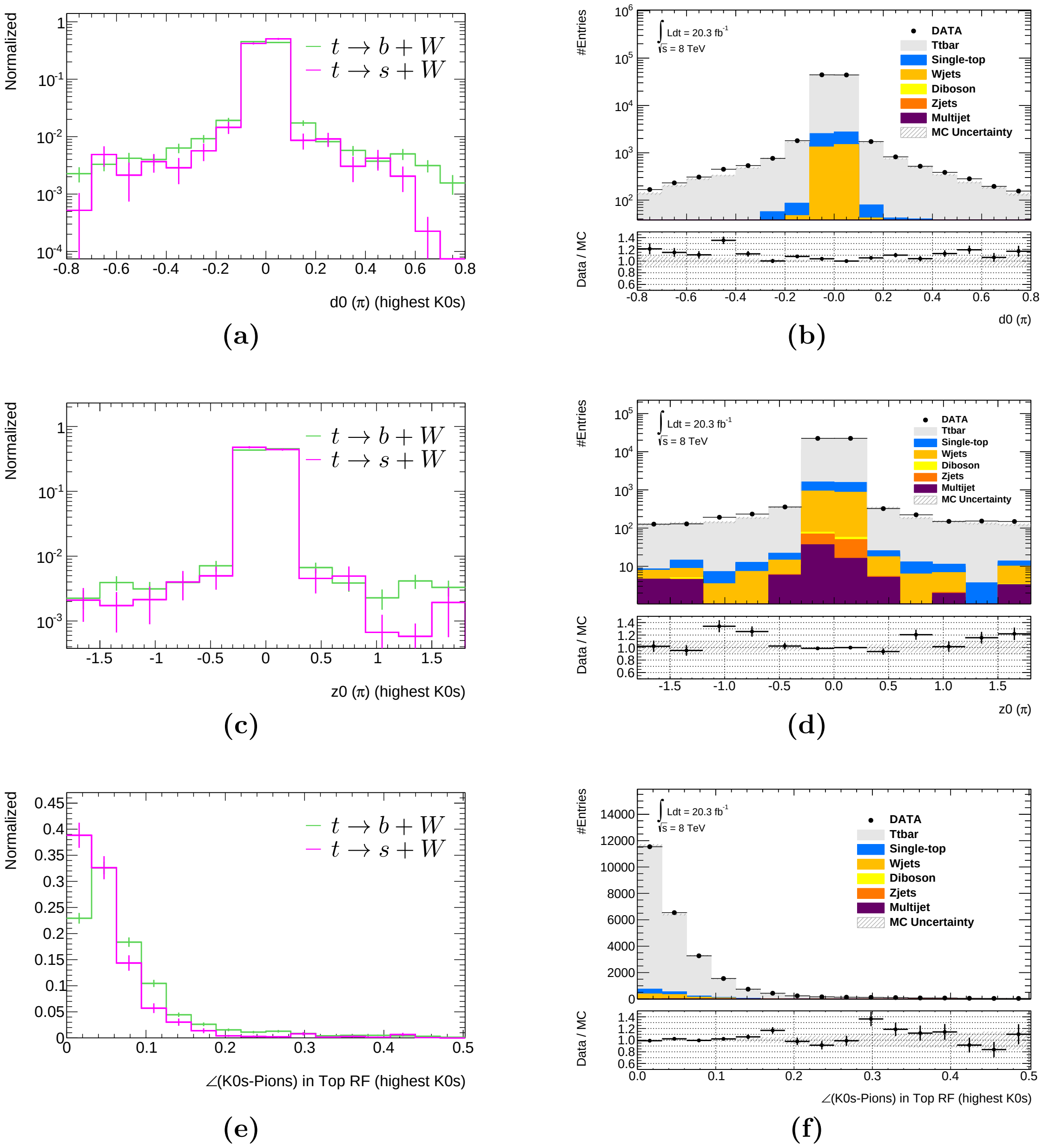


(a) **(b)**

(c) **(d)**

(e) **(f)**

Figure 7.18.: Transverse momentum distributions for the reconstructed leading-p_T K^0_S particles. A slight difference in the distributions of (a) is visible but the two figures in (c) and (e) enhance that effect clearly. In (c) a ratio with respect to the jet p_T is set whereas in (e) a comparison of the "remain. jet p_T" and the K^0_S p_T is performed. The `MC@NLO` data-MC comparisons on the right-hand side are in good agreement.

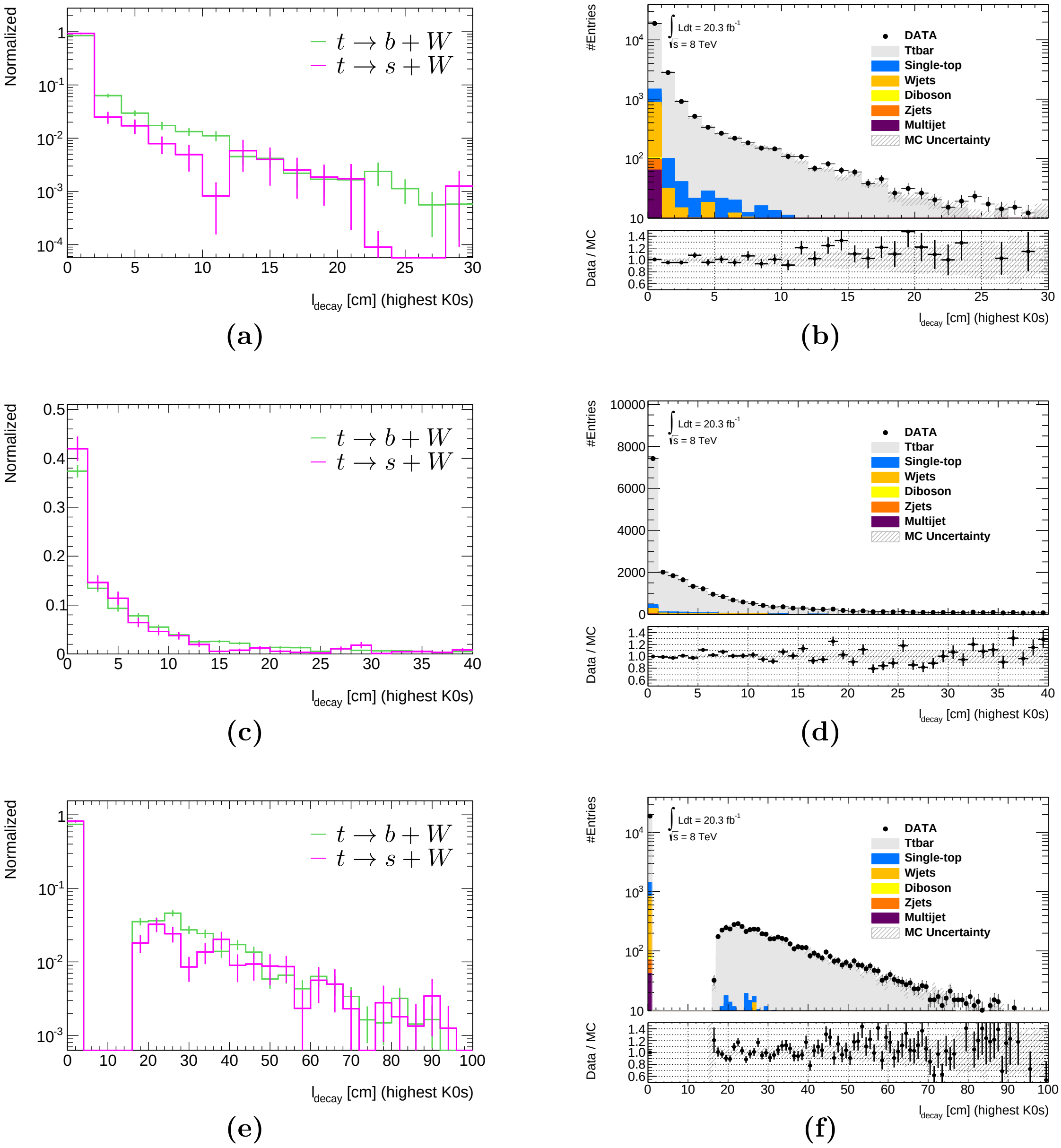


(a) **(b)**

(c) **(d)**

(e) **(f)**

Figure 7.19.: The properties of the tracks stemming from the reconstructed leading K_S^0 particles are illustrated. (a) and (c) represent the longitudinal impact parameter z_0 and the transverse impact parameter d_0 of the pion tracks, respectively. Small differences are visible in these semi-logarithmic plots between s-quarks (pink) and b-quarks (green) which might be used for the further differentiation. The corresponding data-MC comparisons on the right-hand side (b,d) are in good agreement.
(e+f) Analysis of the opening angle between the two tracks in the top quark rest frame. Although, the decays $t \rightarrow b + W$ and $t \rightarrow s + W$ differ in (e), no clear consensus of (e, green) with (f) is observable (especially in the 1st bin), even though both distributions are dominated by $t \rightarrow b + W$ decays. Thus, apparently `PROTOS` seems to model the here presented data imperfect. Among other reasons, that is why the BDT method and the final results are based on the `MC@NLO` generator (cf. chapters 8 and 9).



(a) (b)

(c) (d)

(e) (f)

Figure 7.20.: Overview of three different decay length measurements: Small (a,b), long (c,d) and midway (e,f). They correspond to the three outcomes of Equation 7.15 and are based on the intersection of the corresponding daughter tracks. If two tracks (circles in 2 dimensions) are crossing each other, the two intersection points are associated to the decay length "small" and "long", with respect to the PV. Without an intersection point, the point of minimal distance to both tracks is associated to the "midway" decay length (e,f).
Overall, K^0_S which stem from b-quark *probe jets* (green) decay at a larger distance from the PV. The data-MC comparisons on the right-hand side are in good agreement. All plots, except (c,d), are drawn semi-logarithmically.

The decay length distributions of Figure 7.20 are based on the calculations below. The aim is to reconstruct the intersection point of the two daughter tracks which should correspond to the decay point of the K_S^0 particle.
Assuming two dimensional planes in a transverse orientation, the tracks of charged particles in a homogeneous solenoidal magnetic field form simple circles. The calculations are thus much easier to solve in two dimensions than in three dimensions, since either an exact solution exists or the midpoint between those circles can easily be calculated. The following parameters describe the topology of the charged tracks:

- q/p $[\frac{e}{GeV}]$: Charge over momentum values to describe the radii, including a sign referring to the orientation of the curvature with respect to the PV.
- d_0 [cm]: Transverse impact parameters with respect to the PV.
- ϕ_0: Angles at the point of the closest approach to the PV.

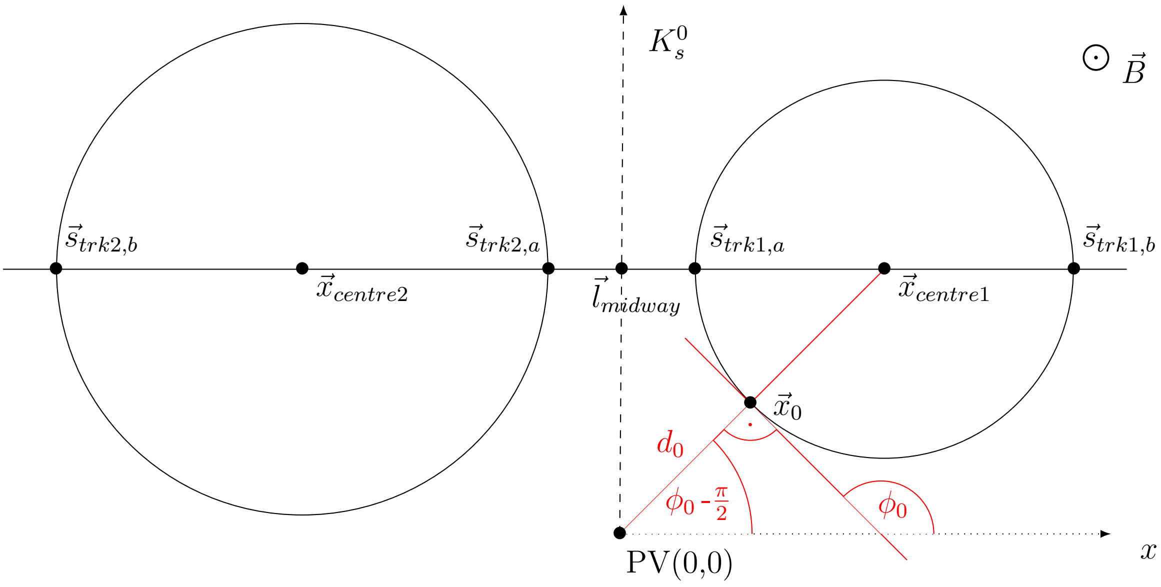


Figure 7.21.: Technical description of a K_S^0 decaying into two charged tracks ($trk1$, $trk2$) in a magnetic field $\vec{B}$. Here, the reconstructed tracks (circles) do not intersect each other, which implies a calculation of the midpoint.
Furthermore, the parameters for a track description in two dimensions are illustrated which includes the angle ϕ_0, the impact parameter d_0 as well as the PV.

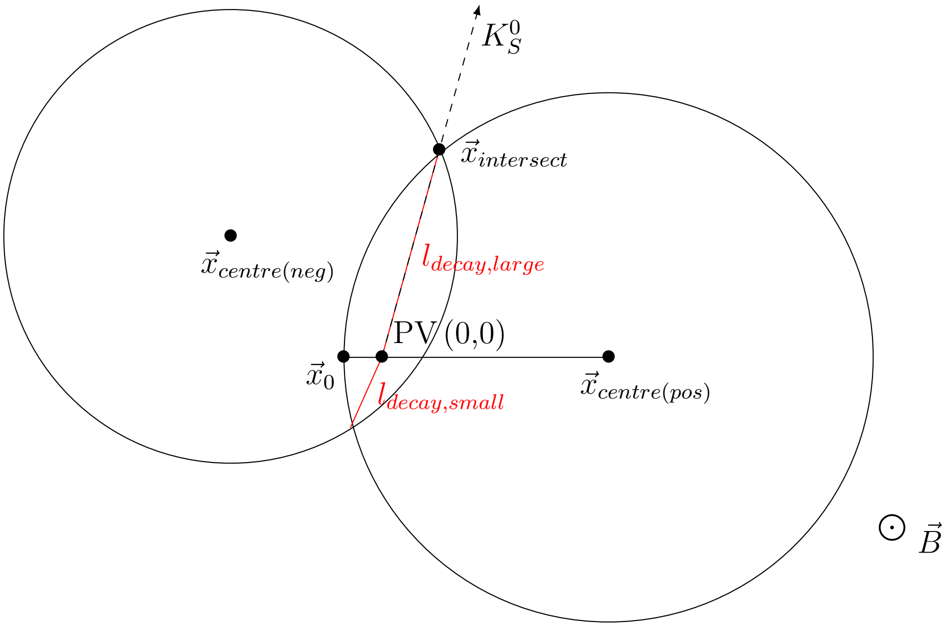


Figure 7.22.: Technical description of a K_S^0 decaying into two charged tracks, $trk\,(pos)$, $trk\,(neg)$. Here, the reconstructed tracks (circles) have two intersection points which are interpreted as decay point candidates of the K_S^0 particle.

The angles and impact parameters are outlined in Figure 7.21. The track's point of closest approach to the PV, $\vec{x}_0$, is defined as:

$$x_0 = -d_0 \sin(\phi_0) \tag{7.6}$$

$$y_0 = d_0 \cos(\phi_0) \tag{7.7}$$

$$R_{trk}\ [cm] = \frac{(10^9/3 \cdot 10^8)}{|q/p|} \tag{7.8}$$

The radius R_{trk} is deduced from the track's curvature (q/p) and the B-field in the inner detector which can be approximated with $B = 2$ T. The corresponding centre of the track's orbit can then be expressed as

$$\begin{pmatrix} x_{\text{centre}} \\ y_{\text{centre}} \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \cdot \left(1 \pm \frac{R_{trk}}{d_0}\right) \tag{7.9}$$

with a plus sign ("+") for positive charged tracks and a minus sign ("−") for negative charged tracks. The resulting intersection points $\vec{x}_{\text{intersect}}$ are described by the two equations

$$\begin{aligned}&\left(x_{\text{intersect}} - x_{\text{centre (pos)}}\right)^2 + \left(y_{\text{intersect}} - y_{\text{centre (pos)}}\right)^2 = R^2_{trk\,\text{(pos)}} \qquad (7.10)\\ &\left(x_{\text{intersect}} - x_{\text{centre (neg)}}\right)^2 + \left(y_{\text{intersect}} - y_{\text{centre (neg)}}\right)^2 = R^2_{trk\,\text{(neg)}}, \quad \text{which yields}\end{aligned}$$

$$x_{\text{intersect}} = a - y_{\text{intersect}} \cdot b \tag{7.11}$$

$$y_{\text{intersect}} = \frac{a \cdot b - b \cdot x_{\text{centre (pos)}} + y_{\text{centre (pos)}} \pm \sqrt{C}}{1 + b^2}, \quad \text{with } C, a, b\text{:}$$

$$\begin{aligned}C = &\left(a \cdot b - b \cdot x_{\text{centre (pos)}} + y_{\text{centre (pos)}}\right)^2 \\ &- \left(1 + b^2\right) \cdot a^2 - 2a x_{\text{centre (pos)}} + x^2_{\text{centre (pos)}} + y^2_{\text{centre (pos)}} - R^2_{trk\,\text{(pos)}}\end{aligned} \tag{7.12}$$

$$a = \frac{R^2_{trk\,\text{(pos)}} - x^2_{\text{centre (pos)}} - y^2_{\text{centre (pos)}} - R^2_{trk\,\text{(neg)}} + x^2_{\text{centre (neg)}} + y^2_{\text{centre (neg)}}}{2 \cdot \left(x_{\text{centre (neg)}} - x_{\text{centre (pos)}}\right)} \tag{7.13}$$

$$b = \frac{y_{\text{centre (neg)}} - y_{\text{centre (pos)}}}{x_{\text{centre (neg)}} - x_{\text{centre (pos)}}}\ . \tag{7.14}$$

The two possible values of $y_{\text{intersect}}$ (Equation 7.11) yield two decay lengths l_{decay} which are defined with respect to the PV (cf. Figure 7.22):

$$l_{\text{decay}} = \sqrt{x^2_{\text{intersect}} + y^2_{\text{intersect}}} \tag{7.15}$$

For the "small" decay length distribution the lower value $l_{\text{decay, small}}$ is chosen whereas for the "large" decay length distribution the higher one $l_{\text{decay, large}}$ is taken, Both are displayed in Figure 7.20 (a,b) and Figure 7.20 (c,d), respectively. In case of a negative C value in the square root, the two tracks are not crossing each other. Thus, the midway between the tracks' closest points is chosen, as detailed in Figure 7.21 and Equations 7.16, 7.17. At this, the vector $\vec{s}_{trk1}$ describes the track's closest point with respect to the second track and vice versa for $\vec{s}_{trk2}$.

$$\vec{s}_{trk1} = \left(1 \pm \sqrt{\frac{R^2_{trk1}}{(x_{\text{centre}\,1} - x_{\text{centre}\,2})^2 + (y_{\text{centre}\,1} - y_{\text{centre}\,2})^2}}\right) \cdot (\vec{r}_{\text{centre}\,1} - \vec{r}_{\text{centre}\,2}) + \vec{r}_{\text{centre}\,2} \tag{7.16}$$

The two solutions are opposite on the track's circle ($\vec{s}_{trk1,a}$ and $\vec{s}_{trk1,b}$ in Figure 7.21). To determine the point on track 1, which is closer to track 2, the distance to the centre of this second circle ($\vec{x}_{centre2}$) is evaluated. The point with the smaller distance is supposed to be the correct one (i.e. $\vec{s}_{trk1,a}$ in Figure 7.21). The identical procedure is performed for $\vec{s}_{trk2,a}$ which results in a "midway" point of the two tracks

$$\vec{l}_{\text{midway}} = \frac{\vec{s}_{trk1,a} + \vec{s}_{trk2,a}}{2} \tag{7.17}$$

with the corresponding distributions displayed in Figure 7.20 (e,f).

As a further discriminator, the relation of a K^0_S particle and the corresponding jet is examined. The p_{T} difference between the jet's momentum vector and the K^0_S momentum vector can be expressed by $p_{\text{T,rel}}$ of Equation 7.18, which is also illustrated in Figure 7.23. The results are depicted in Figure 7.24.

$$p_{\text{T,rel}} = |\vec{p}_{K^0_S} - \frac{\vec{p}_{K^0_S} \cdot \vec{p}_{jet}}{|\vec{p}_{K^0_S}|^2} \cdot \vec{p}_{jet}|. \tag{7.18}$$

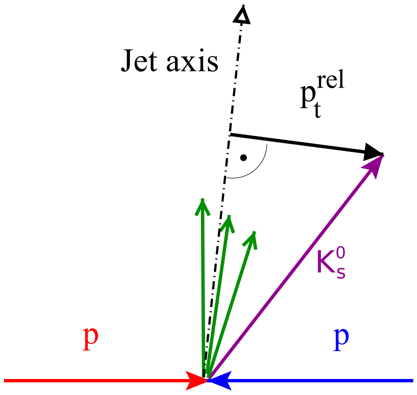
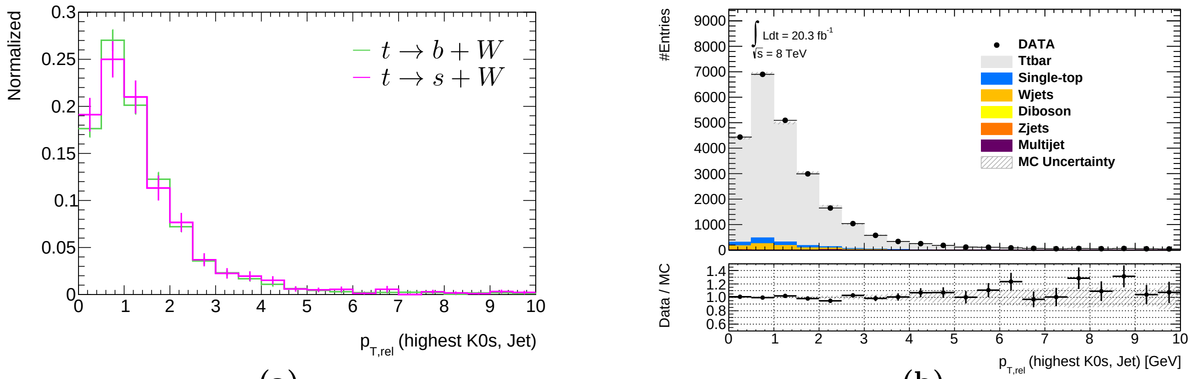
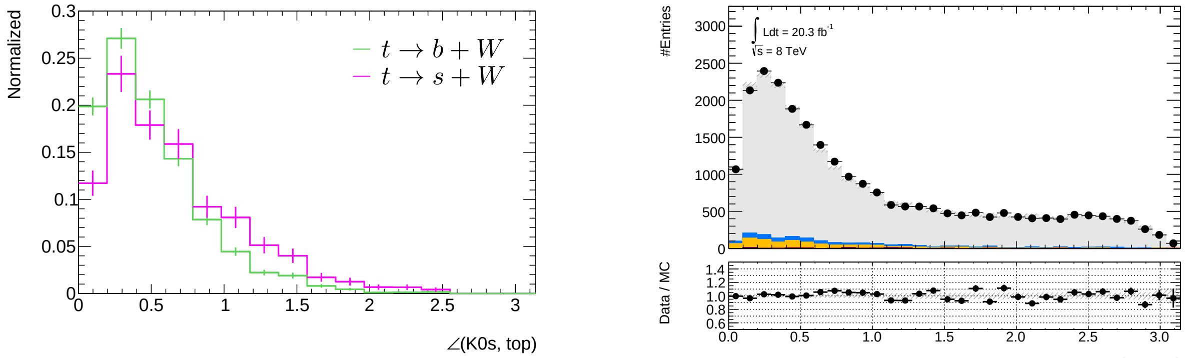


Figure 7.23.: The $p_{\text{T}}^{\text{rel}}$ variable used in the distributions of Figure 7.24 is shown. It is the vectorial difference between a K^0_S particle and its projection on the jet vector, calculated in the momentum space. Adopted from [117].



(a) **(b)**

Figure 7.24.: The p_T^{rel} variable is displayed which is calculated with respect to the jet axis as detailed in Figure 7.23. (a) A small tendency is visible for K_S^0 stemming from s-quark *probe jets* (pink) to be closer to the jet axis. (b) The data-to-MC comparison on the right-hand side is in good agreement.



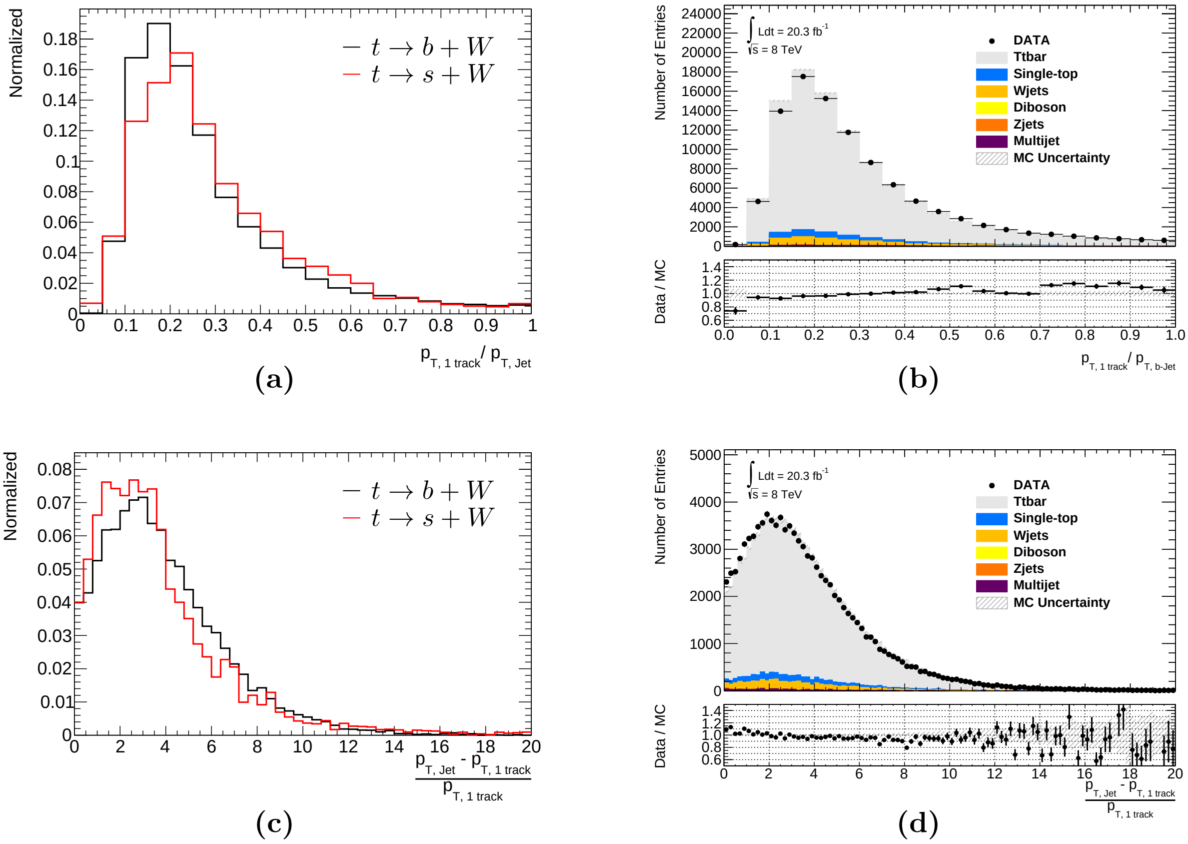
(a) **(b)**

Figure 7.25.: The angle between the reconstructed K_S^0 particle and the direction of flight of its mother top quark is displayed. (a) The angle between the top quark and the K_S^0 seems to be larger for K_S^0 from s-quark *probe jets* (pink) and smaller for b-quarks (green). The data-MC comparison on the right-hand side is in good agreement.

The last s- vs. b-quark discrimination parameter based on K_S^0 particles uses the relation to the mother top quark. A distribution of the angles between these particles are is in Figure 7.25 and indicate a good separation power.

7.4. Further parameters to identify $t \rightarrow s + W$ decays

Apart from the fundamental jet properties, discussed in the 7 TeV truth study (section 7.2), and the K_S^0 variables (section 7.3) which require at least one reconstructed K_S^0 per event, additional parameters can be used for a separation of $t \rightarrow s + W$ and $t \rightarrow b + W$ decays.

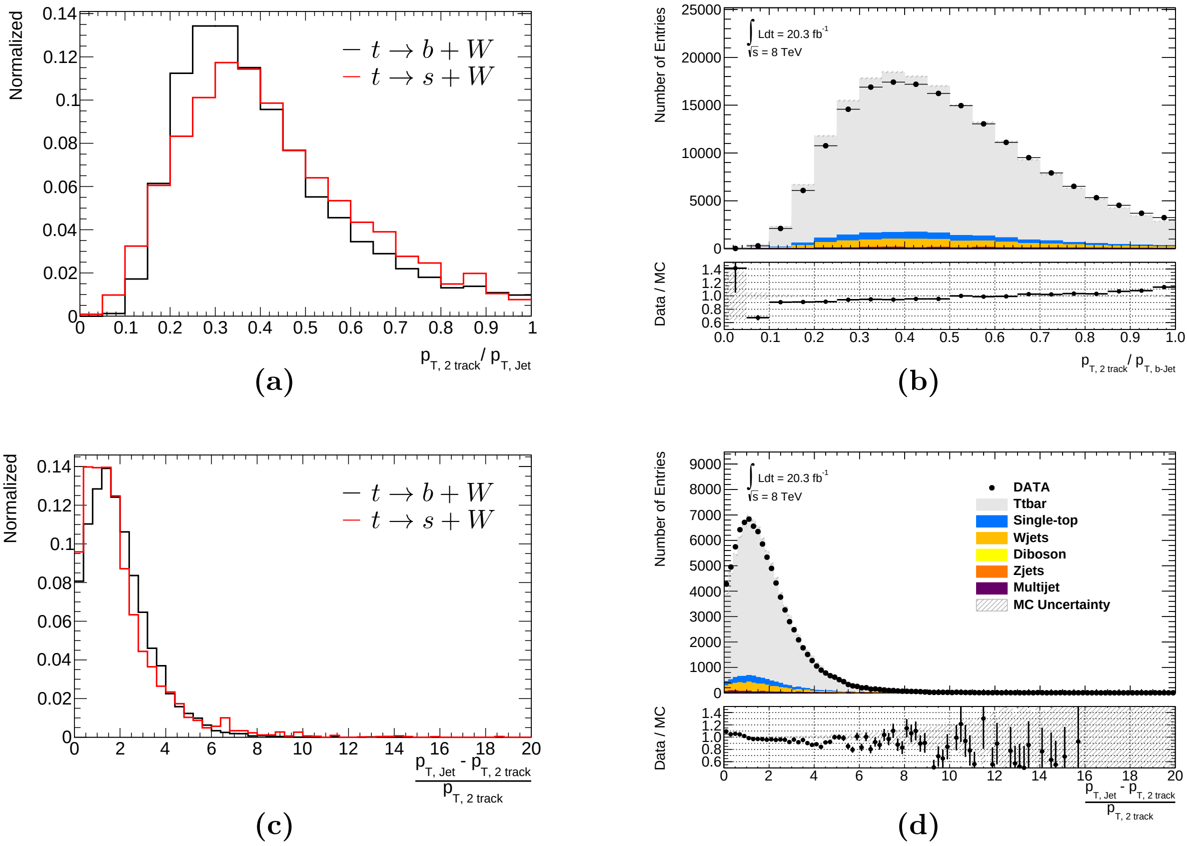


(a) **(b)**

(c) **(d)**

Figure 7.26.: The p_{T} of the leading track in a jet is compared to the jet p_{T}, similar to the K_S^0 parameters of section 7.3.2. The s-quark *probe jets* are indicated in red, whereas those from b-quarks are drawn in black. In (a), a simple ratio is shown whereas in (c) the "remain. jet p_{T}" is compared to the leading track p_{T} (cf. Equations 7.19 and 7.20), resulting in a slightly better discrimination power. The corresponding data-to-MC comparisons (b,d) indicate a reasonable agreement, even though small deviations are visible.

The first parameter is based on the track properties. Similar to the K_S^0 parameters (cf. Figure 7.18), the p_{T} of the leading tracks are compared to the jet p_{T}. This is performed for the three leading tracks with the highest p_{T}, which is written as formulas in Equations 7.19 and 7.20:



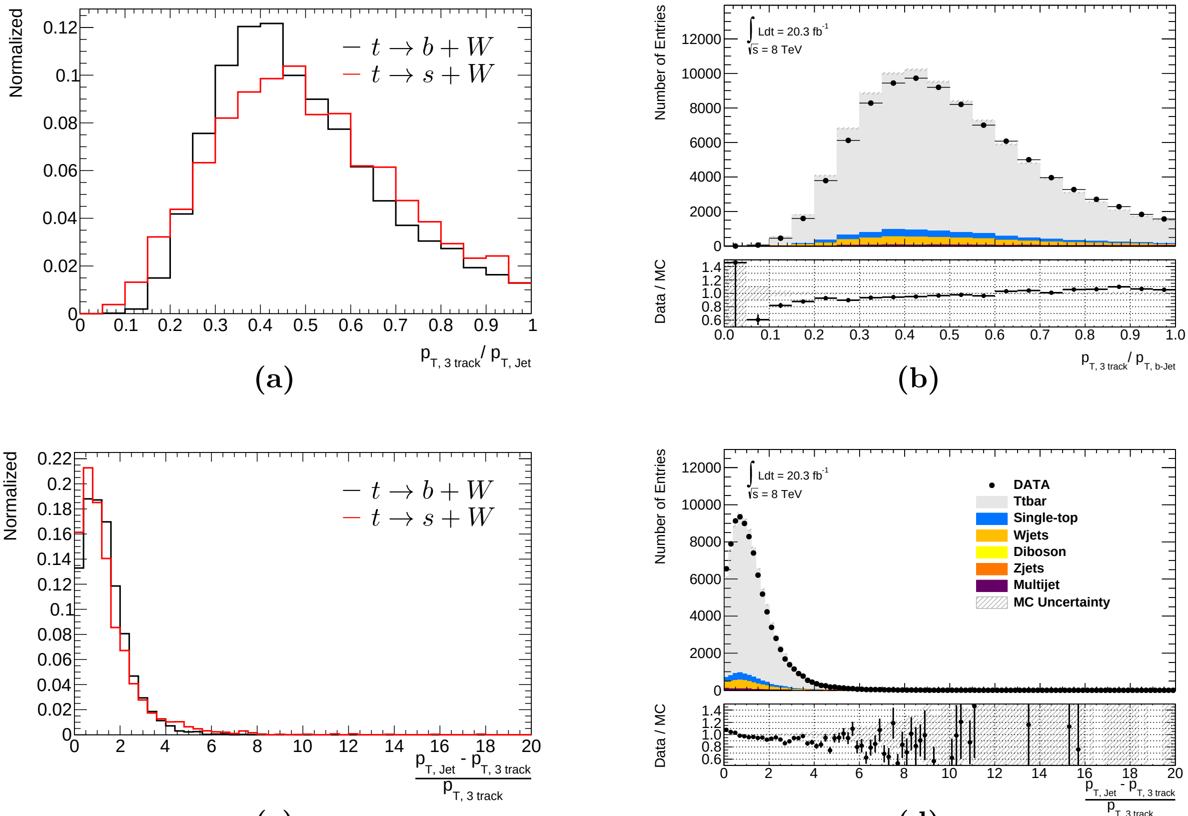
(a) **(b)** **(c)** **(d)**

Figure 7.27.: The p_T sum of the leading track and the 2$^\mathrm{nd}$ leading track in a jet is compared to the full jet p_T. The s-quark *probe jets* are indicated in red, whereas those from b-quarks are drawn in black. In (a), a simple ratio is shown whereas in (c) the “remain. jet p_T” is compared to the tracks p_T (cf. Equations 7.19 and 7.20), resulting in a slightly better discrimination power. The corresponding data-to-MC comparisons (b,d) indicate a reasonable agreement even though small discrepancies are visible, similar to Figure 7.26.

$$\frac{\sum\limits_{i=1}^{N} {p_\mathrm{T}}^{i^\mathrm{th}track}}{p_\mathrm{T}^{jet}} \tag{7.19}$$

$$\frac{p_\mathrm{T}^{jet} - \sum\limits_{i=1}^{N} {p_\mathrm{T}}^{i^\mathrm{th}track}}{\sum\limits_{i=1}^{N} {p_\mathrm{T}}^{i^\mathrm{th}track}}, \tag{7.20}$$

with $N = 1 \ldots 3$. The two different formulas describe the simple ratio of p_T sum of the tracks versus the jet p_T in Equation 7.19, and a ratio between the jet p_T without



(a) (b)

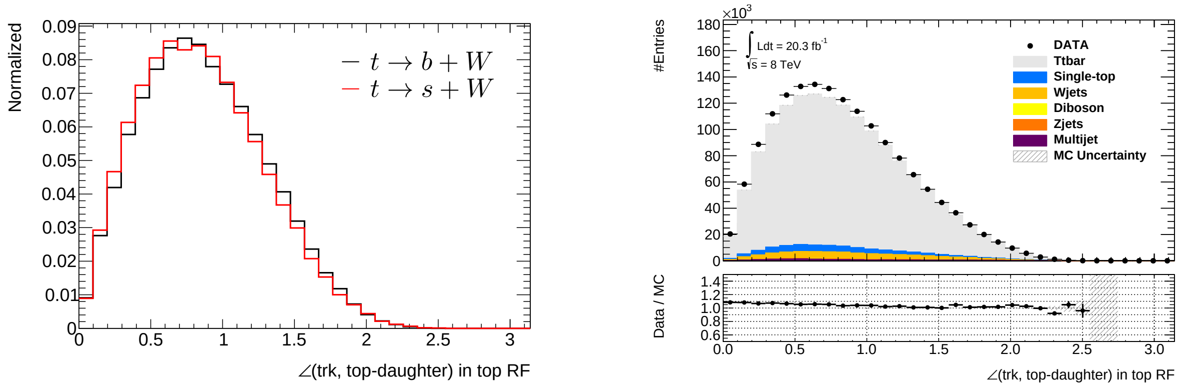
(c) (d)

Figure 7.28.: The p_{T} sum of the three leading tracks in a jet is compared to the full jet p_{T}. The s-quark *probe jets* are indicated in red, whereas those from b-quarks are drawn in black. In (a), a simple ratio is shown whereas in (c) the jet p_{T} without the tracks p_{T} ("remain. jet p_{T}") is compared to the tracks p_{T} (cf. Equations 7.19 and 7.20), resulting in a slightly better discrimination power. Similar as before, the corresponding data-to-MC comparisons (b,d) indicate a reasonable agreement, even though a few discrepancies are visible.

the tracks' p_{T} (further called "remain. jet p_{T}") versus the plain p_{T} of the tracks in Equation 7.20.

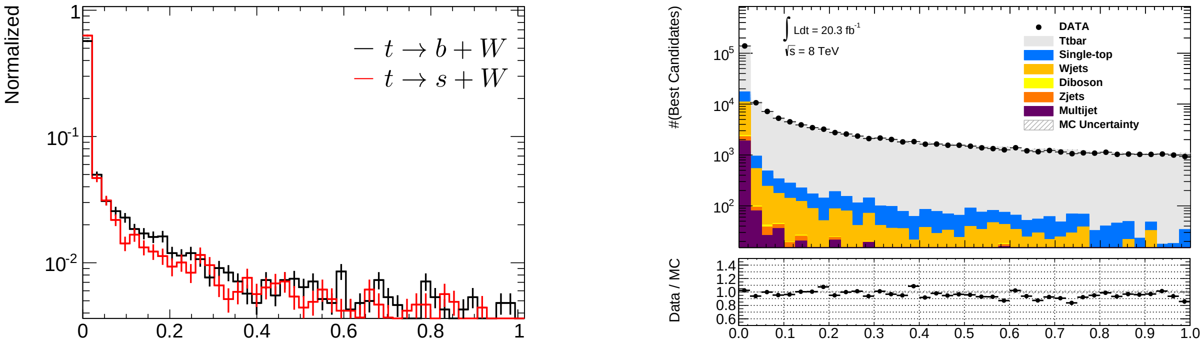
The corresponding plots are drawn in Figure 7.26 for the leading track, in Figure 7.27 for the two leading tracks and in Figure 7.28 for the three leading tracks of a single jet. Studies with 4 and 5 leading tracks do not show any more improvements. As before, the discrimination power between s- and b-quark *probe jets* is visible on the left-hand side, based on `PROTOS` generated events with enhanced $t \to s + W$ BR. The corresponding data-to-MC comparison, based on `MC@NLO` generated $t\bar{t}$ events, is on the right-hand side. The combination of data-MC consistency and discrimination power attains its optimum for three tracks in the "remain. p_{T}" distribution, Figure 7.28 (c,d). Thus, this parameter is incorporated in the BDT analysis of chapter 8.

Next, the angles between the *probe jet's* central axis and the direction of all jet tracks are considered, transformed into the rest frame of the mother top quark. The separation power as well as the data-to-MC comparison are depicted in Figure 7.29, which indicate a good discrimination. However, the BDT of chapter 8 only accepts one value per event, i.e. not multiple tracks per event, which means that only the RMS or the mean values of the distributions can be used. With considering discrepancies in the data-MC comparison and after an exemplary BDT training (cf. chapter 8), the RMS outcome is chosen as the final track-angle parameter for the subsequent analysis.



(a) **(b)**

Figure 7.29.: Angles between the *probe jet's* axis and the corresponding tracks, both transformed into the rest frame of the mother top quark. The discrimination power between $t \rightarrow s + W$ (red) and $t \rightarrow b + W$ decays (black) is clearly visible in (a), which indicates smaller angles for s-quark tracks. (b) The data-MC distributions are in reasonable agreement, small differences are considered by just using the RMS value for the BDT.



(a) **(b)**

Figure 7.30.: The p-value of the top pair χ^2 reconstruction method, introduced in section 6.3. As indicated in (a), events from $t \rightarrow s + W$ decays (red) have a slightly lower p-value than those from $t \rightarrow b + W$ decays (black). (b) The data-MC distributions are in good agreement.

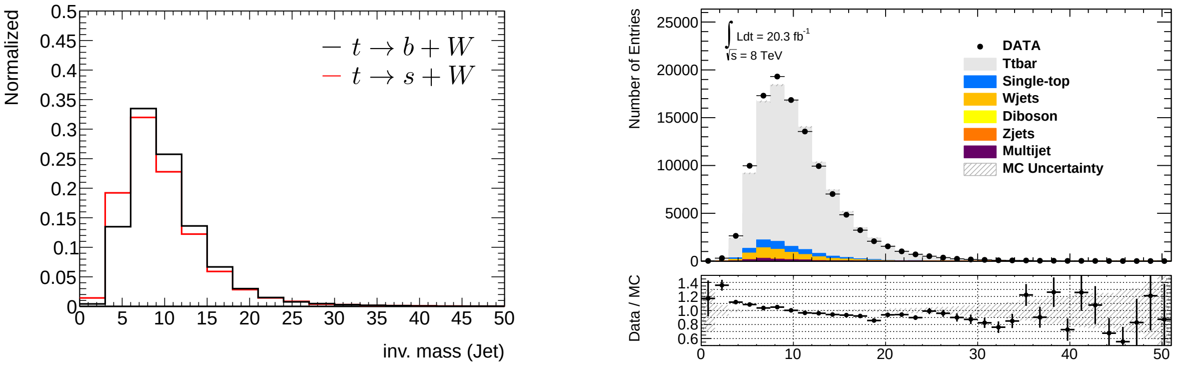
The last parameter which is taken into consideration for the BDT analysis of chapter 8 is the top pair reconstruction p-value which was introduced in section 6.3 with the kinematic χ^2 fitter. The corresponding distributions for the separation and the data-MC agreement are depicted in Figure 7.30 which both show a good performance.

All the differences discussed in the sections above indicate a possible discrimination of $t \rightarrow b + W$ and $t \rightarrow s + W$ decays which are further tested in the BDT of the next chapter 8.

7.5. Discrimination variables omitted in the final analysis

Several variables which could theoretically distinguish $t \rightarrow s + W$ from $t \rightarrow b + W$ decays can not be used for the final BDT analysis of chapter 8 due to different reasons.

First, the track multiplicity (number of tracks per jet) as tested in the 7 TeV truth study (section 7.2) would in principle improve the separation between s- and b-quarks but the data distribution can not fully be modelled by the MC generators as illustrated in Figure 7.1 (b). As already outlined, the number of tracks in the *probe jet* seems to be slightly higher than in the predictions.
Due to that improper description, further variables which are influenced by the track multiplicity also have to be omitted for the BDT. This includes the impact parameters d_0 and z_0 as well as the p_{T} of the individual tracks.



(a) **(b)**

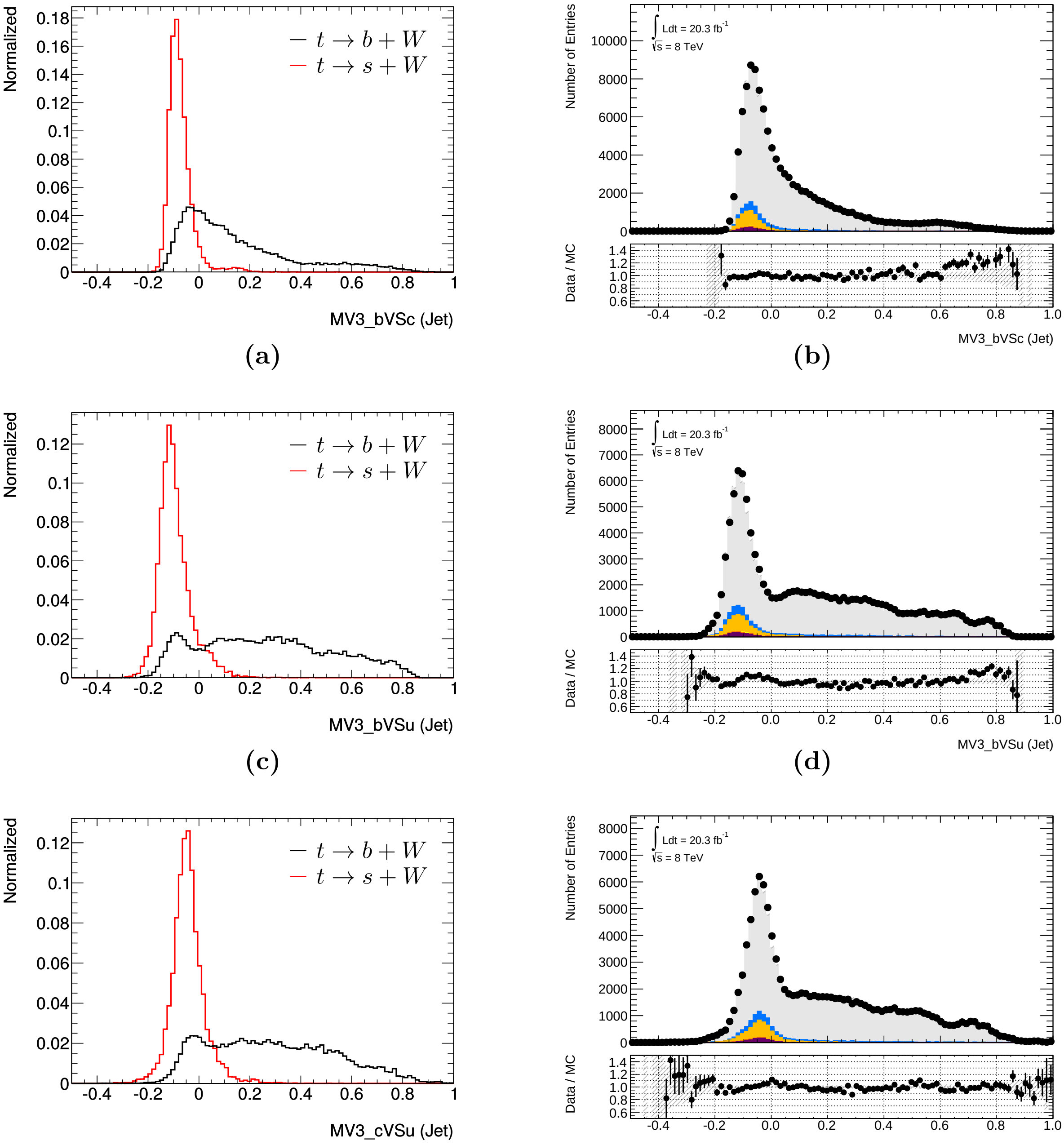
Figure 7.31.: Invariant mass of the *probe jets*. In (a) the possible discrimination power is shown with $t \rightarrow s + W$ decays in red and $t \rightarrow b + W$ decays in black, based on the `PROTOS` generated events. (b) The observed differences between data and generated events can not be explained reasonable and thus this parameter is excluded from the further analysis.

Moreover, the invariant masses of the *probe jets* seem to be also influenced by the improper MC modelling. The related distributions, showing the discrimination power and the data-to-MC comparison, are depicted in Figure 7.31. These can also be seen as exemplary plots for the tracks' d_0, z_0 and p_{T} distribution, illustrating a good separation power but an improper agreement between data and simulated events.

In addition to that, several other accessible parameters are tested but all result in a moderate data-to-MC comparison. Those include most variables which are associated to a reconstructed SV (invariant vertex mass, number of SV tracks, etc.), as well as all variables describing V0 properties. An exemplary distribution, the summed p_T of the V0 particles in an event, is displayed in Figure A.11. A discrepancy of 30 - 50 % between data and MC is visible.

Furthermore, another discriminating algorithm, the MV3 b-tagger, was considered for the BDT analysis. Similar to the MV1 algorithm, it is a newly developed b-tagger based on a multivariate method. The training was split, aiming for three different outcomes: First, to separate b- from c-quark jets, second, to separate b- from light-quark jets and third, to separate c- from light-quark jets. All three distributions are depicted in Figure 7.32.
Despite the fact that the MV3 was trained to discriminate b-quark jets from light-quark jets, it was not optimised to separate b-quark jets from high energetic s-quark jets arising from $t \to s + W$ decays. Instead, the central aim was to reduce the b-tag misidentification, as detailed in section 5.6. Thus, most of the discrimination power it achieves is already implied in the discrimination power of the MV1 variable, as several tests with the BDT method indicate. Moreover, the MV3 parameters are not officially approved by the ATLAS top working group, even though the data-to-MC agreement is quite reasonable. Thus, the MV3 variables are not used for the final analysis.

In addition to the above discussed and studied parameters, there are several more event properties which could in principle increase the separation between s- and b-quarks, stemming from top quarks, but are not accessible for the analysis in hand. Almost all of these are summarised in [106].
One of the most crucial ones is the variable describing the number of charged leptons in the *probe jets*. Some B hadrons decay semileptonically, resulting in additional charged leptons which can be used for the hadron identification [107]. However, in this analysis, pre-selection cuts are performed, demanding exactly one good charged lepton per event (cf. chapter 6). Thus, all information about additional “loose” charged leptons (cf. section 5.1) is discarded and can not be used anymore. In principle, charged lepton properties could also be used by considering properties of tracks, e.g. by counting the number of charged tracks per *probe jet*, as discussed in section 7.2, but unfortunately, the corresponding distribution has a moderate data-MC agreement as detailed above. Thus, for a future $|V_{ts}|$ analysis it would be an improvement to save the information of all charged leptons for further studies.
Last but not least, s-quarks also have the tendency to transfer their quantum numbers to the corresponding leading hadrons, as stated in [106]. Thus, a more detailed K^0_S analysis could further improve the separation power of the full $|V_{ts}|$ analysis.



(a) (b) (c) (d) (e) (f)

Figure 7.32.: A further development of the MV1 b-tagger results in the MV3 algorithm. It consists of 3 sub-taggers which are specialised to discriminate b- from c-quark jets (a,b), b- from light-quark jets (c,d) and c- from light-quark jets (e,f). The discrimination power, based on `PROTOS` generated events, is depicted on the left-hand side whereas the data-to-MC comparison is illustrated on the right-hand side, based on the `MC@NLO` datasets for $t\bar{t}$ events.

8. Multivariate analysis

For most of the parameters introduced in chapter 7 no simple cuts are applicable to discriminate $t \to s + W$ from $t \to b + W$ decays. The found differences are distributed among many variables and quite small for each individual one. Thus a multivariate analysis is introduced which combines the separation power of multiple variables into a single output variable.

For this, several concepts are available as summarised in [118]. For the $|V_{ts}|$ analysis, the *boosted decision trees* (BDT) algorithm is chosen which is also often referred to as the best "out of the box" classifier due to its simplicity, high performance as well as the robustness against possible perturbations [118].

The basic idea of that technique is outlined in section 8.1. A summary of the BDT parameters and the selected input variables is given in section 8.2. This is followed in section 8.3 by the technical description of the training and testing procedure. Finally, the responses of the classifier are drawn for all events and for events with at least one K^0_S candidate in sections 8.4 and 8.5, respectively, which can then be used for further analysis.

All conceptional descriptions in this chapter as well as the algorithm itself are based on the code and the explanations of the "Toolkit for Multivariate Data Analysis with ROOT" (TMVA) [118].

8.1. Concept of Boosted Decision Trees

A single *decision tree* is a collection of cuts, which has to be trained by events that are either tagged as signal or background, in order to discriminate signal-like events from background-like events when finally applied to data. The concept is illustrated in Figure 8.1. Each cut at a *node* is based on a binary separation of the data into signal- and background-like events. The selection is performed using the optimum discrimination variable at this stage in the tree by an automatic search for the corresponding cut value. After a separation into sub-samples, these are transferred to the next nodes for another distinction, forming a tree like structure.

The usage of input variables in the cutting nodes is not restricted, thus, they can either be used once, multiple times or not at all. The algorithm ends if a specified stop condition is reached which can for instance be a maximum number of nodes or if a minimum number of events per node is undershot. The final nodes, also referred to as *leafs*, are eventually tagged as *background* or *signal leafs*, depending on the fraction of events per category.

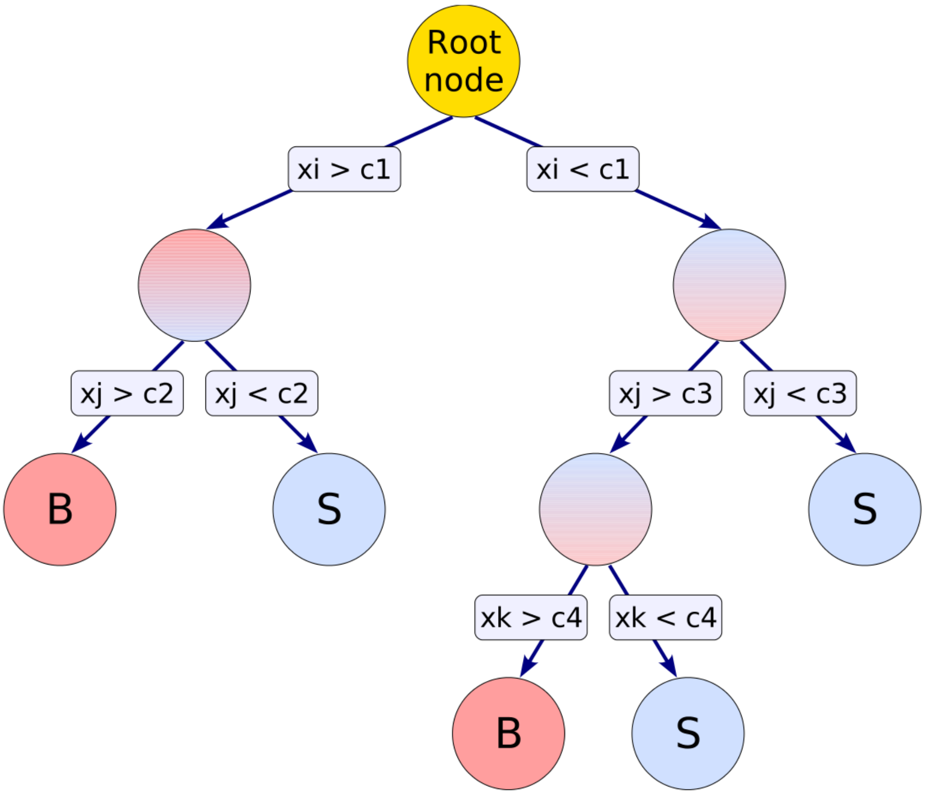


Figure 8.1.: Principle idea of a single *decision tree*. At each node, starting with the "root node", the algorithm performs a binary selection based on one variable (e.g. *xi*) with a corresponding cut (e.g. *c1*) yielding an allocation of each event to be background like (B) or signal like (S). To end up with *boosted decision trees*, multiple of these trees are grown resulting in a *forest*. Figure taken from [118].

The concept of *boosting* follows the idea of improving the first *decision tree* by generating further trees. Signal events, which are eventually misclassified as background (and vice versa) are weighted with a higher priority. The reweighted training sample is then passed to a new *root node* to restart the growing of a *decision tree* with the same events as before but with different priorities for the selection cuts. This procedure is called *boosting* since events with an incorrect assignment are boosted, to end up in the correct category in the newly grown tree. The boost weights themselves are defined with respect to the previous misclassification rate.
This procedure can be repeated several times (typical 50 - 500), thus, instead of a single *decision tree*, one ends up with a multitude of iterated *boosted decision trees*. This achieves a large improvement of the separation performance and the stability regarding statistical fluctuations, as first stated in [119]. The downside, however, is the loss of a clear decision tracing, since the number of individual trees is too large. A detailed

explanation of the here applied boosting method, called “Gradient Boost”, can be found in [118].
Speaking in the example of Figure 8.1, the final BDT can be visualised as a *forest* of the exemplary tree with each having different node properties. The subsequent signal and background event classification is based on a majority vote, taking all individual *decision trees* into account. The event’s output weight, called BDT response, is defined in the range $[-1, +1]$, in which -1 represents background-like events and $+1$ signal-like events. The BDT response itself can then be interpreted as a new variable, that can be subjected to a final analysis step, like in section 9.2.

To sum up, compared to a simple application of cuts, multivariate algorithms as used in the $|V_{ts}|$ analysis have the advantage to select multiple hypercubes in the variable space which can be identified as signal-like events. The BDT algorithm is chosen due to its high performance combined with a good robustness with respect to overtraining and imprecise variables. Furthermore, a relatively large number of input variables is allowed. [118]

8.2. Determination of input parameters

Multiple observable combinations are examined to find the best variables for the BDT method. For these tests, all 35 variables with a reasonable data-to-MC agreement from chapter 7 are included. To obtain the optimum performance of the BDT not all of those can finally be used. Thus, an indispensable reduction, based on a “minimal loss variable pruning” procedure, is accomplished:[1]
Each variable is in turn once excluded and the BDT is trained with the remaining $N-1$ variables. The N different compositions are compared in terms of influence on the BDT response and the least important variable is discarded. An execution of this pruning procedure for several times reduces the total number of variables substantially. This iterative reduction procedure stops if a significant BDT performance drop takes place.
As detailed in section 8.3, `PROTOS` generated events are used in the BDT training, but `MC@NLO` generated events in the subsequent final analysis step. To account for this approach, an estimated rating based on the separation power of all 35 variables is performed, using `MC@NLO` events. The outcome of that `MC@NLO`-based rating also influences the decisions of the pruning procedure, even though it is mainly based on `PROTOS` events.
As a result, seven variables are selected for the final BDT training which are summarised in Table 8.1. Three additional K_S^0 variables are used for events with at least one K_S^0 candidate, listed in Table 8.2.

---

[1] During this reduction process, the BDT parameters, introduced in section 8.3, are fixed.

Table 8.1.: Discriminating variables which are used for the BDT training, including references to their distributions. All occurring jet properties are defined with respect to the *probe jet*.

MV1 b-tagging weight (jet)	Figure 7.4
p-value (top reconstruction)	Figure 7.30
p_T (jet)	Figure 7.3
jet width (jet)	Figure 7.2
RMS of distribution $\measuredangle$(tracks, jet) in the top quark's rest frame	Figure 7.29
"remain jet-p_T" vs. "3x tracks' p_T"	Figure 7.28 (c,d)
"remain jet-p_T" vs. "K_S^0 p_T"	Figure 7.18 (e,f)

Table 8.2.: Additional variables for events with at least one K_S^0 candidate, which are used in the BDT training.

Opening angle of K_S^0 pions in the top quark's rest frame[2]	Figure 7.19 (e,f)
Decay length (small) of K_S^0	Figure 7.20 (a,b)
Decay length (large) of K_S^0	Figure 7.20 (c,d)

The p-value of the K_S^0-finder is not included in the method since no improvement in the discriminative BDT output is seen. One reason for this behaviour might be the demand for just one K_S^0 particle per event, in general a high energetic particle, which should easily be reconstructable by the fitting algorithm.

Since correlations between the input variables could reduce the BDT performance and are thus undesirable [118], high correlation values are considered as further decision guidance during the pruning process. The correlations of the final variables are depicted in Figures 8.2 and 8.3 for all events and events with at least one K_S^0 candidate, respectively. It can be seen that in none of the figures strong correlations are present, and thus, all finally selected variables match the requirements of the BDT method.

---

[2] Even though this variable shows an imperfect agreement, when comparing `PROTOS` generated and `MC@NLO` generated $t \to b + W$ decays in Figure 7.19 (e,f), it enhances the BDT decision. This improvement is especially visible for the final application on `MC@NLO` events.

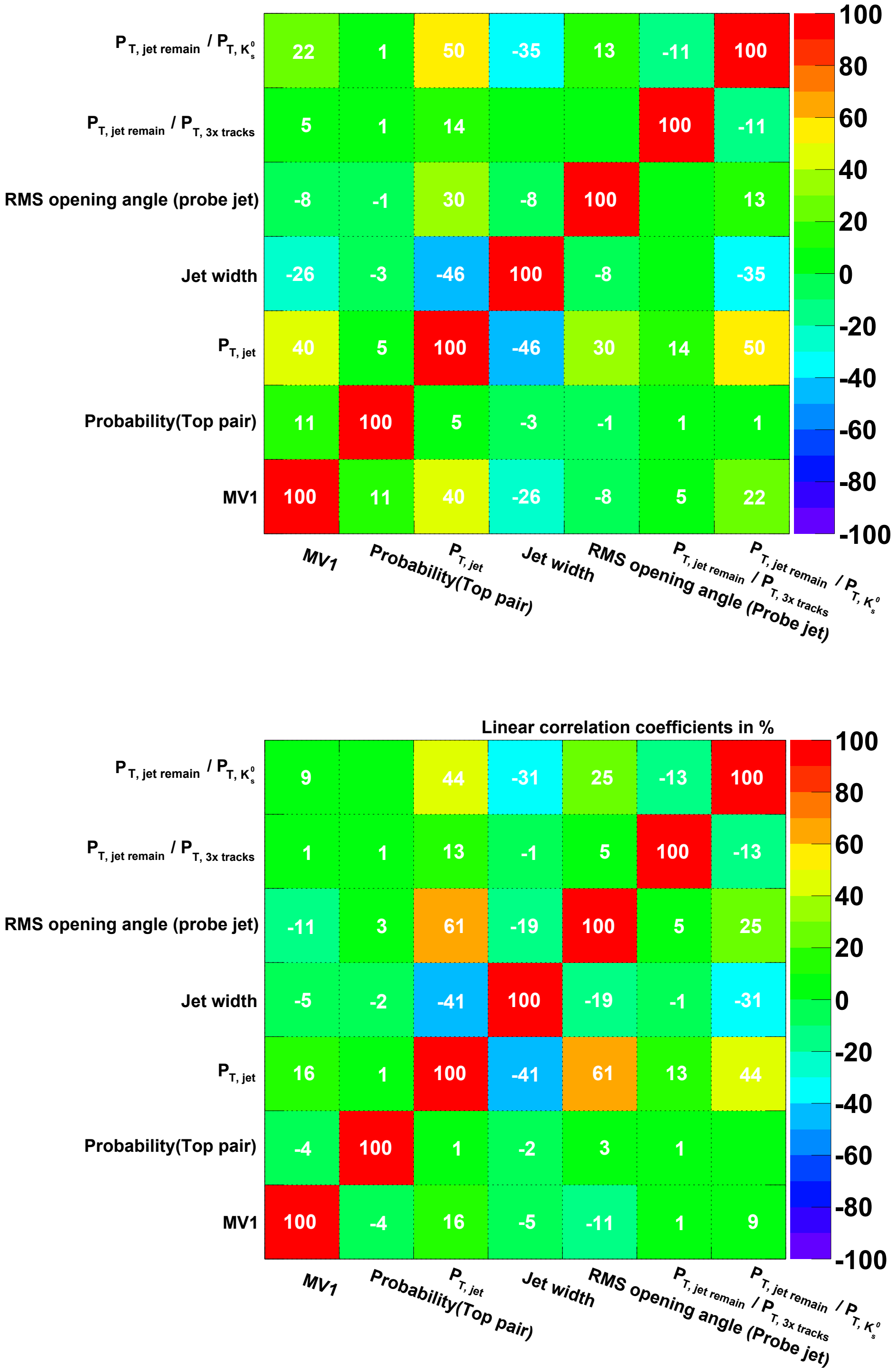
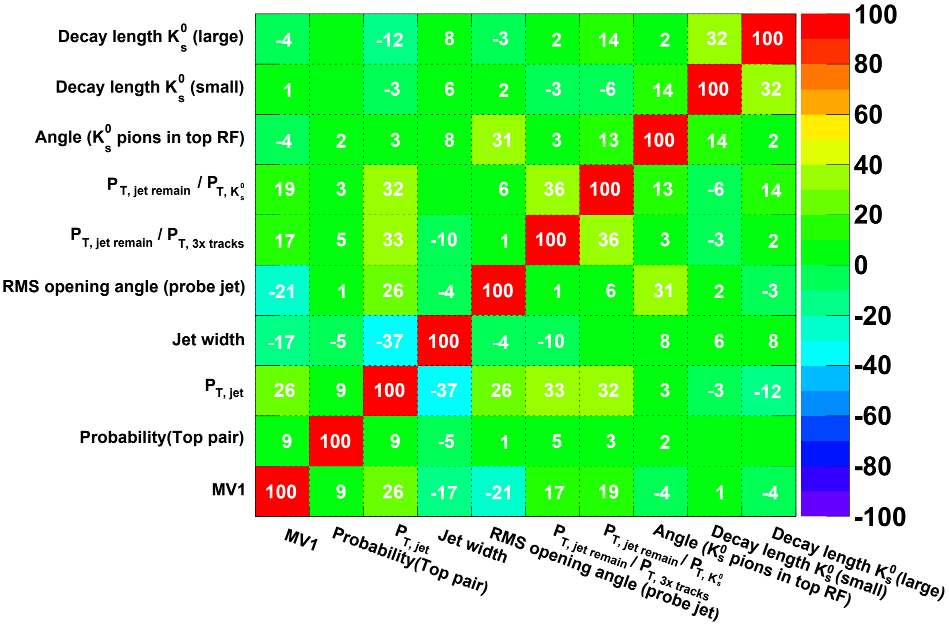


Figure 8.2.: Correlation matrices of input variables for all events. The upper diagram shows the background datasets, the lower one the signal dataset. Both are based on `PROTOS` generated $t\bar{t}$ events.



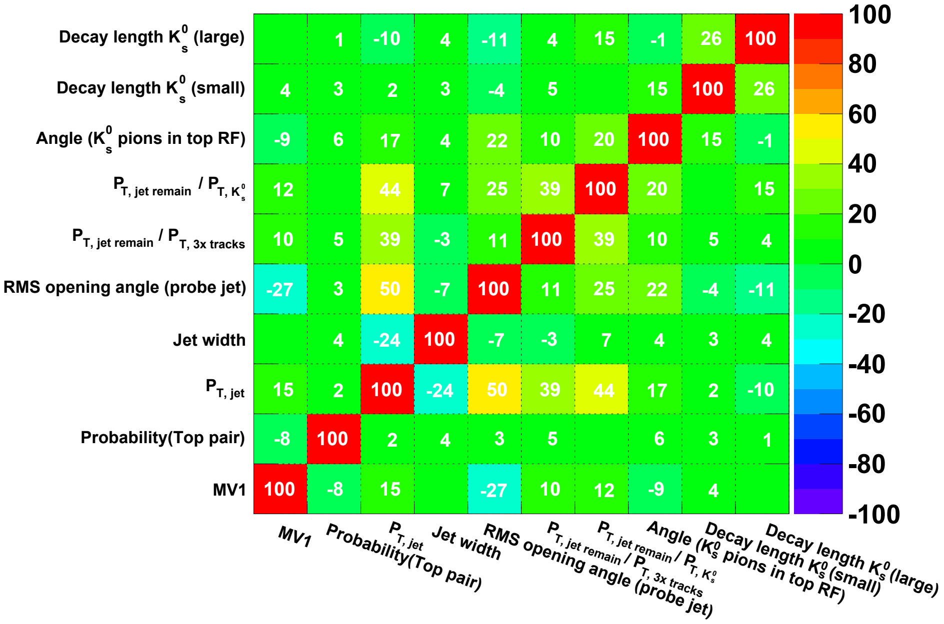


Figure 8.3.: Correlation matrices of input variables for events with at least one K_S^0 candidate. The upper diagram shows the background datasets, the lower one the signal dataset. Both are based on `PROTOS` generated $t\bar{t}$ events.

8.3. Training and testing procedure

In the next step, several classifier parameters are set, which includes e.g. the number of nodes and the corresponding cut values. Those are defined by training the BDT with respect to the identified input variables of the section above. The `MC@NLO` dataset does not contain enough $t \to s + W$ events for an adequate training, thus the BDT decisions are based on the `PROTOS` datasets (cf. chapter 4). However, the subsequent analysis of chapter 9, which requires a perfect data-to-MC agreement, is performed with `MC@NLO` since a reasonable number of signal events is present for this final analysis step.
The $t\bar{t}$ `PROTOS` dataset with enhanced BR is used to describe the signal events $t\bar{t} \to WWb\bar{s}$, with 8436 events which passed the selection criteria of chapter 6. For the background sample, the $t\bar{t}$ `PROTOS` dataset with SM BR is used to describe the events $t\bar{t} \to WWb\bar{b}$, with 45 438 events in total. It contains about four times more $t \to b + W$ decays than the dataset with enhanced BR. Moreover, the processes of $W + jets$ are taken into account as further backgrounds due to their prevalence in the signal region. All other smaller background contributions are not included since this would reduce the overall BDT performance significantly without improving the discrimination performance.
Before the actual training is carried out, apart from the selection cuts of chapter 6, additional cuts can be set. Two trainings are hence performed, the first without any additional cut applied, the second demands at least one K^0_S candidate per event. Consequently, the BDT outputs are separated into two parts, presented in section 8.4 and section 8.5.

Apart from these cuts, further method related parameters are set for the BDT algorithm, specified below and in Table 8.4.

- NTrees:
  Number of decision trees in the classification method.

- Shrinkage:
  Parameter representing the learning rate of the BDT which adjusts the weight of the single trees. Setting this parameter to a small value can improve the prediction accuracy since more trees are equally taken into account.

- nCuts:
  Number of sampling points, represented by bins in the histogram, to find the optimum cut value for a variable. Higher granularity can improve the differentiation at the cost of CPU time.

- MaxDepth:
  Maximum depth of the individual decision trees.

- MinNodeSize:
  Minimum number of events which are allowed per leaf node, given in percentage.

- Training fraction (Signal):
  Fraction of signal events used for the training procedure. The remaining events are used for a subsequent test.

- Training fraction (Background):
  Fraction of background events used for the training procedure. The remaining events are used for a subsequent test.

Table 8.4.: Optimised parameter combination for the two BDT trainings, the first without any cuts applied, the second with at least one K_S^0 candidate.

BDT parameters	All events	N(K_S^0) ≥ 1
NTrees	80	85
Shrinkage	0.18	0.20
nCuts	20	20
MaxDepth	4	3
MinNodeSize	5.66 %	5.00 %
Training fraction (Signal)	50 %	66 %
Training fraction (Background)	50 %	50 %

For the final BDT training the parameters of Table 8.4 are chosen. Similar to the variable decisions of section 8.2, this final selection emerges from plenty of training cycles. The central aim is always to obtain the best separation between signal and background events in order to improve the figure of merit $S/\sqrt{S+B}$. A close to perfect classifier distribution would look like the training samples in Figure 8.4 (dots with error bars).
Besides accounting for the separation power, it is also crucial to achieve a good agreement between training distribution and the output of an independent test run.

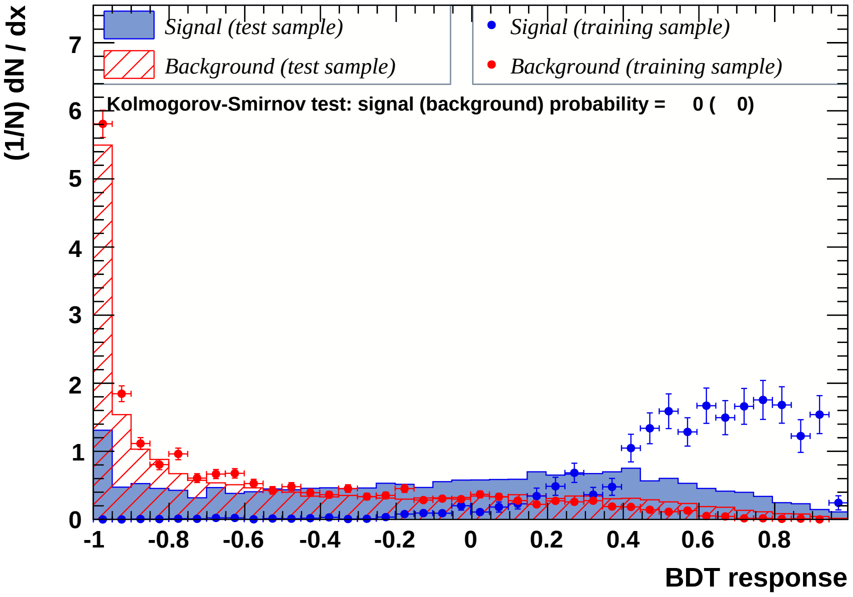


Figure 8.4.: Typical behaviour of an overtrained BDT classifier with signal events depicted in blue and background events in red. The training samples (dots with error bars) indicate a very good discrimination power of signal and background. The test samples instead (filled histograms), which should in principle show the same behaviour, are totally different with a quite vague discriminating efficiency.

Large deviations are a clear indication for overtraining, as visible in Figure 8.4 when comparing the training and test samples. This behaviour can occur if too many degrees of freedom in the classification algorithm are adjusted to only a few training events. Thus, fluctuations can be recognised erroneously as real physical differences. The resulting effect is a seemingly increase of the separation power in the training sample, in contrast to a clear loss of performance in the test sample and so in the final analysis. The corresponding plots for the finally obtained BDT responses are depicted in Figures 8.5 and 8.9.

Due to the high number of nodes, the BDT method in general has a strong tendency for overtraining if no counteracting operations are performed. Those operations always aim for a reduction of the degrees of freedom by e.g. diminishing the number of nodes, the tree depths or varying other parameters. As a support for detecting overtrained outputs, the TMVA package [118] provides a ”Kolmogorov-Smirnov test“ [120], which determines the differences between training and test samples. If its outcome deviates from zero, the tendency for overtraining reduces dramatically.

8.4. BDT response for all events

For the BDT training based on events without additional cuts, the discriminating variables from Table 8.1 as well as the parameters from Table 8.4 are used. The corresponding linear correlation coefficients are summarised in the matrices of Figure 8.2.

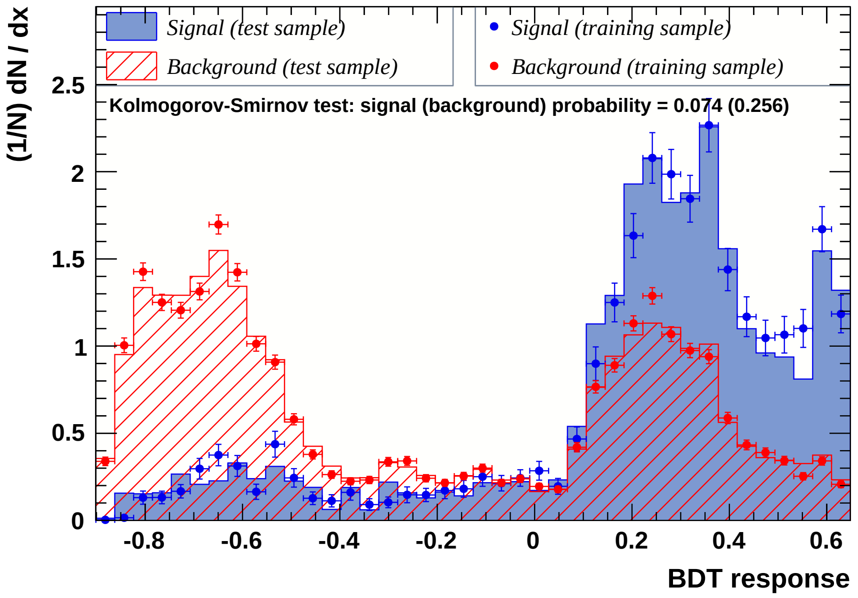
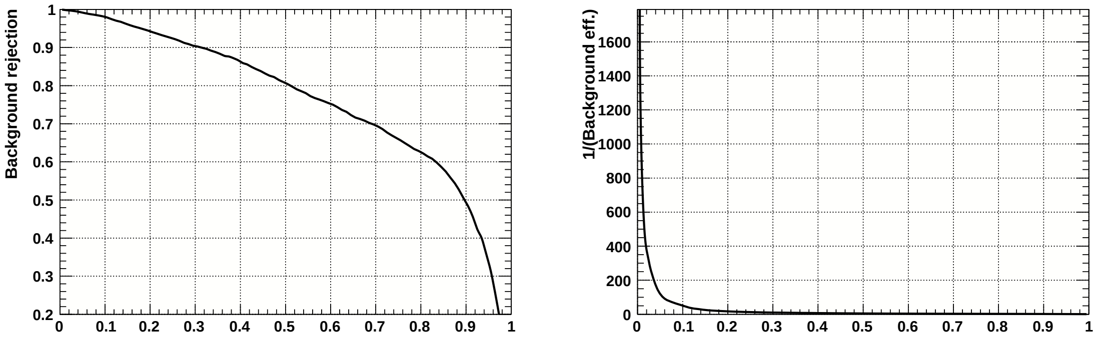


Figure 8.5.: Response of the final BDT training based on all events, with signal events depicted in blue and background events in red. The training samples (dots with error bars) and the test samples (filled histograms) show a good agreement within statistical errors. The separation power is clearly improved, compared to the individual variable inputs. The distributions are based on `PROTOS` datasets.

With these inputs, the BDT produces the response distributions of Figure 8.5. To avoid cluttering of the plot, no statistical uncertainties are drawn for the test sample, but they have nearly the same height as the training samples. Thus, the distributions agree quite well within statistical uncertainties and no indication for overtraining is visible which is supported by the Kolmogorov-Smirnov test with probability values of 0.074 for the signal and 0.256 for the background distributions. Small discrepancies between the curves at lower response values are negligible since the major discrimination can be achieved at high values where most of the signal is present.

The discrimination power in Figure 8.5 shows a nice behaviour and is strongly improved compared to the individual input variables, which can be further assessed by the curve of Figure 8.6 (a), displaying the signal efficiency vs. the background rejection. Moreover, the signal efficiency vs. the inverse background efficiency is shown in Figure 8.6 (b) which can be interpreted as the mis-identification rate when searching for $t \to s + W$ decays.

In Figure 8.7, the expected significances for cuts on the BDT response variable, are drawn. The best cut at 0.05 for an approximated statistical significance yields a maximum of $S/\sqrt{S+B} \approx 0.59$ when assuming 100 signal (S) and 50 000 background (B) events. This result includes only the main background contributions from $t\bar{t}$ and W + *jets* and does not account for any systematic uncertainties. These two aspects can reduce the final significance substantially and are considered in chapter 9.



(a) (b)

Figure 8.6.: (a) Signal efficiency vs. background rejection rate. (b) Signal efficiency vs. the inverse background efficiency for the BDT output which can be interpreted as the mis-identification rate when searching for $t \rightarrow s + W$ decays.

An application of this BDT response to the final `MC@NLO` based selection, including data events from the 2012 run, yields the distribution of Figure 8.8 (a), with a good agreement between data and MC simulated events. The relative distributions for signal and backgrounds are drawn on the right-hand side (b). The colour code is identical to the left diagram, just the background distribution of b-quark probe jets is drawn in black instead of light grey. For a better clarity, no MC uncertainties are included but the visible differences are indeed significant due to the large bin width. The QCD multijet events (purple) can result in negative entries caused by the weights in the matrix method (cf. section 4.3), thus the sum of its visible normalised entries can be larger than one. For comparison, the data-to-MC distribution of the BDT response for `PROTOS` generated events is depicted in Figure A.12 (a). As expected, the agreement is quite vague, especially at high values which is caused by the missing NLO terms.

To conclude, the $t \rightarrow s + W$ signal events (red) show a clear maximum at a relatively high BDT response value (Figure 8.8 (b)). The subsequent template fit method, introduced in section 9.2, is using the shape of the different distributions. Therefore, even though a large number of $t \rightarrow b + W$ events (black) and W + *jets* events (light orange) are present in the same region, the signal events should be recognisable since the $t \rightarrow b + W$ background has a clear second peak at low response values and the W + *jets* background is slightly shifted to smaller values compared to the signal. All other backgrounds are either negligible due to their clearly different shape or due to their low number of events compared to the dominant backgrounds and are thus illustrated with dashed lines.

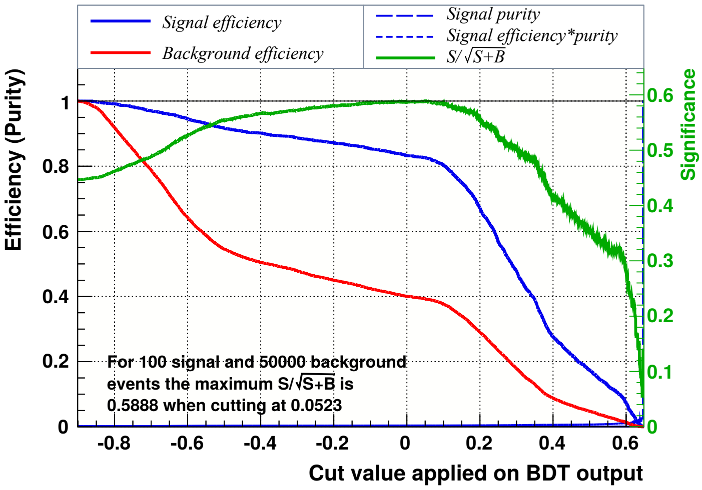
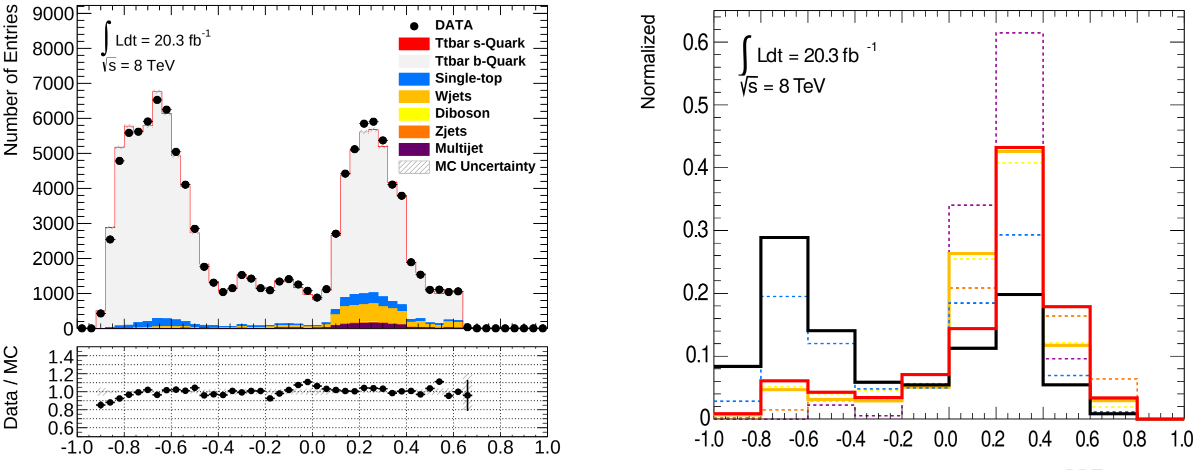


Figure 8.7.: The significance for a possible cut based analysis is drawn in green, based on the `PROTOS` datasets. The signal (blue line) and background efficiency curves (red line) are also drawn with respect to a possible cut value. A cut at around 0.05 would lead to a maximum statistical significance of about 0.59, neglecting the systematic uncertainties and further backgrounds.



(a)

(b)

Figure 8.8.: BDT response based on `MC@NLO` generated $t\bar{t}$ events. (a) The combination with data events results in a nice data-MC agreement, which is based on the full 2012 dataset with $\int \mathcal{L}\, dt = 20.3\, fb^{-1}$. (b) Relative distributions for all signal and background datasets. The colour code is identical to (a) except b-quark probe jets drawn in black. For a better clarity, only the dominant backgrounds are drawn in solid lines and no MC uncertainties are included. The visible differences are indeed significant due to the large bin width.

8.5. BDT response for events with a K_S^0 candidate

For events with the cut $N(K_S^0) \geq 1$, the discriminating variables from Table 8.1 combined with Table 8.2 are used. The values summarised in Table 8.4 are chosen as the classifier parameters. The corresponding linear correlation coefficients are displayed in the matrices of Figure 8.3.

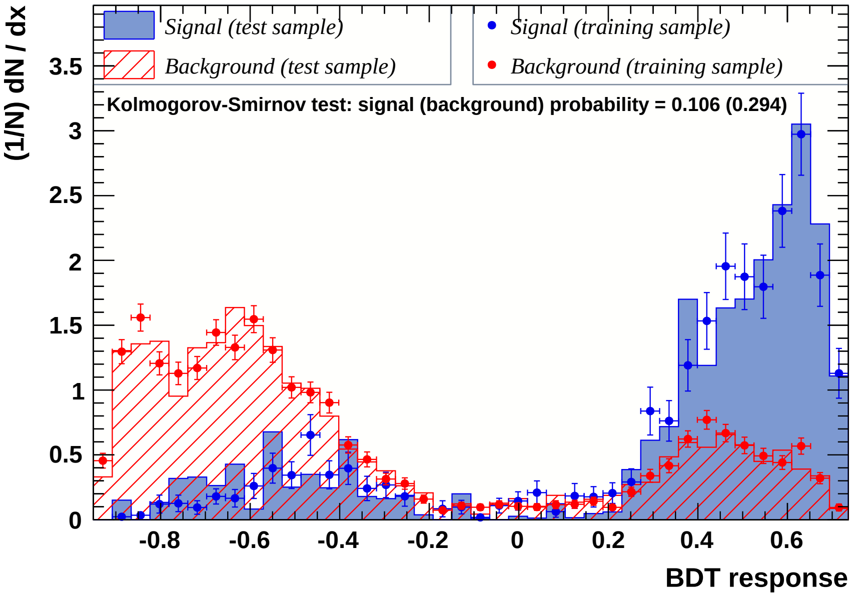
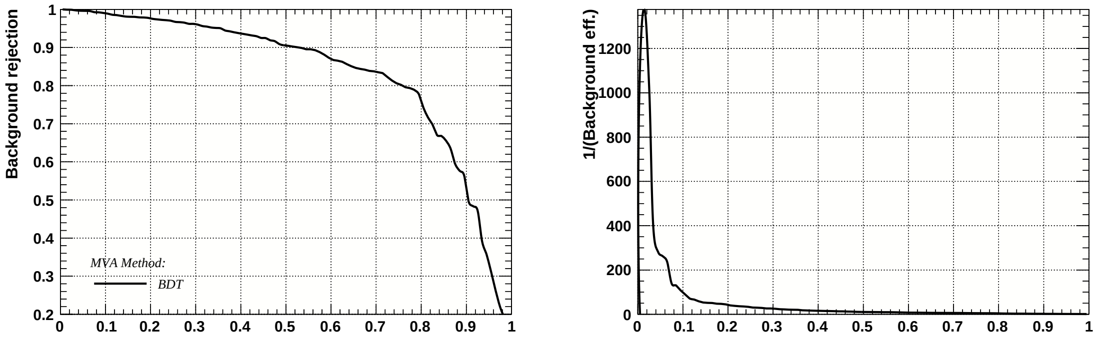


Figure 8.9.: Response of the BDT algorithm based on `PROTOS` generated events with at least one K_S^0 candidate, with signal events depicted in blue and background events in red. The training samples (dots with error bars) and the test samples (filled histograms) show a good agreement within statistical errors. The separation power is better than the BDT response of Figure 8.5 without a required K_S^0 candidate.

With these settings, the BDT response results in Figure 8.9. Again, no statistical uncertainties are drawn for the test sample, but they have nearly the same height as the training samples. Thus, the distributions agree well within statistical uncertainties and no indication for overtraining is visible as supported by the Kolmogorov-Smirnov test with probability values of 0.106 for the signal and 0.294 for the background distributions. The discrimination power shows even better results than those without a cut, which is emphasised by the corresponding comparison between signal efficiency and background rejection (Figure 8.10 (a)) and the signal efficiency vs. the inverse background efficiency (Figure 8.10 (b)).

In Figure 8.11, the expected significances for cuts on the BDT response variable, are drawn. The best cut at 0.72 for an approximated statistical significance yields a maximum of $S/\sqrt{S+B} \approx 0.52$ when assuming 21 signal (S) and 12 000 background (B) events. However, this cut is a result of the peak structure at high values, which might be attributable to statistical fluctuations. Therefore, the more reliable cut lies at a BDT output of 0.52 with a maximum statistical significance of $S/\sqrt{S+B} \approx 0.34$. This value is lower than in Figure 8.7 due to the expected lower number of events caused by the additional K_S^0 candidate. As before, this includes only the main background contributions from $t\bar{t}$ and W + *jets* and does not account for any systematic uncertainties.



(a) **(b)**

Figure 8.10.: (a) Signal efficiency vs. background rejection rate. (b) Signal efficiency vs. the inverse background efficiency for the BDT output which can be interpreted as the misidentification rate when searching for $t \rightarrow s + W$ decays.

An application of this BDT response to the final `MC@NLO` based selection with at least one K_S^0 candidate, including data events from the 2012 run, yields the distribution of Figure 8.12 (a). A reasonable agreement between data and MC simulated events is achieved, especially in the two important peak regions. The relative distributions for signal and backgrounds are drawn on the right side in Figure 8.12 (b). The colour code is in principle identical to the left diagram, only the background distribution of b-quark probe jets is drawn in black instead of light grey. Again, no MC uncertainties are shown but the visible differences are significant. The QCD multijet events (purple) can also result in negative entries, thus the sum of its visible normalised entries can be larger than one. For comparison, the data-to-MC distribution of the BDT response for `PROTOS` generated events is depicted in Figure A.12 (b). As expected the agreement is moderate, especially at high values which is caused by the missing NLO terms.

For events which require at least one K_S^0 candidate, the $t \rightarrow s + W$ signal events (red) show an even more distinct response at higher values than without a K_S^0 cut since the main background of $t \rightarrow b + W$ events (black) is strongly reduced in the signal region. The W + *jets* background (light orange) however shows a quite similar shape as the signal events which could lead to more vague results since the subsequent template fit method (section 9.2) uses the shape of the distributions for the discrimination. The lower number of events caused by the additional K_S^0 cut might worsen the final template result.

All other backgrounds, which are not discussed here, are either negligible due to their clearly different shape or due to their low number of events compared to the dominant backgrounds and are thus illustrated with dashed lines.

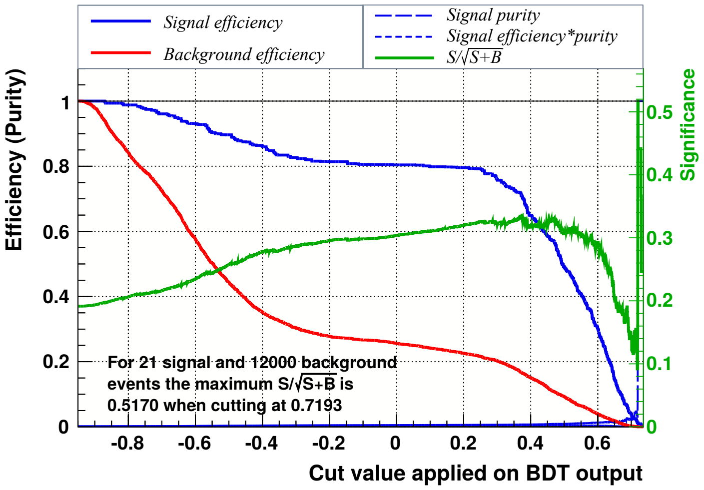
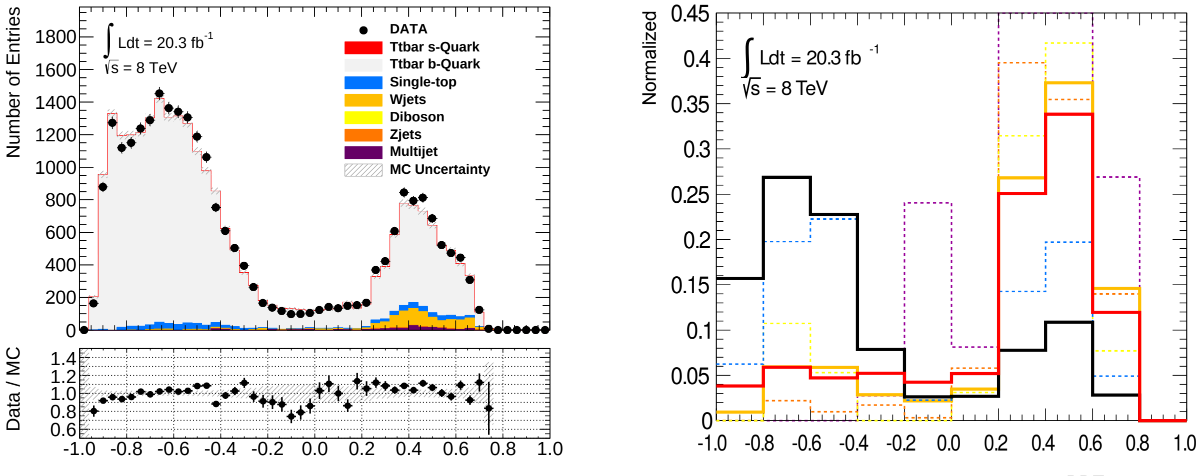


Figure 8.11.: The significance for a possible cut based analysis is drawn in green, based on the `PROTOS` datasets. The signal (blue line) and background efficiency curves (red line) are also drawn with respect to a possible cut value. The indicated cut at 0.72 with statistical significance of about 0.52 is a result of the peak structure at high values, which might be attributable to statistical fluctuations. Therefore, the more reliable cut lies at a BDT output of 0.52 with a maximum statistical significance of about 0.34, neglecting the systematic uncertainties and further backgrounds.



(a) (b)

Figure 8.12.: BDT response based on `MC@NLO` generated $t\bar{t}$ events when applying a cut $N(K_S^0) \geq 1$. (a) The combination with data events results in a reasonable data-MC agreement, which is based on the full 2012 dataset with $\int \mathcal{L}\, dt = 20.3\, fb^{-1}$. (b) Relative distributions for all signal and background datasets. The colour code is identical to (a) except b-quark probe jets drawn in black. For a better clarity, only the dominant backgrounds are drawn in solid lines and no MC uncertainties are included. The visible differences are indeed significant due to the large bin width.

9. Results

Based on the BDT output response, all ingredients can now be assembled. For the final results, both sources of uncertainties, statistics and systematics are taken into account. The statistical uncertainty can easily be calculated by regarding the number of events in the signal region which passed the selection criteria. Systematic uncertainties however arise from parameters which can not be determined confidently and thus require detailed studies of the influences on the final result, as summarised in section 9.1. Afterwards, the template fit method is explained in section 9.2, which takes the BDT response and includes all uncertainties as nuisance parameters, aiming for the best description of the data events by varying the signal and background samples within given constraints. The output can thus be used for a limit calculation that leads to the final results (section 9.3), which are subsequently compared to similar CKM matrix measurements in section 9.4. All details discussed in this chapter are based on `MC@NLO` generated $t\bar{t}$ events.

9.1. Systematic uncertainties

For the final analysis, the effects of systematic uncertainties have to be considered. Apart from changes in the event yields, this includes shape variations of the distributions, which is even more complex. The main sources are mainly detector effects, theoretical predictions or selection cuts, which can not be modelled perfectly. For each systematic parameter, the relevant weight or scaling factor is varied up and down with respect to the default value, where applicable, by 1 standard deviation ($\pm 1\sigma$). In addition, also data fluctuations can contribute to the final uncertainty values, which can either be due to internal calibrations (e.g. in JES components) or due to the usage of data samples in the derivation of backgrounds (e.g. QCD multijet). Furthermore, some systematics like those from the luminosity or the MC normalisation have individual variations, which have been evaluated by different ATLAS working groups.

For each of these variations, the full analysis is repeated, including the event selection and reconstruction steps, in order to determine the final influence. The resulting values are stated with respect to the variation of the event yields in the signal region of the BDT output, which is represented by a positive BDT response value in Figures 8.8 and 8.12. All sources of systematic uncertainties, which are integrated in the template fit, are listed in the following. The corresponding assumptions are in agreement with the recommendation of the ATLAS top working group [67].

- **Integrated luminosity**
  The integrated luminosity of the 2012 Run I period at $\sqrt{s} = 8$ TeV was evaluated to be 20.28 fb^{-1}, including an uncertainty of ±2.8 %, as described in section 4.4.1 and [70, 71]. This value has a direct effect on the amount of signal events and thus is included as an individual systematic uncertainty.

- **$t\bar{t}$ modelling**
  The simulation of $t\bar{t}$ events at NLO, including $t \to s + W$ decays, can only be achieved with the `MC@NLO` generator. To take the modelling and theory uncertainties into account a total normalisation uncertainty of ±11 % is included. This covers also the uncertainties of ISR/FSR modelling, parton distribution functions and non-perturbative processes like the underlying event descriptions and the colour reconnection, including hadronisation effects. [121]

- **Additional background modelling**
  All background models include different uncertainties due to the different MC generators and theoretical bases.
  The description of the cross section of the Single Top MC generator is assumed to have a normalisation uncertainty of ±7 % [122].
  For the MC generated backgrounds W+jets and Z+jets, theoretical normalisation uncertainties of 4 % [123] as well as 24 % for each additional jet [124] have to be incorporated, which is added in quadrature. With typically four jets in the signal region (section 4.3), a total normalisation uncertainty of ±48 % for W+jets and Z+jets results.
  Similar, the di-boson events arise with a theoretical uncertainty of 5 % as well as 24 % for each extra jet [125]. In this case, events with two additional jets are the most dominant processes (section 4.3), resulting in an uncertainty of ±34 % for di-boson+jets.
  For the QCD multijet background, a variation of 50 % with respect to its default value is chosen. This assignment is based on the outcomes of alternative methods for an estimation of QCD multijets events [126]. The relatively large uncertainty is due to the data driven approach, which is detailed in section 4.3 and [65, 66].

  All background modelling uncertainties have a limited effect on the final results, due to the high purity of $t\bar{t}$ events. Moreover, the template fit also rescales the distributions to data all over the parameter space, including the signal region (positive BDT response) and the background region (negative BDT response), which reduces the modelling effects even further.

- **Jet related uncertainties**
  In the $|V_{ts}|$ analysis with four to six jets, the correct understanding of the jet

reconstruction procedure is crucial, which also includes the understanding of all uncertainties.

The uncertainty associated to the jet reconstruction efficiency (**JEff**) is derived by a comparison of the default calorimeter based jets (cf. section 5.4) with jets reconstructed from charged tracks [85]. No large discrepancies are seen since most differences arise for low-p_{T} jets only [127], which are excluded by the reconstruction steps. Consequently, the JEff uncertainty in the signal region is found to be smaller than 0.1 %.

The determination of the jet energy resolution (**JER**) uncertainty is based on a comparison between truth and reconstructed resolutions of simulated jet energies. At this, each truth jet energy is smeared with a Gaussian width similar to the jet energy resolution [128], which results in a symmetric uncertainty of ±2.1 %.

The Jet Vertex Fraction (**JVF**) systematic uncertainty is obtained by increasing and decreasing the cut value of the JVF cut as detailed in section 5.4, which is strongly related to the effect of pile-up modelling (cf. section 5.5). The difference between the variations and the default value is taken as the systematic uncertainty, resulting in a range of -1.0 % to $+2.3$ %.

The last jet related uncertainty is due to the jet energy scale (**JES**) (section 5.4). It is composed of 22 sub-effects, which can be classified in the following uncertainty categories: η intercalibration (difference between detector and calibrated η), single particle (behaviour of jets with high-p_{T} single particles), flavour (composition and response of gluon and light quark jets), bJES (b-jet energy scale), pile-up and so called effective uncertainties. The latter one can again be split into statistical, detector and physics modelling categories. A detailed description of all components is given in [129]. For each of those, a scaling procedure of one standard deviation is performed, basically describing an up and down shift of the jet p_{T} by 1σ with a subsequent re-execution of the full analysis. All described sub-effects of the JES uncertainties, added in quadrature, result in a total uncertainty of ±9.5 %.

- **b-tagging uncertainty**
  For the correct assignment of b-tags, a data-to-MC calibration is used, based on a combination of three different methods (section 5.6). The corresponding scale factors and efficiencies are varied by applying separate shifts of $\pm 1\sigma$ on the factors of b-tagged jets (**btagSF**), c-tagged jets (**ctagSF**) (including τ leptons identified as jets) as well as on the mistagging rates of light jets (**mistagSF**). The corresponding uncertainties in the signal region lead to effects of -1.9 % to 1.6 %, ±0.2 %, as well as -0.2 % to 0.1 %, respectively.

- **Uncertainties related to charged leptons (e, μ)**
  To obtain a better data-to-MC agreement, the kinematic properties of charged leptons (e, μ) are calibrated (sections 5.1 and 5.2). This procedure is accompanied by additional uncertainties which include the electron energy scale (**EES**) and

the muon momentum scale (**MuSc**) uncertainties, resulting in absolute values of 0.2 % and smaller. Furthermore, the limited energy and momentum resolution is accounted for, by smearing the central values, which results in an uncertainty component for the electron energy resolution (**EER**), and two uncertainty components for muons, one for the resolution of the inner detector (**MuID**) and one for the resolution of the muon spectrometer (**MuMS**). All absolute uncertainty values are 0.1 % or smaller.
In addition to the correction of the kinematic properties, the uncertainties of the lepton trigger efficiency, identification and reconstruction are considered (sections 5.1 and 5.2). The three corresponding scale factors are once more varied by one standard deviation, which results in uncertainties of ±0.7 % for the trigger scaling factor (**LepTrigSF**), ±1.3 % for the identification scaling factor (**LepIdSF**) and ±0.2 % for the reconstruction scaling factor (**LepRecoSF**).

- **Uncertainties related to the missing transverse energy**
  The $\not{E}_T$ is determination is based, among others, on the p_{T} sum of all reconstructed physics objects ("hard terms") in an event (cf. section 5.3). Thus, all above discussed uncertainty aspects are propagated to the $\not{E}_T$ components. Apart from those "hard terms" also so called "soft terms" (e.g. low energetic calorimeter deposits) contribute to the $\not{E}_T$ calculation and influence the uncertainty. This includes scale factor (**$\not{E}_T$ sc_soft**) and resolution (**$\not{E}_T$ res_soft**) aspects of the soft term's energy deposits, resulting in uncertainties of ±0.2 % and lower than 0.3 %, respectively.

All uncertainty values stated above are with respect to events without any K_S^0 cut. A summary, including also the uncertainties for events requesting at least one K_S^0 candidate, is given in Table 9.1. Moreover the **data statistic** uncertainty is included, with a value of ±0.5 % in the signal region.

Table 9.1.: Overview of all uncertainties included in the template fit as nuisance parameters. The values are stated with respect to the variation of the event yields in the signal region of the BDT output, which is represented by a positive BDT response value in Figures 8.8 and 8.12. For uncertainties labelled with a dash (-), the absolute uncertainty values are smaller than 0.1 %.
All terms describing uncertainties of model normalisation, luminosity as well as JEff and JER are symmetric by definition. For the data statistic, the effect in the signal region and the effect on all events are labelled separately.

Source	No cut (↓)	No cut (↑)	$N(K_S^0) \geq 1$ (↓)	$N(K_S^0) \geq 1$ (↑)
Luminosity	symmetric	2.8 %	symmetric	2.8 %
$t\bar{t}$ model	symmetric	11.0 %	symmetric	11.0 %

Source	No cut (↓)	No cut (↑)	$N(K_S^0) \geq 1$ (↓)	$N(K_S^0) \geq 1$ (↑)
SgTop model	symmetric	7.0 %	symmetric	7.0 %
W+jets model	symmetric	48.0 %	symmetric	48.0 %
Z+jets model	symmetric	48.0 %	symmetric	48.0 %
Di-boson model	symmetric	34.0 %	symmetric	34.0 %
QCD method	symmetric	50.0 %	symmetric	50.0 %
JEff	symmetric	< 0.1 %	symmetric	0.2 %
JER	symmetric	2.1 %	symmetric	0.3 %
JVF	- 1.0 %	2.3 %	-1.0 %	1.7 %
JES	- 9.5 %	9.5 %	-4.7 %	3.7 %
btagSF	-1.9 %	1.6 %	-2.1 %	1.9 %
ctagSF	-0.2 %	0.2 %	-0.1 %	0.1 %
mistagSF	-0.2 %	0.1 %	-0.2 %	0.2 %
EES	-0.2 %	-	-0.3 %	-
EER	-	-	-0.3 %	-
MuSc	-	-	-0.1 %	-
MuID	-	0.1 %	-0.2 %	-
MuMS	-	-	-0.1 %	0.1 %
LepTrigSF	-0.7 %	0.7 %	-0.7 %	0.7 %
LepIdSF	-1.3 %	1.3 %	-1.3 %	1.3 %
LepRecoSF	-0.3 %	0.3 %	-0.3 %	0.3 %
$\not{E}_T$ sc_soft	-0.2 %	0.2 %	-0.3 %	-
$\not{E}_T$ res_soft	-	0.3 %	-0.4 %	-
Data statistic (all events)	symmetric	0.3 %	symmetric	0.6 %
Data statistic (signal region)	symmetric	0.5 %	symmetric	1.2 %

To incorporate adequately the impact of all individual uncertainties in the final result, these enter the fit procedure as nuisance parameters of the likelihood calculations (cf. section 9.2). The subsequent total uncertainty can only be determined after the execution of the fit, where a concurrent optimisation of all input parameters, including uncertainties, is performed, which also takes correlation effects into account. The final influences of the uncertainties on that fit are shown below in Table 9.3, clustered in categories due to the low effects of the individual contributions. For the limit setting,

the uncertainties are included in the overall final limit value.
Apart from the above discussed values, also the influences of the χ^2 and p-values of the top pair finder (cf. section 6.3) and the K_S^0 finder (cf. section 7.3) have principally to be considered, but different tests do not show any significant effects on the BDT response and thus on the final result. One reason might be the good data-MC agreement as shown in section 6.3 and section 7.3. Therefore, these possible uncertainties are omitted in the final analysis step.

9.2. Template fit and limit calculation

The final results of the $|V_{ts}|$ analysis are obtained by a template fit procedure. The corresponding templates are based on the individual signal and background distributions of the BDT response, displayed in Figures 8.8 and 8.12. By varying all functions' rates, the best fit to the data can be achieved.
To determine the actual signal, the *parameter of interest* μ, is introduced. By definition it describes the signal strength, with $\mu = 1$ representing the theoretically expected number of signal and background events, whereas $\mu = 0$ represents a background-only model. In the $|V_{ts}|$ analysis, it is used for a description of the cross section of $t \rightarrow s + W$ decays with respect to the PDG [1] value. Thus, also hypotheses with a μ-value above 1 are possible, which would indicate a larger than expected cross section value.

$$\sigma^{\text{measured}} = \mu \cdot \sigma^{\text{PDG}} \tag{9.1}$$

The aim of this procedure is either to determine the best fitted value μ for the signal strength, which results in a direct measurement of $|V_{ts}|$, or to set a strong limit on μ so that no adequate fit to data can be performed anymore within the constraints. The latter one can then be used for an upper limit on the $|V_{ts}|$ value. Compared to a cut based analysis, this approach considers the full range of the BDT response, including the negative control region where background information is present, which is normally lost in a cut based analysis.
During the fit, the overall rate of each background can be derived. This is especially important for the W+jets distribution, since W+jets MC calibration, as proposed by the top working group [67], was not implemented in the precedent analysis steps. Now, with a variation of the background distribution to data, the corresponding scaling factor is directly implied in the final analysis, which results in an equivalent output as the proposed calibration. This principle can also be transferred to $t\bar{t}$ events.

All descriptions of this section are only explained in a compressed and simplified way. A more detailed explanation of these analysis steps can be found in the “HistFactory” manual [130], the “RooStats” documentation [131] and in a description of likelihood-based statistical tests of G. Cowan et al. [132].

The method starts with the distributions of the BDT response which are employed to extract probability density functions, further used as templates. From these, a likelihood function $\mathcal{L}$ is build, which describes all parameters of the different models, including the normalisation of the distributions, the signal strength parameter μ and the uncertainties of section 9.1. The inputs for $\mathcal{L}$ are all based on either Gaussian constraints for the systematic uncertainties or Poisson distributions $\mathcal{P}$ for the remaining event parameters. Combined, a descriptive equation for the full model emerges, which can be written as

$$\mathcal{L}\left(x_e, a_p|\mu, \alpha_p\right) \;=\; \mathcal{P}(n_{tot}|\mu S + B)\;\cdot \prod_{e\,\in\,events}\left(\frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B}\right)\cdot \prod_{p\,\in\,uncert} f_p(a_p|\alpha_p). \tag{9.2}$$

The Gaussian based terms $f_p(a_p|\alpha_p)$ describe the uncertainties, which are included as auxiliary measurements a_p constrained by the actual nuisance parameters α_p. The number of signal events S and the number of background events B are modelled by the probability density functions to form the shapes $f_S(x_e)$ and $f_B(x_e)$, with x_e representing the discrimination observable for the events e. The data are described by n_{tot}. [130]

It is noteworthy that technically all distributions are used as binned distributions instead of continuous ones, like stated in Equation 9.2, which results in more complex formulas. Nevertheless, the likelihood function $\mathcal{L}$ can be still described correctly. Moreover, shape variations are included in these more complex discrete description, since some of the systematic uncertainties can change the shape of the input samples, for which a detailed study was performed in [121]. Based on this, the following samples are varied in shape with respect to the corresponding systematic uncertainties:

Samples	Systematic uncertainties with shape influences
$t\bar{t}$	JES
$t\bar{t}$, Single top	$\not{E}_T$ sc_soft
$t\bar{t}$, Single top, W+jets	$\not{E}_T$ res_soft

The shape variances of all other uncertainties are smaller than the statistical fluctuations of the MC generation and can thus be neglected. Only the rate variations of the templates have then to be considered. [121]

All the above described steps, to obtain $\mathcal{L}$, are based on the "HistFactory“ method [130].

With the final likelihood function $\mathcal{L}$ for the model descriptions available, hypothesis tests can now be performed by a comparison to the observed data, while also considering the correlations between the uncertainties. This analysis step is based on the “RooStats” package [131]. For the signal significance test, the null hypothesis $H(b,\ \mu = 0)$ (describing background only) should be rejected whereas the alternative, competitive hypothesis $H(s+b$, with μ not fixed) (describing an additional signal) should be proved. To have an indication, if the observed data is more signal-like or more background-like, the hypotheses are tested using the following test statistic t_μ:

$$t_\mu = -2 \cdot \ln \lambda(\mu) \qquad \text{with} \qquad \lambda(\mu) = \left.\frac{\mathcal{L}(\mu, \boldsymbol{\alpha})}{\mathcal{L}_{max}}\right|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}_\mu} \tag{9.3}$$

$\mathcal{L}_{max}$ indicates the best fit of the likelihood function with no constraint on μ or the nuisance parameters $\boldsymbol{\alpha}$. $\mathcal{L}(\mu, \boldsymbol{\alpha})$ instead is fitted with a fixed μ value, which is either signal or background like. The nuisance parameters $\boldsymbol{\alpha}$ are set to have the best values $\hat{\boldsymbol{\alpha}}_\mu$ for an optimum fit, taking the fixed signal strength μ into account. If the so called *profile likelihood ratio* $\lambda(\mu, \boldsymbol{\alpha})$ [131] is close to unity (t_μ small), a good agreement between the observed data and the hypothesis is achieved, values closer to zero (t_μ large) indicate higher disagreement. To transfer this result to a probability value, that describes the level of disagreement between data and hypothesis, the following p-value is introduced:

$$p_\mu = \int\limits_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu) dt_\mu \tag{9.4}$$

with $t_{\mu,obs}$ being the test statistic t_μ from observed data and $f(t_\mu|\mu)$ being the probability density function of t_μ with respect to the signal strength μ. [132]

To determine if a positive signal $\mu > 0$, i.e. a measurable signal cross section, is present in the observed data, it is attempted to reject the null hypothesis ($\mu = 0$). The resulting background-only p_0-value can then be converted to a signal significance value Z. [131]

To declare a limit on the signal strength μ and thus on the $|V_{ts}|$ value, the roles of null model and alternative model are exchanged for the significance test. The null model is then defined as $H(s+b)$, the alternative as $H(b)$. The resulting two p-values can subsequently be expressed with respect to the confidence level of the background hypothesis $CL_b = 1 - p_{\mu=0}$, and of the signal + background hypothesis $CL_{s+b} = 1 - p_\mu$. The subsequent upper limit of the signal strength μ_{obs} is then defined by the confidence level

$$CL_s = \frac{CL_{s+b}}{CL_b} \quad [133]. \tag{9.5}$$

An hypothesis is excluded, if the CL_s value exceeds 95 %, equivalent to a p-value of below 0.05. It is extracted by increasing the signal strength parameter μ successively and performing multiple confidence level checks, until this 95 % limit is exceeded.
For the $|V_{ts}|^2$ exclusion limit, the rescaling is based on the PDG input value of 0.16 % [1], which is multiplied by increasing factors of μ until $CL_s = 0.95$ is reached. This final value is then called μ_{limit}.

9.3. $|V_{ts}|$ limit and final results

Having all inputs and techniques discussed, this culminates in the final results, which are all interpreted with respect to SM decay topologies of $t\bar{t} \to WWb\bar{s}$ and $t\bar{t} \to WWb\bar{b}$. Both systematic and statistical uncertainties are included here.

112 171 events, thereof an estimated number of 11 204 background events, passed the $t\bar{t}$ selection criteria and are used for the full analysis, as stated in section 6.4. In these candidate events, approximately 183 $t\bar{t} \to WWb\bar{s}$ decays are expected. The event numbers are all normalised to a total integrated luminosity of 20.3 $\text{fb}^{-1} \pm 2.8$ %, which is derived in section 4.4. With the s-quark discrimination studies of chapter 7 and the subsequent multivariate analysis of chapter 8, the signal and background events can be separated. The template method of section 9.2 is then the last analysis step which reveals the final results.

The outcome can be interpreted as the signal significance Z, the cross section σ_s of $t \to s + W$ decays, and the ratio

$$\mathcal{R}_s = \frac{\text{BR}(t \to W + s)}{\text{BR}(t \to W + d, s, b)}, \quad \text{which is equal to} \quad |V_{ts}|^2, \tag{9.6}$$

when assuming a unitary CKM matrix. Here, the contribution from $|V_{td}|^2$ is neglected, since it is expected to be about 20 times smaller than $|V_{ts}|^2$ [1].

Results for all events

Taking the BDT output from section 8.4, describing all observed data events without any K_S^0 cut, a probability value of $p = 0.28$ and a significance for $t \to s + W$ decays of

$$Z = 0.59\ \sigma \tag{9.7}$$

are obtained for the best fitted signal strength of $\mu = 2.4^{+2.9}_{-2.4}$.
Assuming a total cross section $\sigma(t\bar{t}) = 252.9^{+13.1}_{-14.2}$ pb for all decay channels [134, 135] and a PDG value $|V_{ts}|^2_{\text{PDG}} = 0.16\ \% \pm 0.01\ \%$ [1], the corresponding measured CKM matrix element and cross section thus yield $|V_{ts}|^2 = 0.40^{+0.53}_{-0.40}\ \%$ and $\sigma_s = 1.0^{+1.3}_{-1.0}$ pb.

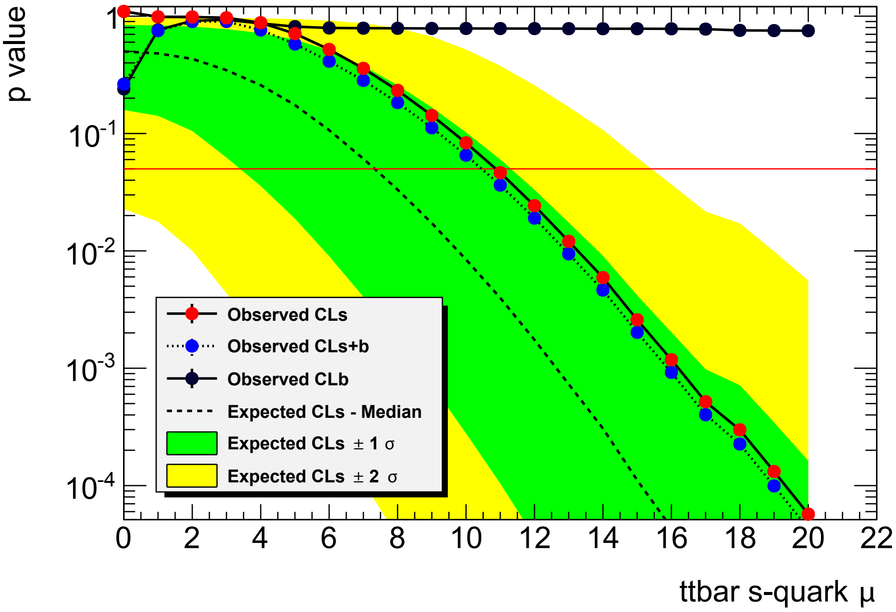


Figure 9.1.: Exclusion limit for all events. The red line represents the p-value of 0.05, which is equivalent to a CL_s value of 95 %. The point, at which the red-dotted CL_s curve crosses this line is the measured limit value $\mu = 10.9$.
The dashed line, combined with the corresponding green and yellow bands, indicate the expected values for CL_s with 1σ and 2σ deviations, based on “Asimov” [136] generated events.

The corresponding 95 % CL_s upper limit with respect to the maximum signal strength is computed to be

$$\mu_{limit} = 10.9 \tag{9.8}$$

The upper limit on the CKM matrix element and the cross section thus yield in the end

$$|V_{ts}|^2 < 1.74\ \% \quad \text{and} \quad \sigma_s < 4.4\ \text{pb}, \tag{9.9}$$

both at 95 % confidence level. An expected limit, based on “Asimov” [136] generated events for $\mu_{\text{input}} = 1$, is found to be $\mu_{exp.limit}\,(\pm 1\sigma) = 7.4^{+4.0}_{-3.9}$ or $\mu_{exp.limit}\,(\pm 2\sigma) = 7.4^{+8.0}_{-7.4}$ as illustrated in Figure 9.1.

Results for events with at least one K_S^0 candidate

For events with at least one K_S^0 candidate, the BDT outputs from section 8.5 are used as templates. They result in a probability value of $p = 0.24$ and a significance for $t \to s+W$ decays of

$$Z = 0.70\ \sigma \tag{9.10}$$

for the best fitted signal strength of $\mu = 4.5^{+6.4}_{-4.5}$.
Assuming again $\sigma(t\bar{t}) = 252.9^{+13.1}_{-14.2}$ pb for all decay channels and $|V_{ts}|^2_{\text{PDG}} = 0.16\ \% \pm 0.01\ \%$, the corresponding measured CKM matrix element and cross section thus yield $|V_{ts}|^2 = 0.73^{+1.05}_{-0.73}\ \%$ and $\sigma_s = 1.8^{+2.7}_{-1.8}$ pb.

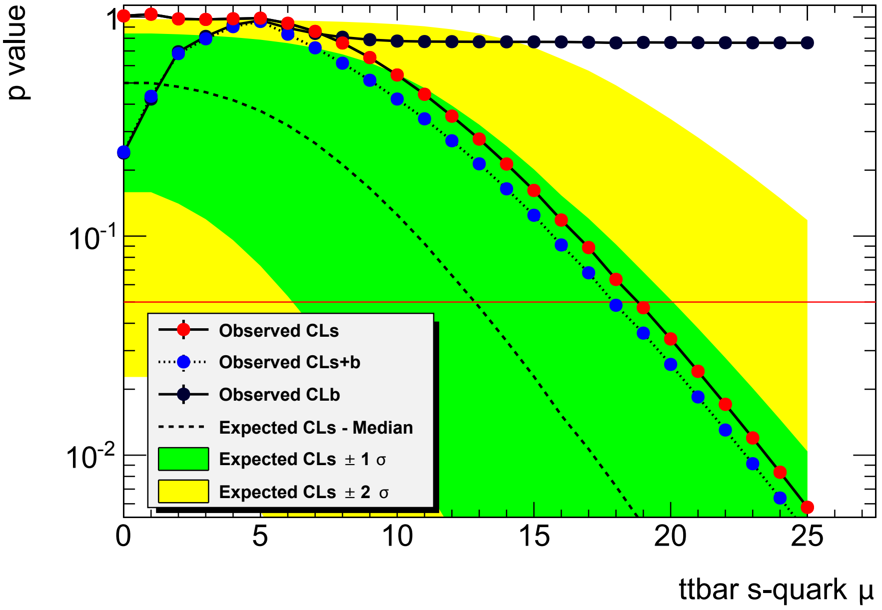


Figure 9.2.: Exclusion limit for events with at least one K_S^0 candidate. The red line represents the p-value of 0.05, which is equivalent to a CL_s value of 95 %. The point, at which the red-dotted CL_s curve crosses this line is the measured limit value $\mu = 18.8$.
The dashed line, combined with the corresponding green and yellow bands, indicate the expected values for CL_s with 1σ and 2σ deviations, based on “Asimov” [136] generated events.

The corresponding 95 % confidence level upper limit with respect to the maximum signal strength is computed to be

$$\mu_{limit} = 18.8 \tag{9.11}$$

The upper limit on the CKM matrix element and the cross section thus yield in the end

$$|V_{ts}|^2 < 3.09\ \% \quad \text{and} \quad \sigma_s < 7.8\ \text{pb}, \tag{9.12}$$

both at 95 % confidence level. An expected limit, based on “Asimov” generated events, is found to be $\mu_{exp.limit}\,(\pm 1\sigma) = 12.9^{+6.9}_{-6.7}$ or $\mu_{exp.limit}\,(\pm 2\sigma) = 12.9^{+10.8}_{-11.9}$ as illustrated in Figure 9.2.

The less strict limit from events with at least one K^0_S candidate can be explained by the lower number of available $t \to s + W$ decays in the final selection since only about 22 % of all events are available (cf. Figure 7.13 (a)). Even though a better separation strength is achieved in section 8.5, it can not compensate this effect.

No uncertainties have to be stated for the two obtained limit values, since, by definition, limit values integrate all aspects of an hypothesis in the final value of 95 % CL_s, including uncertainties. The individual influence of a systematic uncertainty can, however, still be obtained, by repeating the limit calculation without the uncertainty in question. Since the single influences are small, these results are combined with respect to their origin for reason of better clarity (Table 9.3).

Table 9.3.: Influences of systematic uncertainties on the final limits, for events with no cut and events with at least one K^0_S candidate. Since the single contributions are small, the uncertainties are clustered in categories. It can be clearly seen, that the template fit method reduces absolute uncertainties by at least a factor of two compared to the BDT response (cf. Table 9.1). Negative values indicate a lower, i.e. more stringent limit when omitting this group of uncertainties.

Source of uncertainty	No cut	$N(K^0_S) \geq 1$
Luminosity	– 0.2 %	< 0.1 %
All signal and backround models	– 7.4 %	1.2 %
Jet algorithms (without JES)	– 0.1 %	0.3 %
JES components	4.9 %	0.5 %
b-tagging algorithm	– 0.4 %	0.1 %
Leptons (Charged leptons and $\not{E}_T$)	– 0.4 %	−0.8 %

To summarise, all obtained results and the first direct limit of $|V_{ts}|^2 < 1.74\ \%$ at 95 % CL_s are compatible within the SM predictions and the PDG values [1].

9.4. Comparison with other measurements

This $|V_{ts}|$ analysis, based on the detection of $t\bar{t} \to WWb\bar{s}$ events, is the first direct analysis. Thus, no further comparisons with other direct measurements can be performed. Still, $|V_{ts}|$ can be extracted indirectly from available measurements, which are briefly discussed in the following.

B-$\bar{B}$ meson oscillations

The cited $|V_{ts}|$ value in the previous section 9.3 arises from a combination of theoretical predictions and indirect measurements (cf. section 2.3). The most precise indirect measurements are based on $B_s^0 - \bar{B}_s^0$ oscillation studies which yield a mass difference of $\Delta m_s = (17.761 \pm 0.022)\ \mathrm{ps}^{-1}$ [1]. Combined with assumptions on the $|V_{tb}|$ value and on lattice QCD calculations, this can then be interpreted as $|V_{ts}| = (40.0 \pm 2.7) \times 10^{-3}$ and $|V_{ts}|^2 = (0.16 \pm 0.02)\ \%$. Even though this value is more precise than the one of the analysis presented, it is only an indirect measurement which requires several assumptions. [1]

In addition to the $B_s^0 - \bar{B}_s^0$ oscillation, the branching ratio of rare, loop-mediated decays of K and B mesons also contribute to the PDG value of $|V_{ts}|$. However, those values have considerable uncertainties or depend on uncertain CKM matrix element combinations, like $|V_{ts}/V_{cb}|$, which both lead to less precise $|V_{ts}|$ values than the oscillation studies [1].

Extraction from top-quark measurements

Apart from the indirect measurements, Tevatron and LHC experiments performed several direct measurements of the top-quark branching fraction

$$\mathcal{R}_b = \frac{\mathrm{BR}(t \to W + b)}{\mathrm{BR}(t \to W + d, s, b)} = |V_{tb}|^2, \tag{9.13}$$

which can be interpreted with respect to $|V_{ts}|^2$. Those measurements are generally based on the multiplicity of b-tagged jets in $t\bar{t}$ events. The expected number of b-jets is modelled in terms of the b-tagging efficiency, the mistagging rate and $\mathcal{R}_b$. Events with a b-tagging multiplicity between 0 and up to 4 are considered and fitted to data, which yield the variable $\mathcal{R}_b$.
The first of these measurements was performed by the CDF collaboration, achieving a limit of $|V_{tb}| > 0.78$ at 90 % confidence level [137], which is equivalent to $|V_{tb}|^2 > 0.61$. Several further studies have been performed since then. The most recent and precise measurement was published in 2014 by the CMS collaboration, measuring a value of $|V_{tb}| > 0.975$ at 95 % confidence level [96], which corresponds to $|V_{tb}|^2 > 0.95$.

Moreover, measurements based on the single-top-quark production cross section were performed. The production process of single-top-quarks involves a Wtb vertex, which allows a direct extraction of the $|V_{tb}|$ parameter, since it describes the corresponding coupling strength. Those kind of measurements have the advantage of being relatively model independent, due to a sensitivity to anomalous form factors at the Wtb vertex. However, two basic disadvantages emerge: First, with $|V_{tb}| > 0.92$ [138], the final limit is less strict than the one from the most recent $t\bar{t}$ study, and second, an approximation of $|V_{ts}| \ll |V_{tb}|$ is used, which makes a comparison with the presented analysis difficult.

The best limit, based on top-quark measurements, is therefore $|V_{tb}|^2 > 0.95$, which can be interpreted as

$$|V_{ts}|^2 < 5.0\ \% \tag{9.14}$$

at 95 % confidence level. This limit is about 2.9 times worse than the one obtained in section 9.3 of this analysis, which is

$$|V_{ts}|^2 < 1.74\ \%. \tag{9.15}$$

So this analysis, using a direct measurement of the CKM matrix element $|V_{ts}|^2$, provides the best direct limit on $|V_{ts}|^2$ up to now.

10. Conclusion and Outlook

In the year 2012, the ATLAS detector operated with $\sqrt{s} = 8$ TeV pp-collisions and recorded a total integrated luminosity of 20.3 fb^{-1}, which is available for physics studies. By several analysis steps, 112 171 $t\bar{t}$ candidates are reconstructed, with an expected purity of 90.0 %, from which 183 are assumed to decay through $t \to s + W$. This decay establishes the basis to determine the CKM matrix element $|V_{ts}|$ by a comparison to Monte Carlo simulations that contain the relevant theoretical predictions.
To separate $t \to s + W$ decays from the dominant $t \to b + W$ background, several differences are evaluated, covering the properties of jets, tracks and the outcome of b-tagging algorithms. Furthermore, the s-quark hadrons K^0_S are considered, which are reconstructed by a kinematic-fit-based method. In total 35 discriminating variables are thus available, from which 10 are selected for a multivariate analysis, namely "boosted decision trees". The subsequent output can be interpreted as one single discrimination variable for further studies. The multivariate analysis is executed twice, once for all events and once for events with at least one K^0_S candidate, resulting in two independent analyses.
The final step of each of the two analyses is based on a template fit method, at which several signal strength dependent hypotheses are formulated and consequently tested with respect to their agreement with data. The best outcome is a significance value of 0.7σ for the $t \to s + W$ decays and an upper limit

$$|V_{ts}|^2 < 1.74\ \%$$

at 95 % confidence level, which coincides well with the expected values from the PDG [1]. For these results, the contributions from $t \to d + W$ decays are neglected and a unitary CKM matrix is assumed. All relevant uncertainties are considered in the template fit method.
This is the first direct $|V_{ts}|^2$ measurement. Moreover, comparison with direct $|V_{tb}|$ measurements, which can be interpreted with respect to $|V_{ts}|$, show that a sensitivity increase by a factor of 2.9 is achieved.

The main limitation of the measurement arises from the low number of $t \to s + W$ decays, since only about 0.16 % of all reconstructed events are decaying through $t\bar{t} \to W^+W^-b\bar{s}$ (charge conjugation implied). Despite the great effort to optimise this very first direct measurement, there are some proposals which might result in further improved $|V_{ts}|$ analysis results in the future:

- Independent execution of the top pair reconstruction method under the assumption of two b-quarks or one s- and one b-quark, respectively. The different masses of the quarks could be used as fit constraints, resulting in two independent χ^2 values, which could further be used as discriminating variables.
- For an identification of semi-leptonic b-quark decays in $t\bar{t}$ events, all charged lepton properties should be retained as additional information.
- A recalculation of the K_S^0 decay length with additional information (e.g. considering the K_S^0-p_{T}, calculations in three dimensions, etc.) might improve the discrimination power of this specific variable.
- Further s-quark hadrons, like K^+ or K^-, should be incorporated. The reconstruction might be based on effects in the transition radiation tracker (TRT), by separating kaons from pions at high energies due to their characteristic radiations. This approach might be easier to implement at LHC Run II due to the higher centre-of-mass energies $\sqrt{s}$.
- Apart from the most dominant background of $t \to b + W$ decays, W+jets events are misidentified as $t \to s + W$ decays as well. Hence, a discrimination study to identify variables, which differ between W background and the signal events, might improve the full analysis result.
- In this analysis, two main backgrounds are incorporated in the "boosted decision trees" training: W+jets and $t \to b + W$ decays. A separation of the W+jets background in a prior, independent BDT might enhance the total signal strength.
- All discrimination studies are based on the leading order generator `PROTOS`, even though the results from the higher order `MC@NLO` generator achieve a better data agreement. With a newly generated `MC@NLO` sample, based on an enhanced fraction of $t \to s + W$ decays, the discrimination studies could be performed with more reliable NLO events, improving the full analysis substantially.

Furthermore, the analysis in hand neglects $t \to d + W$ decays due to the much smaller expected branching fraction, with only 1/20 of the regarded s-quark signal. An implementation of this contribution is the next consequent step. The separation from $t \to s + W$ decays could then be based once more on kaon reconstruction, since high energetic kaons are less likely to occur in $t \to d + W$ decays than in other top decays. In addition, characteristic high energetic pions or neutrons might be used for the down quark identification.

In 2015, the LHC restarts with higher centre-of-mass energies of $\sqrt{s} = 13$ TeV followed by $\sqrt{s} = 14$ TeV later on. Thus, a future analysis can rely on new data with higher energies, at which several improvements are expectable:
First, the $t\bar{t}$ cross section and thus the number of $t \to s + W$ decays will increase by

about a factor of 4 [134, 135], with respect to a fixed luminosity value. Second, the energy of the particles involved will rise, improving the separation power of kaons in the TRT. Third, an integrated luminosity of about 100 fb^{-1} is expected for LHC Run II, compared to 20.3 fb^{-1} in 2012.
Taking into account all these factors, an approximated significance of more than 3σ can be expected without modifying the actual reconstruction procedures or analysis steps. With an implementation of some of the above suggested improvements, this should easily reach the 5σ threshold for a measurement of the matrix element $|V_{ts}|$.
Deviations from the expected value would have a strong impact on the understanding of nature, since the CKM matrix is one of the major aspects in the Standard Model of particle physics. It e.g. accounts for the CP violation, which can explain the inequality of matter and anti-matter to some extend, and furthermore connects the strong and weak forces via the quark eigenstates. Up to now, all precision measurements confirm its theoretical properties, while any deviation would clearly indicate new physics. Direct measurements of all matrix elements are thus crucial.

However, up to now, this study is the first and only direct measurement aiming for an extraction of $|V_{ts}|$. It thus results in the best limit on the CKM matrix element $|V_{ts}|^2 < 1.74\,\%$ at $95\,\%$ confidence level.

Appendix

A. Appendix

A.1. Validation of the MC generator PROTOS

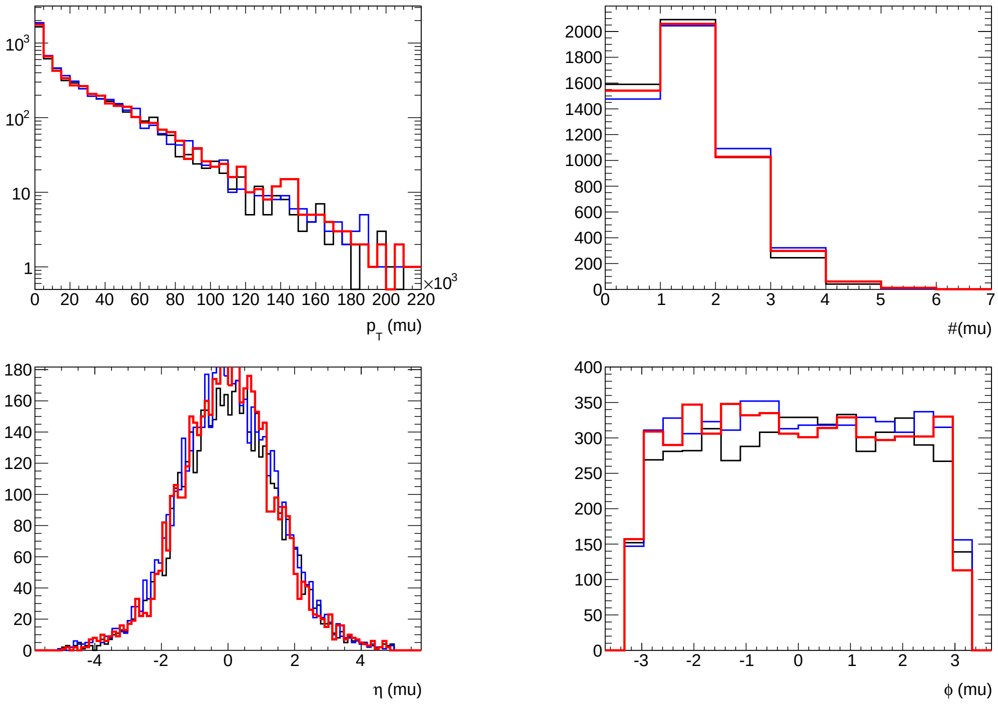


Figure A.1.: Comparison of `PROTOS` generated events (red) with those from `POWHEG` (blue) and `MC@NLO` (black), equivalent to Figure 4.1, but here for muons. All `PROTOS` distributions are in agreement with the two approved generators `POWHEG` and `MC@NLO`.

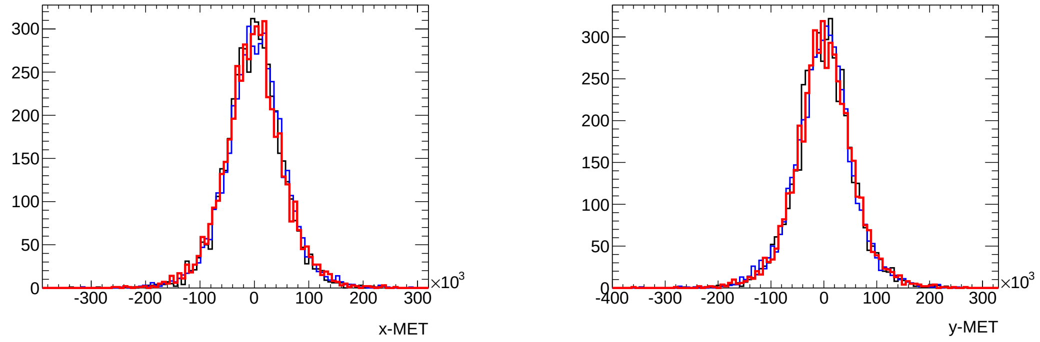


Figure A.2.: Comparison of `PROTOS` generated events (red) with those from `POWHEG` (blue) and `MC@NLO` (black) for the missing transverse energy x-component (left) and y-component (right). All `PROTOS` distributions are in agreement with the two approved generators `POWHEG` and `MC@NLO`.

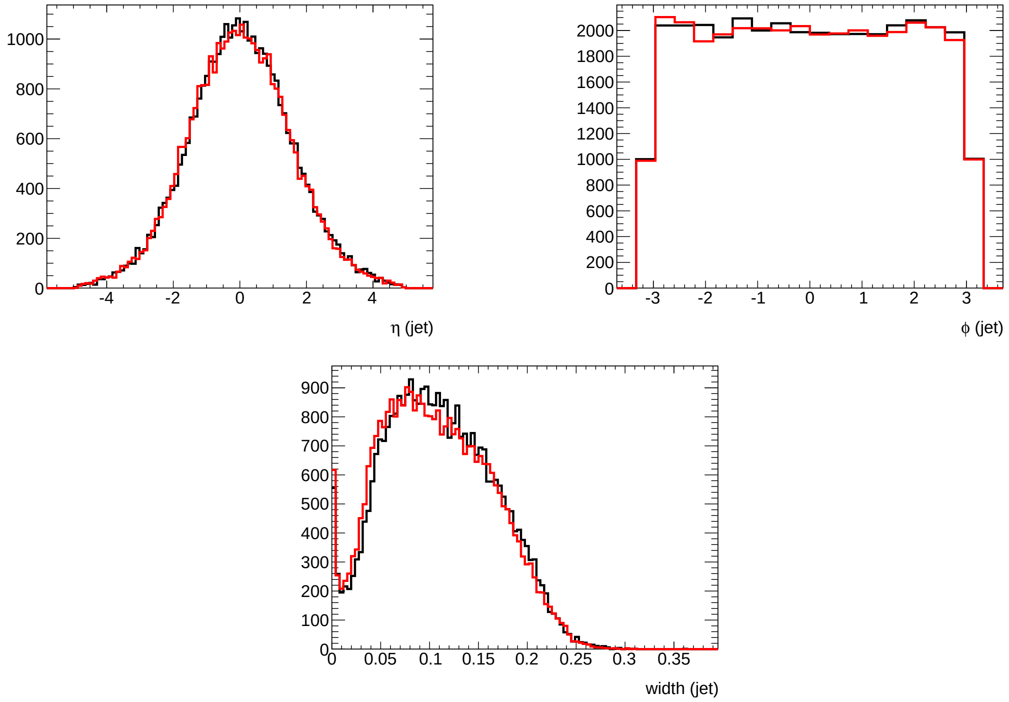


Figure A.3.: Comparison of the two PROTOS generated datasets "SM BR" (black) vs. "non-SM BR" (red) with respect to jet properties, similar to Figure 4.3.

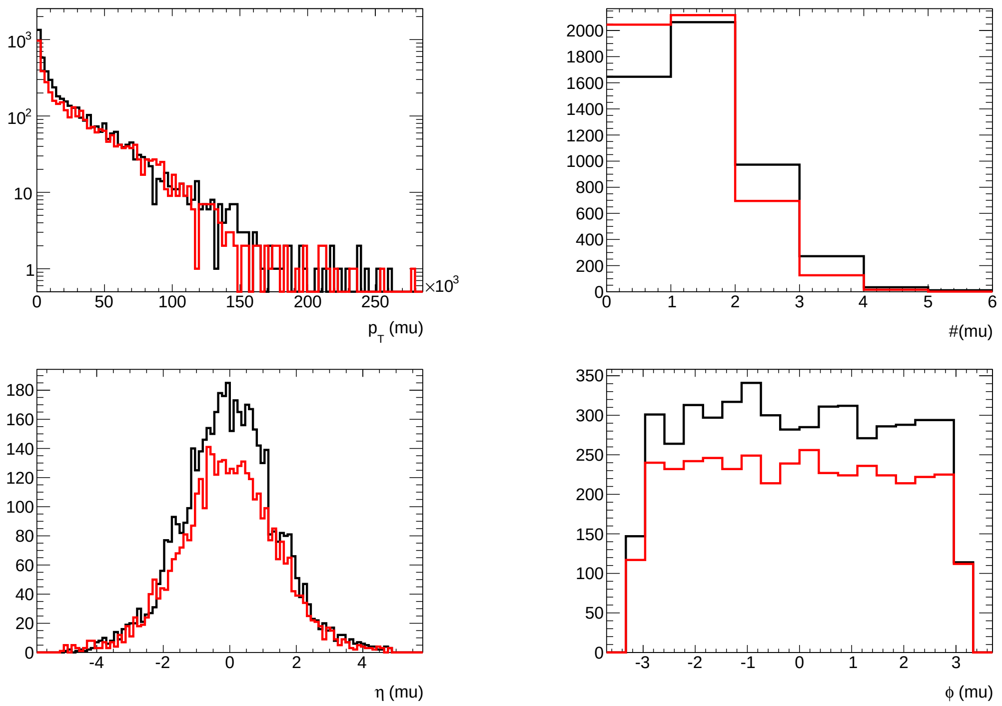


Figure A.4.: Comparison of the two PROTOS generated datasets “SM BR” (black) vs. “non-SM BR” (red) with respect to muon properties, similar to Figure 4.3.

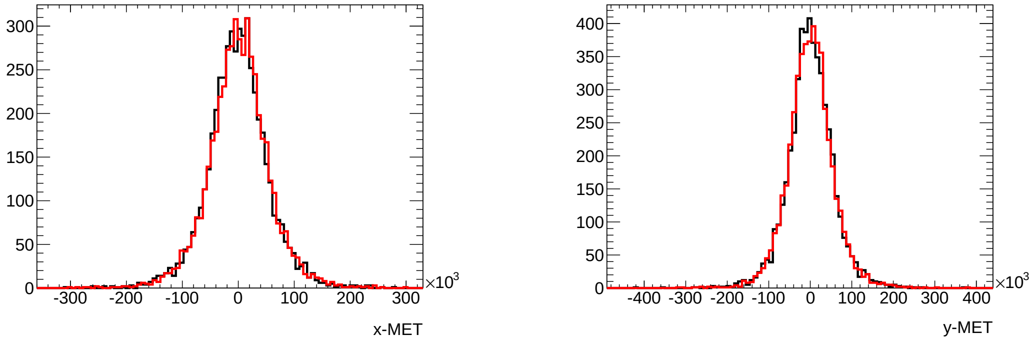
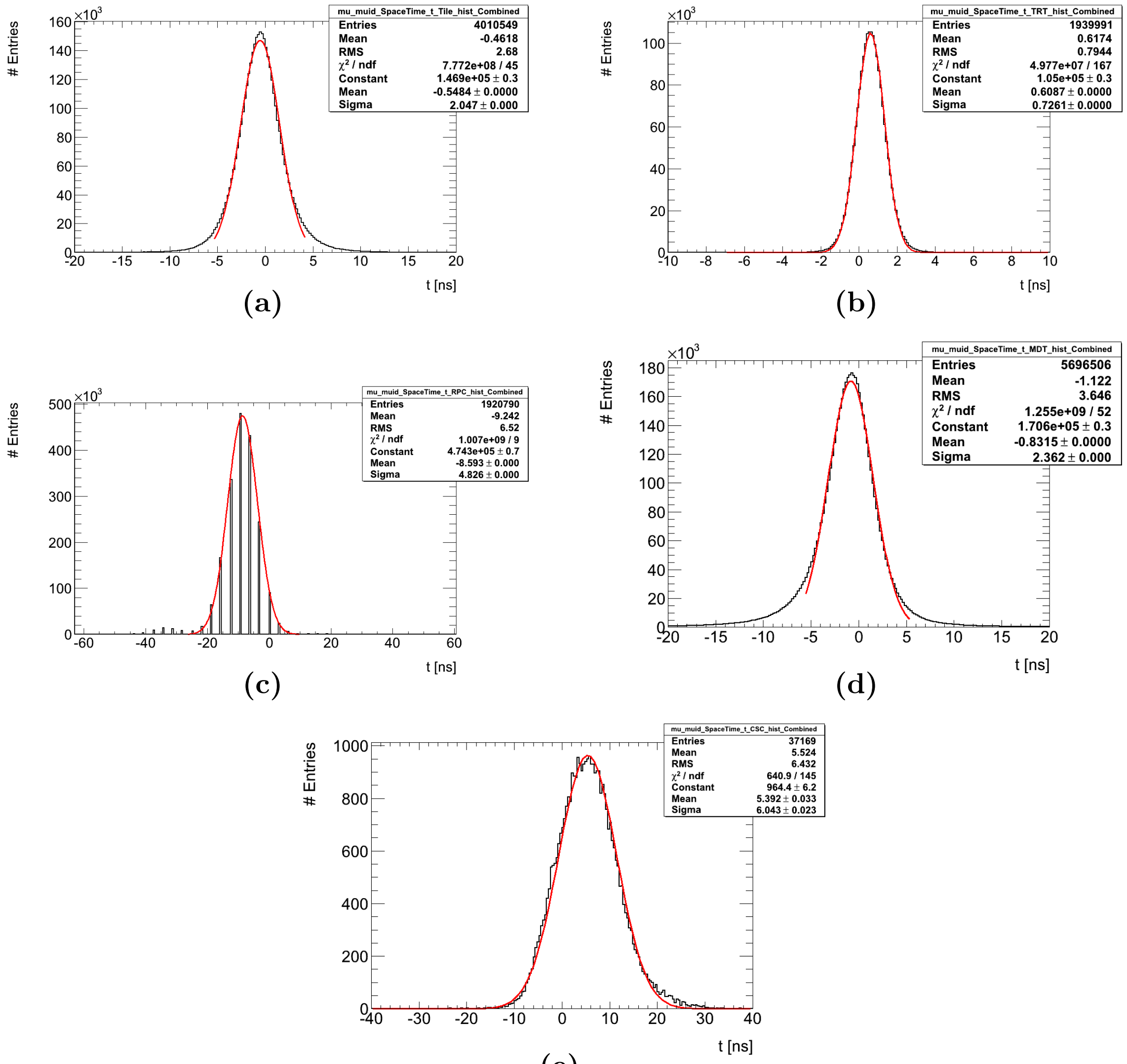


Figure A.5.: Comparison of the two PROTOS generated datasets “SM BR” (black) vs. “non-SM BR” (red) with respect to MET properties, similar to Figure 4.3.

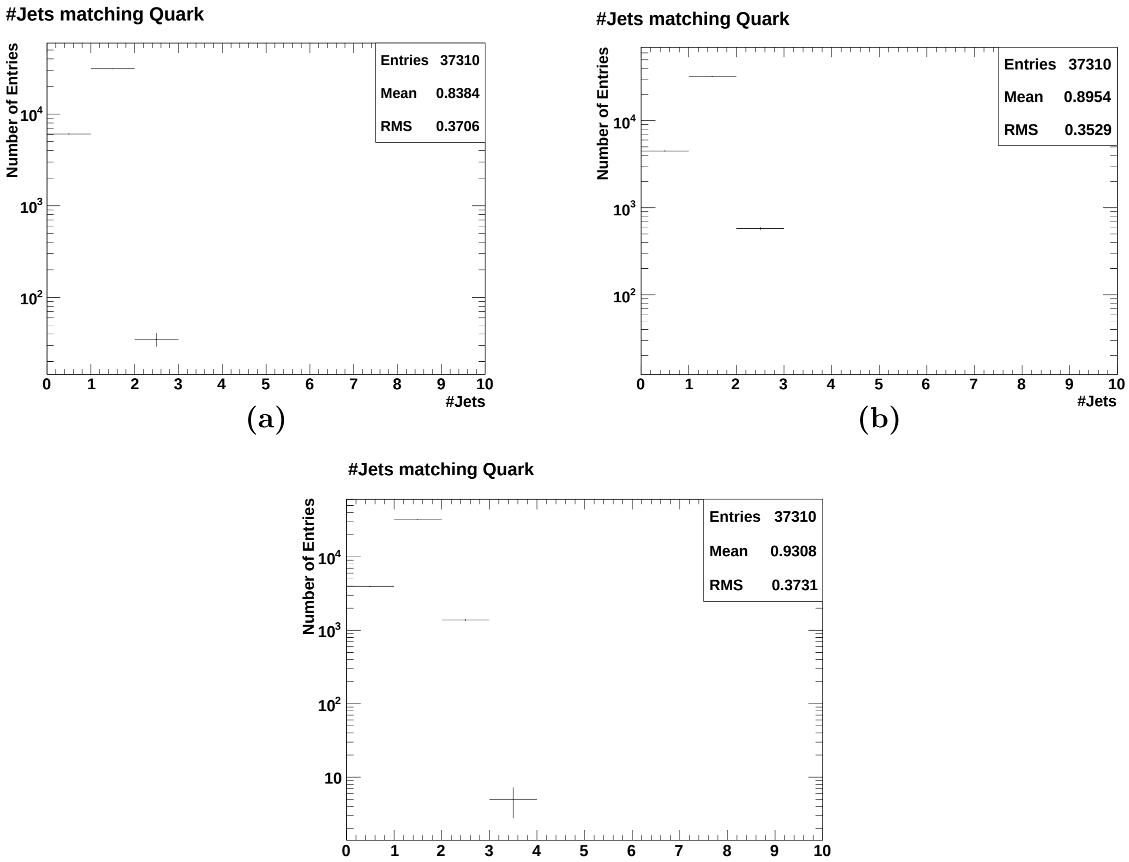
A.2. Time distributions used in the cosmic muon rejection tool



(a) (b) (c) (d) (e)

Figure A.6.: Timing distribution of the detector components used for the development of the cosmic muon rejection tool described in section 5.2.2. These distributions were used for the calibration of the time measurements: (a) Tile calorimeter, (b) TRT, (c): RPC, (d) MDT, (e) CSC

A.3. Optimisation of jet parton matching



(a) (b) (c)

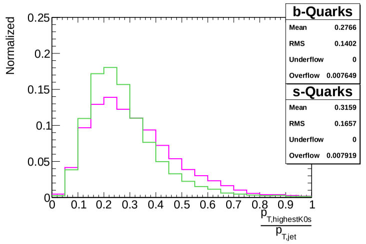
Figure A.7.: Comparison of different dR parameters for the jet parton matching algorithm. (a) $dR = 0.25$, (b) $dR = 0.4$, (c) $dR = 0.5$.
$\Delta R = 0.4$ (b) yields the best performance with a relatively low number of multiple assignments and a relatively low number of quarks with no assignments at all.

A.4. Optimisation of K_S^0 reconstruction

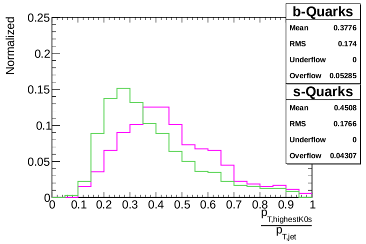
To optimise the K_S^0 reconstruction steps , the track p_T, the track quality values (χ^2) and the p-values of the reconstruction are variated. A selection of 12 different combinations of these tests is given below. They are optimised in order to find the best agreement with truth K_S^0 and in order to achieve the best discrimination between $t \rightarrow s + W$ and $t \rightarrow b + W$ decays.

$p_{T,K0_s}/p_{T,jet}$ (leading K0s)

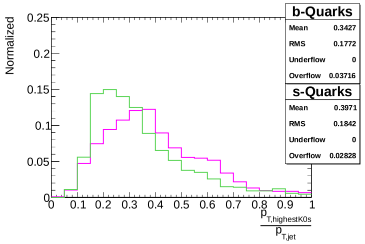
$p_t = 5GeV$
p-val = 0, $\chi_{trk} = 2.5$



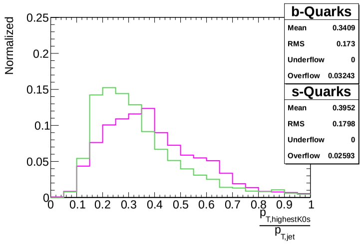
$p_t = 15GeV$
p-value = 0.00005



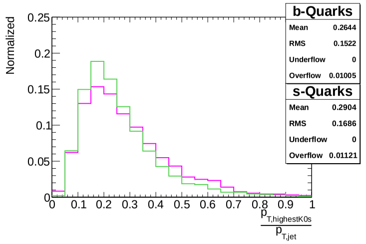
$p_t = 10GeV$
p-value = 0.01



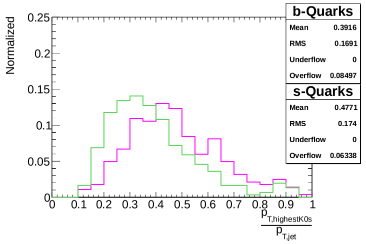
$p_t = 10GeV$
p-value = 0.00005



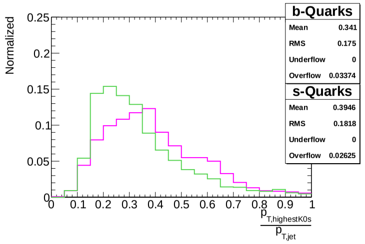
$p_t = 5GeV$
p = 0.05, $\chi_{trk} = 1.5$



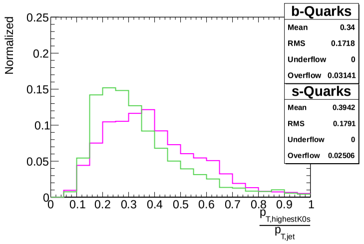
$p_t = 20GeV$
p-value = 0.00001



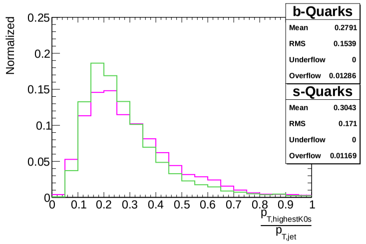
$p_t = 10GeV$
p-value = 0.0005



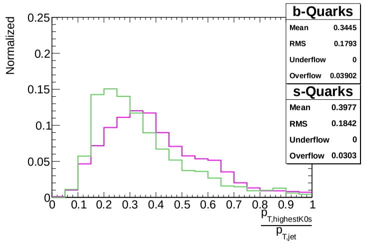
$p_t = 10GeV$
p-value = 0.00001



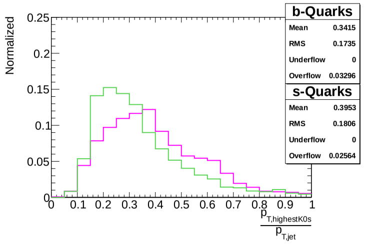
$p_t = 5GeV$
p-value = 0.0005



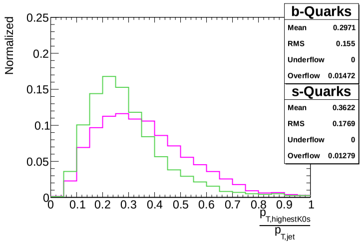
$p_t = 10GeV$
p-value = 0.03



$p_t = 10GeV$
p-value = 0.0001

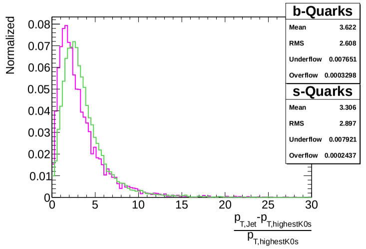


$p_t = 10GeV$
p-value = 0 (No cut)

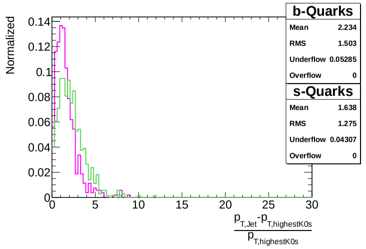


$p_{T,jet} - p_{T,K0_s}/p_{T,K0_s}$ (leading K0s)

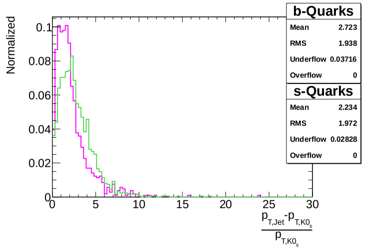
$p_t = 5GeV$
p-val $= 0$, $\chi_{trk} = 2.5$



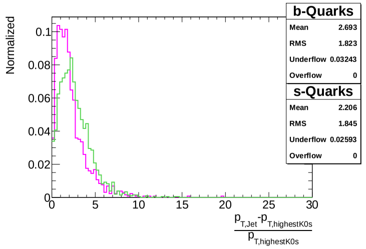
$p_t = 15GeV$
p-value $= 0.00005$



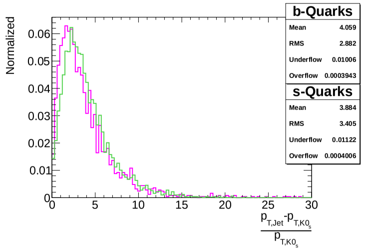
$p_t = 10GeV$
p-value $= 0.01$



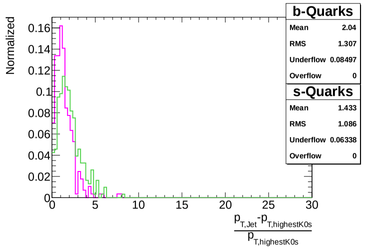
$p_t = 10GeV$
p-value $= 0.00005$



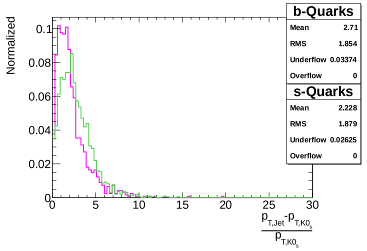
$p_t = 5GeV$
p $= 0.05$, $\chi_{trk} = 1.5$



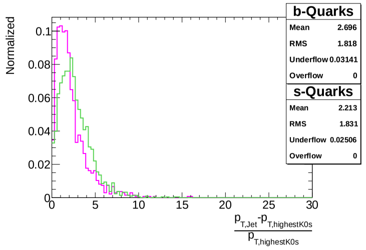
$p_t = 20GeV$
p-value $= 0.00001$



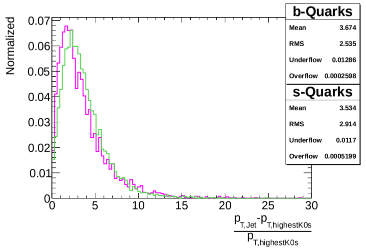
$p_t = 10GeV$
p-value $= 0.0005$



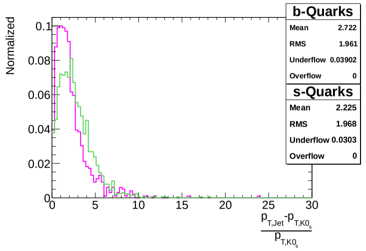
$p_t = 10GeV$
p-value $= 0.00001$



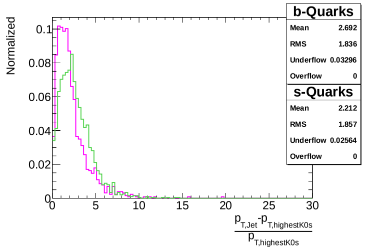
$p_t = 5GeV$
p-value $= 0.0005$



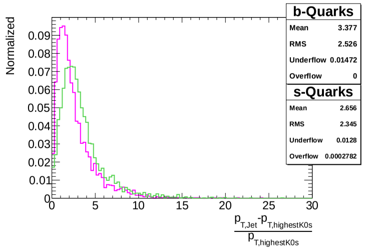
$p_t = 10GeV$
p-value $= 0.03$



$p_t = 10GeV$
p-value $= 0.0001$

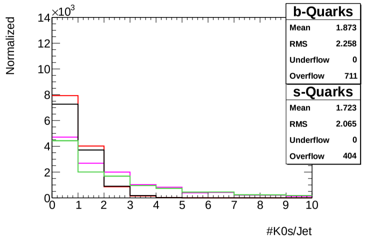


$p_t = 10GeV$
p-value $= 0$ (No cut)

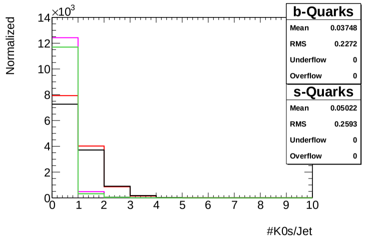


Number of K0s perJet

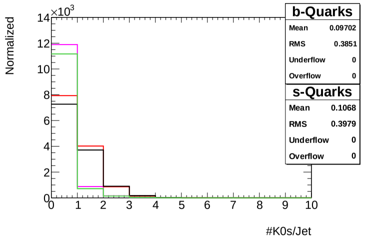
$p_t = 5GeV$
p-val = 0, $\chi_{trk} = 2.5$



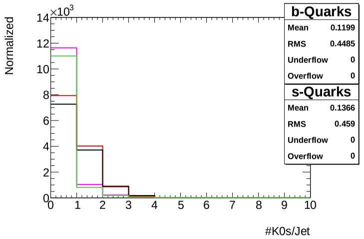
$p_t = 15GeV$
p-value = 0.00005



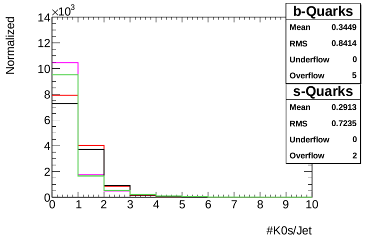
$p_t = 10GeV$
p-value = 0.01



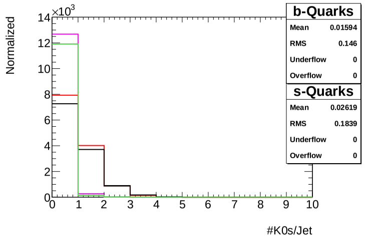
$p_t = 10GeV$
p-value = 0.00005



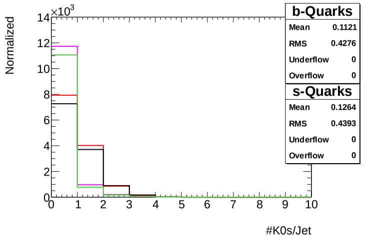
$p_t = 5GeV$
p = 0.05, $\chi_{trk} = 1.5$



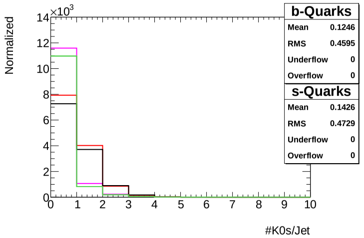
$p_t = 20GeV$
p-value = 0.00001



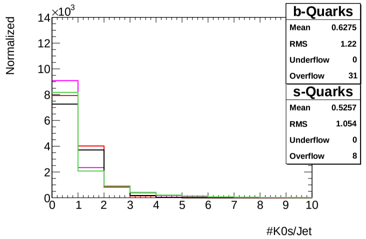
$p_t = 10GeV$
p-value = 0.0005



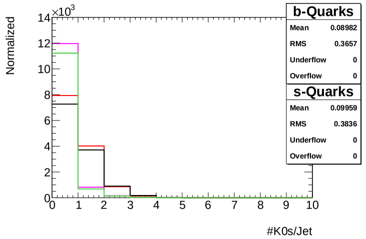
$p_t = 10GeV$
p-value = 0.00001



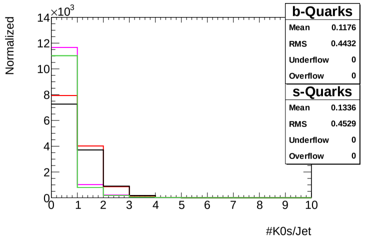
$p_t = 5GeV$
p-value = 0.0005



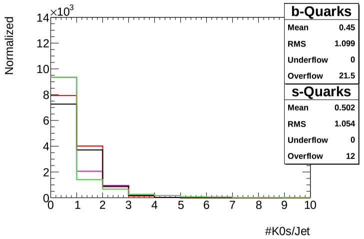
$p_t = 10GeV$
p-value = 0.03



$p_t = 10GeV$
p-value = 0.0001

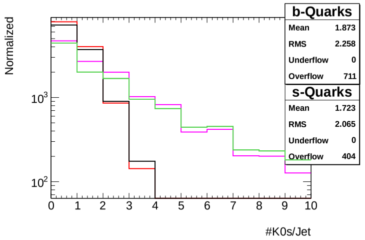


$p_t = 10GeV$
p-value = 0 (No cut)

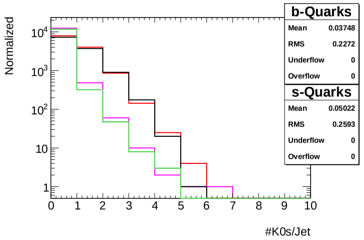


Number of K0s perJet (log Y)

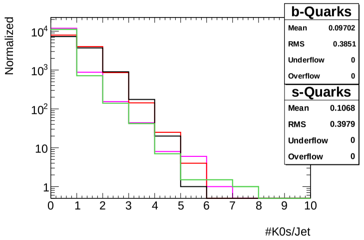
$p_t = 5 GeV$
p-val = 0, $\chi_{trk} = 2.5$



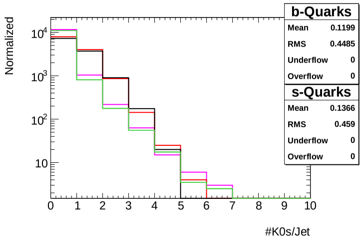
$p_t = 15 GeV$
p-value = 0.00005



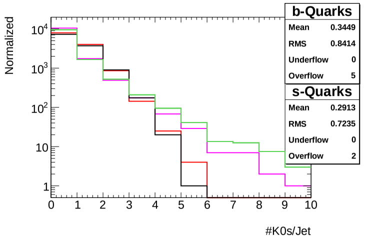
$p_t = 10 GeV$
p-value = 0.01



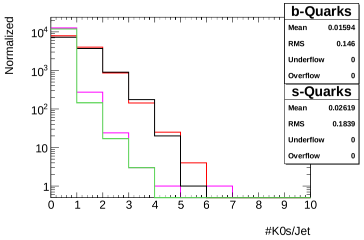
$p_t = 10 GeV$
p-value = 0.00005



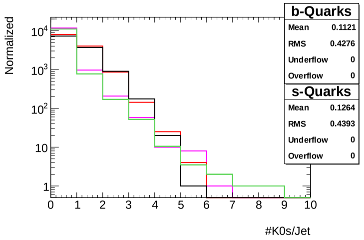
$p_t = 5 GeV$
p = 0.05, $\chi_{trk} = 1.5$



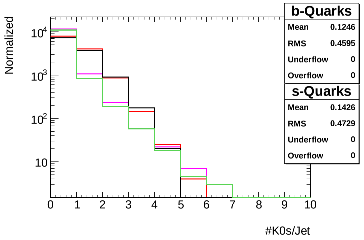
$p_t = 20 GeV$
p-value = 0.00001



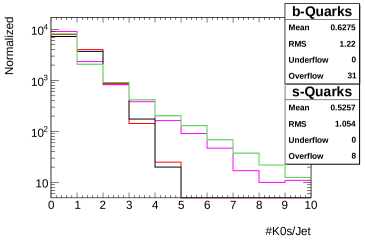
$p_t = 10 GeV$
p-value = 0.0005



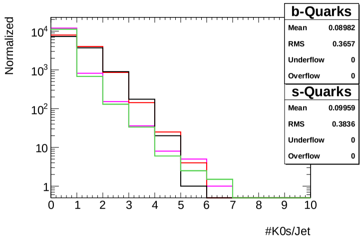
$p_t = 10 GeV$
p-value = 0.00001



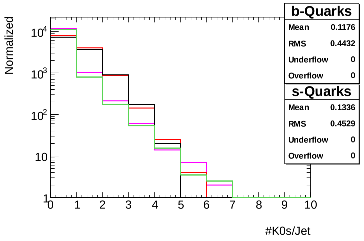
$p_t = 5 GeV$
p-value = 0.0005



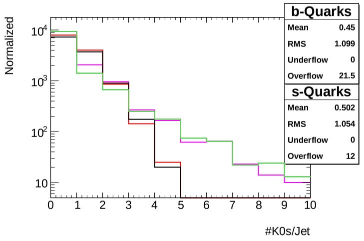
$p_t = 10 GeV$
p-value = 0.03



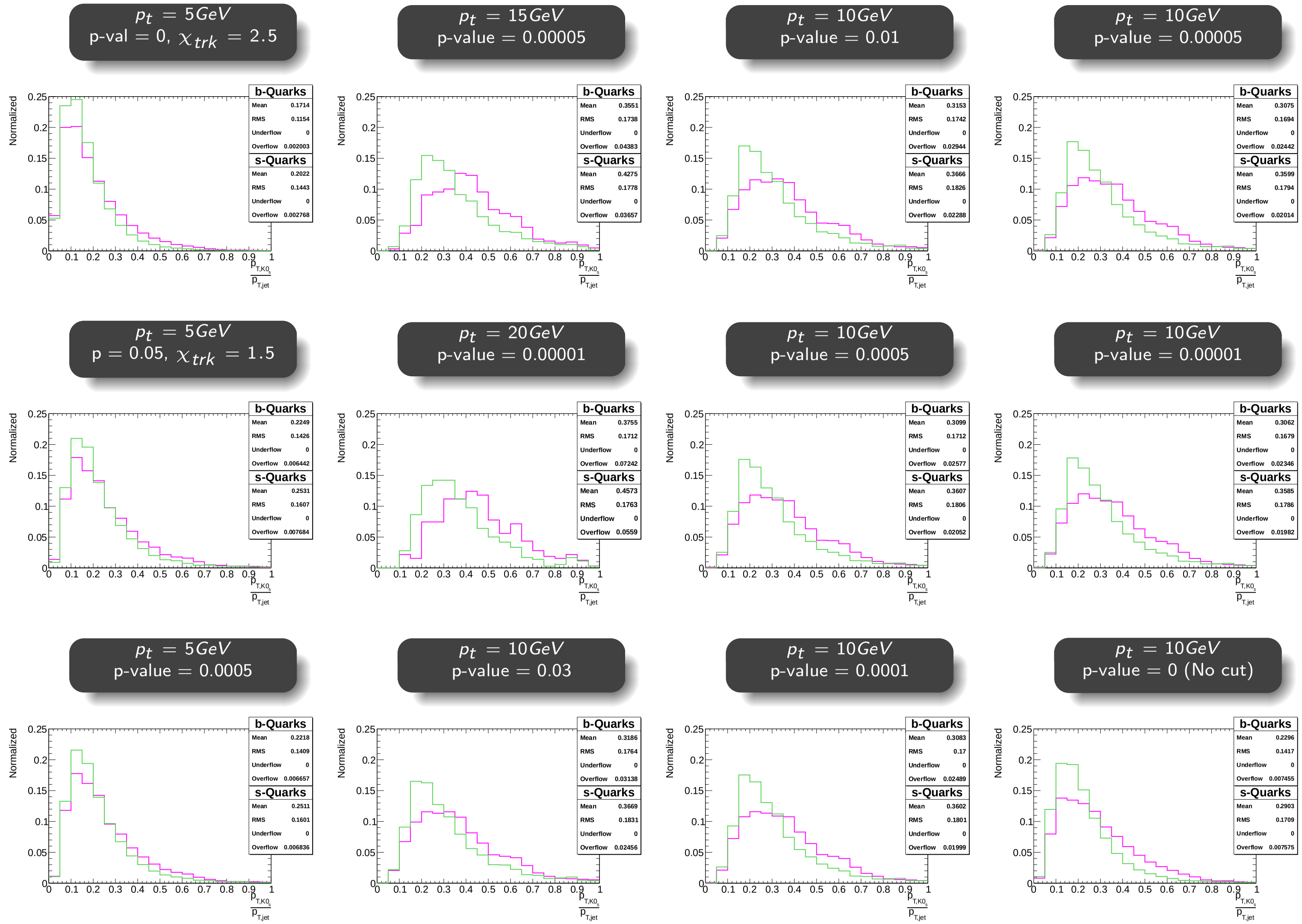
$p_t = 10 GeV$
p-value = 0.0001



$p_t = 10 GeV$
p-value = 0 (No cut)

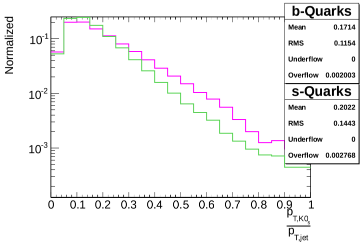


$p_{T,K0_s}/p_{T,jet}$

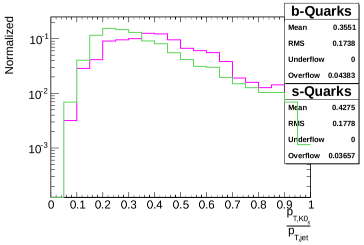

$p_t = 5GeV$
p-val = 0, $\chi_{trk} = 2.5$
b-Quarks Mean 0.1714 RMS 0.1154 Underflow 0 Overflow 0.002003
s-Quarks Mean 0.2022 RMS 0.1443 Underflow 0 Overflow 0.002768
$p_t = 15GeV$
p-value = 0.00005
b-Quarks Mean 0.3551 RMS 0.1738 Underflow 0 Overflow 0.04383
s-Quarks Mean 0.4275 RMS 0.1778 Underflow 0 Overflow 0.03657
$p_t = 10GeV$
p-value = 0.01
b-Quarks Mean 0.3153 RMS 0.1742 Underflow 0 Overflow 0.02944
s-Quarks Mean 0.3666 RMS 0.1826 Underflow 0 Overflow 0.02288
$p_t = 10GeV$
p-value = 0.00005
b-Quarks Mean 0.3075 RMS 0.1694 Underflow 0 Overflow 0.02442
s-Quarks Mean 0.3599 RMS 0.1794 Underflow 0 Overflow 0.02014
$p_t = 5GeV$
p = 0.05, $\chi_{trk} = 1.5$
b-Quarks Mean 0.2249 RMS 0.1426 Underflow 0 Overflow 0.006442
s-Quarks Mean 0.2531 RMS 0.1607 Underflow 0 Overflow 0.007684
$p_t = 20GeV$
p-value = 0.00001
b-Quarks Mean 0.3755 RMS 0.1712 Underflow 0 Overflow 0.07242
s-Quarks Mean 0.4573 RMS 0.1763 Underflow 0 Overflow 0.0559
$p_t = 10GeV$
p-value = 0.0005
b-Quarks Mean 0.3099 RMS 0.1712 Underflow 0 Overflow 0.02577
s-Quarks Mean 0.3607 RMS 0.1806 Underflow 0 Overflow 0.02052
$p_t = 10GeV$
p-value = 0.00001
b-Quarks Mean 0.3062 RMS 0.1679 Underflow 0 Overflow 0.02346
s-Quarks Mean 0.3585 RMS 0.1786 Underflow 0 Overflow 0.01982
$p_t = 5GeV$
p-value = 0.0005
b-Quarks Mean 0.2218 RMS 0.1409 Underflow 0 Overflow 0.006657
s-Quarks Mean 0.2511 RMS 0.1601 Underflow 0 Overflow 0.006836
$p_t = 10GeV$
p-value = 0.03
b-Quarks Mean 0.3186 RMS 0.1764 Underflow 0 Overflow 0.03138
s-Quarks Mean 0.3669 RMS 0.1831 Underflow 0 Overflow 0.02456
$p_t = 10GeV$
p-value = 0.0001
b-Quarks Mean 0.3083 RMS 0.17 Underflow 0 Overflow 0.02489
s-Quarks Mean 0.3602 RMS 0.1801 Underflow 0 Overflow 0.01999
$p_t = 10GeV$
p-value = 0 (No cut)
b-Quarks Mean 0.2296 RMS 0.1417 Underflow 0 Overflow 0.007455
s-Quarks Mean 0.2903 RMS 0.1709 Underflow 0 Overflow 0.007575
Normalized
0 0.05 0.1 0.15 0.2 0.25
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
$\frac{p_{T,K0_s}}{p_{T,jet}}$

$p_{T,K0_s}/p_{T,jet}$ (logY)

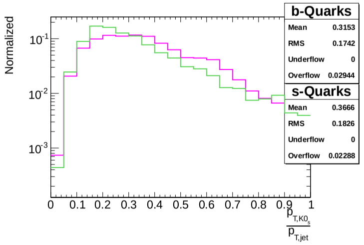
p_t = 5GeV
p-val = 0, χ_{trk} = 2.5



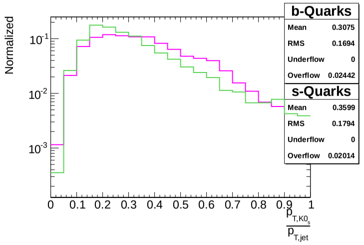
p_t = 15GeV
p-value = 0.00005



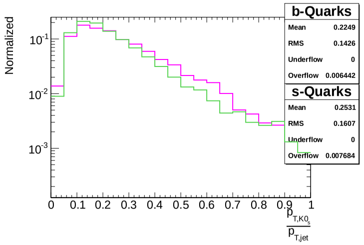
p_t = 10GeV
p-value = 0.01



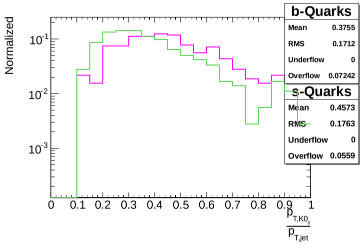
p_t = 10GeV
p-value = 0.00005



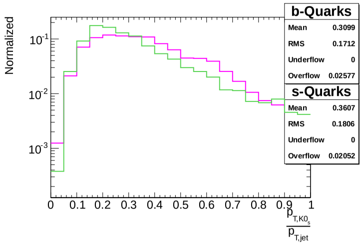
p_t = 5GeV
p = 0.05, χ_{trk} = 1.5



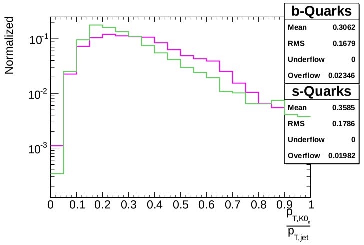
p_t = 20GeV
p-value = 0.00001



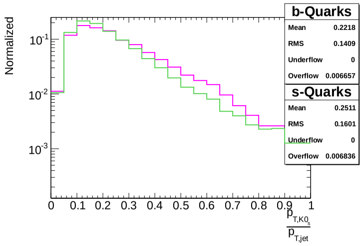
p_t = 10GeV
p-value = 0.0005



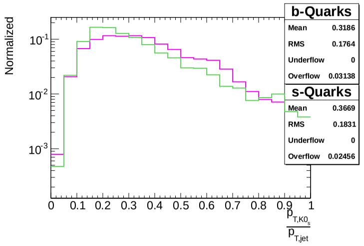
p_t = 10GeV
p-value = 0.00001



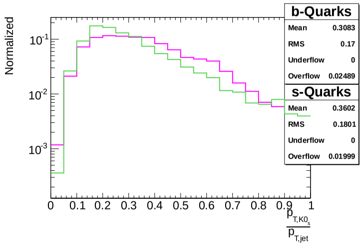
p_t = 5GeV
p-value = 0.0005



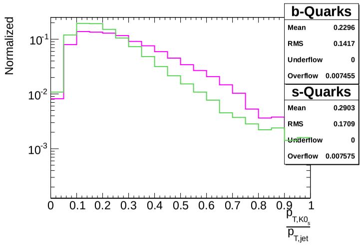
p_t = 10GeV
p-value = 0.03



p_t = 10GeV
p-value = 0.0001

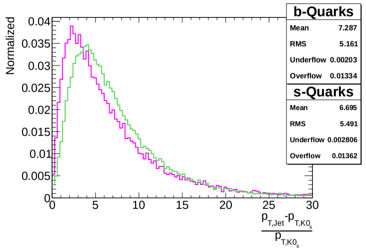


p_t = 10GeV
p-value = 0 (No cut)

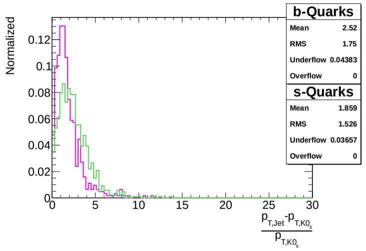


$p_{T,jet} - p_{T,K0_s} / p_{T,K0_s}$

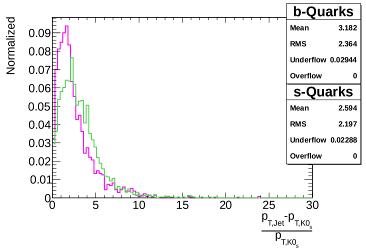
$p_t = 5GeV$
p-val = 0, $\chi_{trk} = 2.5$



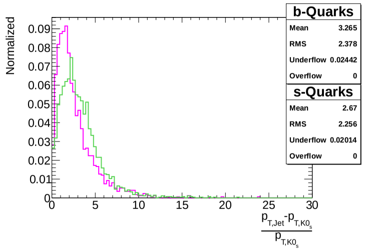
$p_t = 15GeV$
p-value = 0.00005



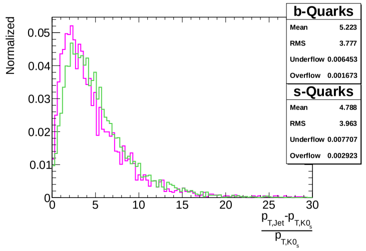
$p_t = 10GeV$
p-value = 0.01



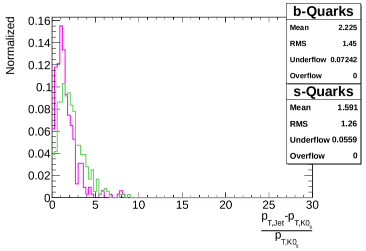
$p_t = 10GeV$
p-value = 0.00005



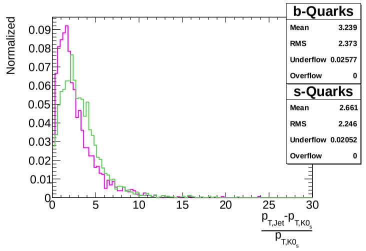
$p_t = 5GeV$
p = 0.05, $\chi_{trk} = 1.5$



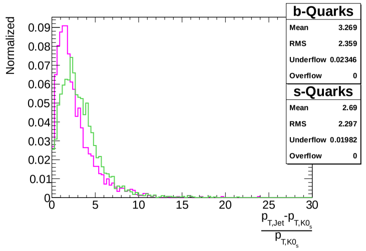
$p_t = 20GeV$
p-value = 0.00001



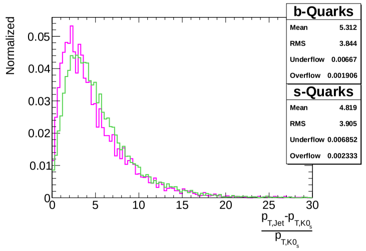
$p_t = 10GeV$
p-value = 0.0005



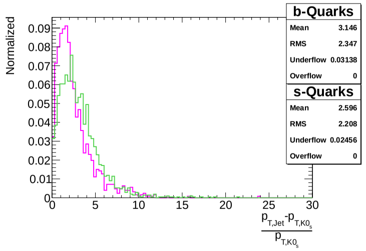
$p_t = 10GeV$
p-value = 0.00001



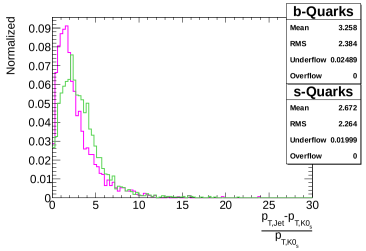
$p_t = 5GeV$
p-value = 0.0005



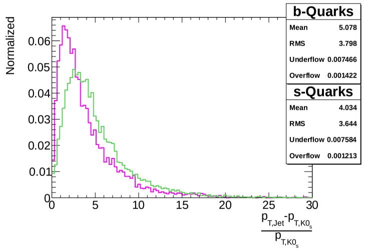
$p_t = 10GeV$
p-value = 0.03



$p_t = 10GeV$
p-value = 0.0001

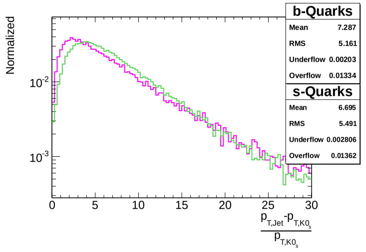


$p_t = 10GeV$
p-value = 0 (No cut)

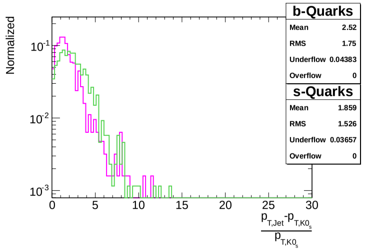


$p_{T,jet} - p_{T,K0_s} / p_{T,K0_s}$ (logY)

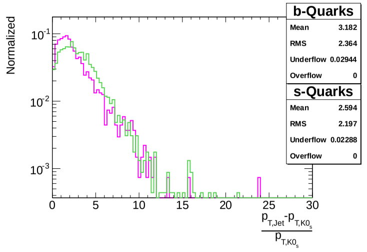
$p_t = 5GeV$
p-val $= 0$, $\chi_{trk} = 2.5$



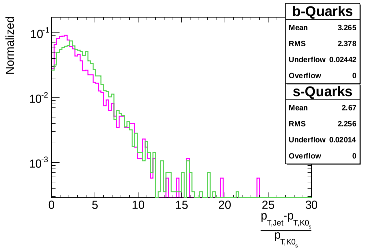
$p_t = 15GeV$
p-value $= 0.00005$



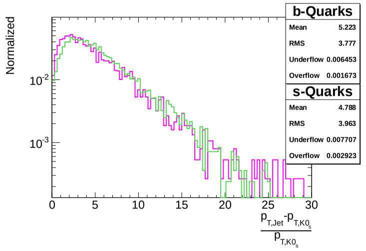
$p_t = 10GeV$
p-value $= 0.01$



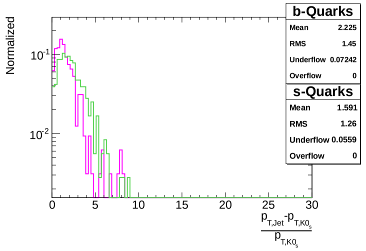
$p_t = 10GeV$
p-value $= 0.00005$



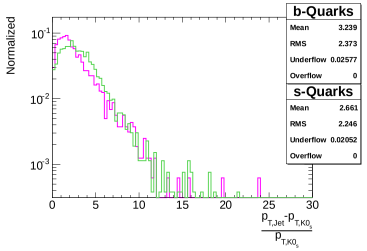
$p_t = 5GeV$
p $= 0.05$, $\chi_{trk} = 1.5$



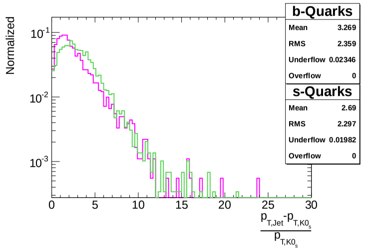
$p_t = 20GeV$
p-value $= 0.00001$



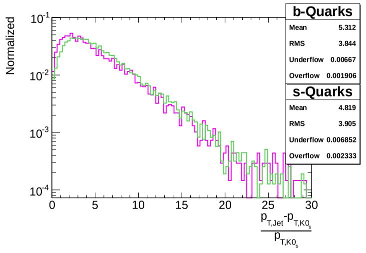
$p_t = 10GeV$
p-value $= 0.0005$



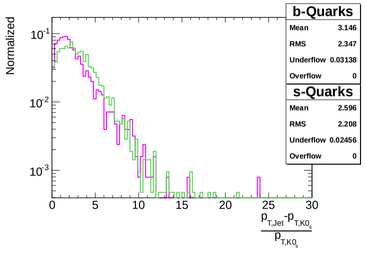
$p_t = 10GeV$
p-value $= 0.00001$



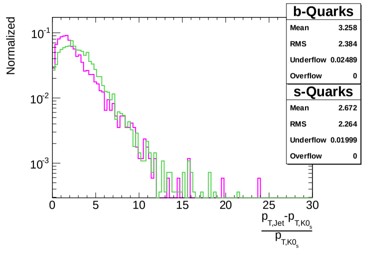
$p_t = 5GeV$
p-value $= 0.0005$



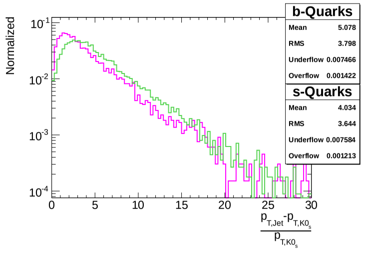
$p_t = 10GeV$
p-value $= 0.03$



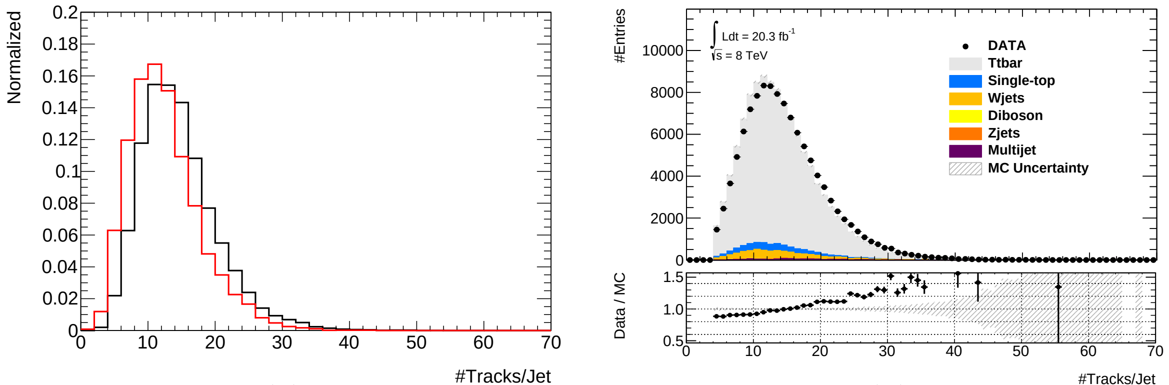
$p_t = 10GeV$
p-value $= 0.0001$



$p_t = 10GeV$
p-value $= 0$ (No cut)

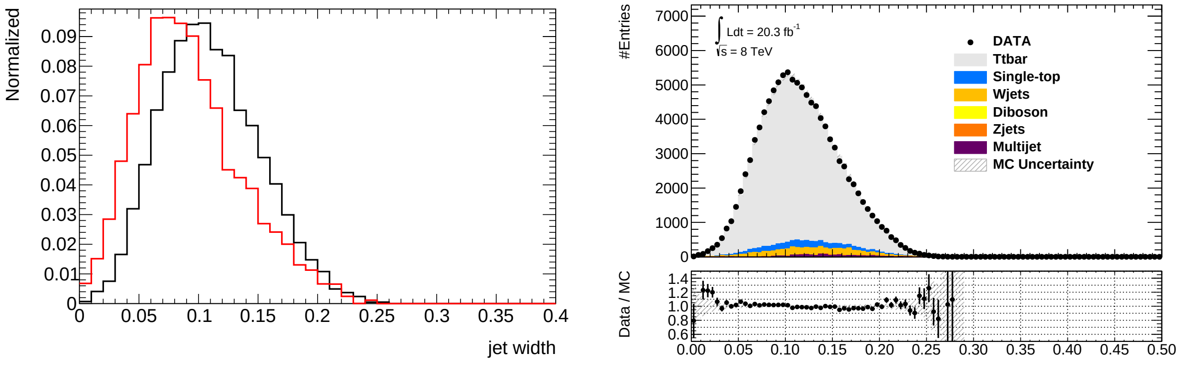


A.5. Supplemental s-quark identification plots



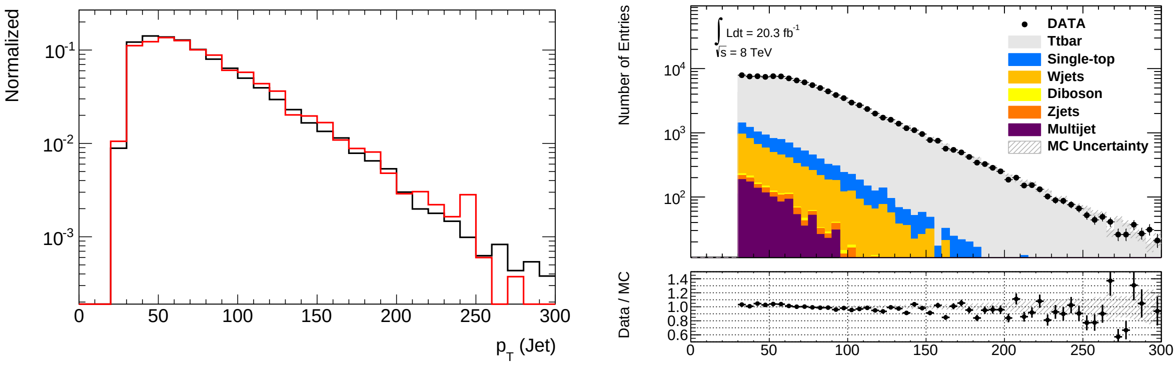
(a) **(b)**

Figure A.8.: (a) Jet multiplicity for jets stemming from $\sqrt{s} = 8$ TeV `PROTOS` events, with $t \to s + W$ decays in red and $t \to b + W$ decays in black. (b) The data-MC agreement shows discrepancies.



(a) **(b)**

Figure A.9.: (a) Jet cone size for jets stemming from $\sqrt{s} = 8$ TeV `PROTOS` events, with $t \to s + W$ decays in red and $t \to b + W$ decays in black. (b) The data-MC agreement shows reasonable agreement.



(a) **(b)**

Figure A.10.: (a) Jet p_{T} in GeV for jets stemming from $\sqrt{s} = 8$ TeV `PROTOS` events, with $t \rightarrow s + W$ decays in red and $t \rightarrow b + W$ decays in black. (b) The data-MC agreement shows reasonable agreement.

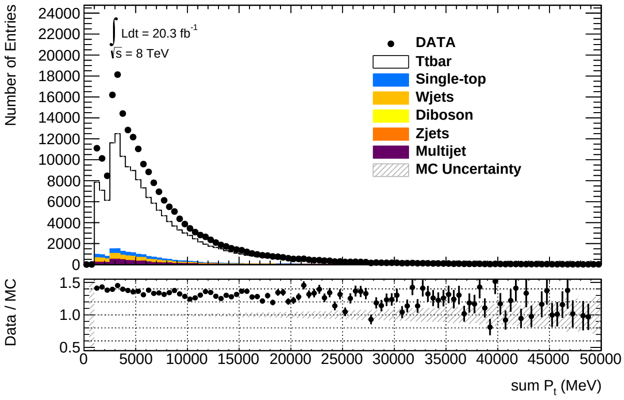
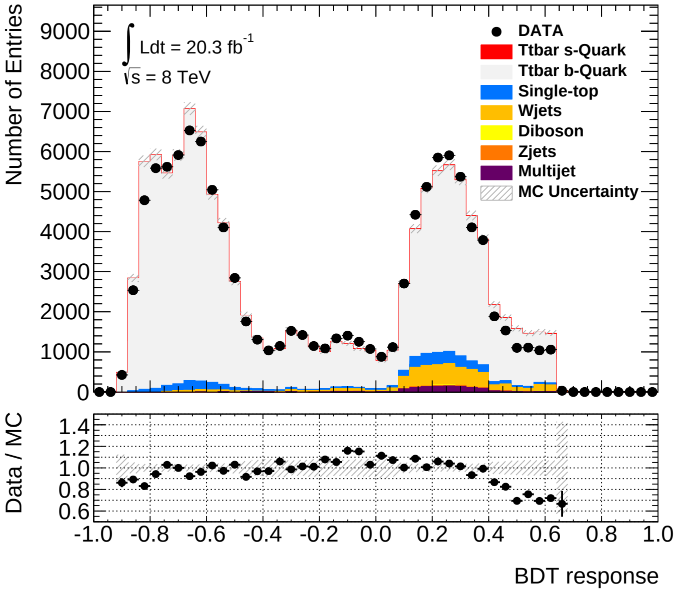
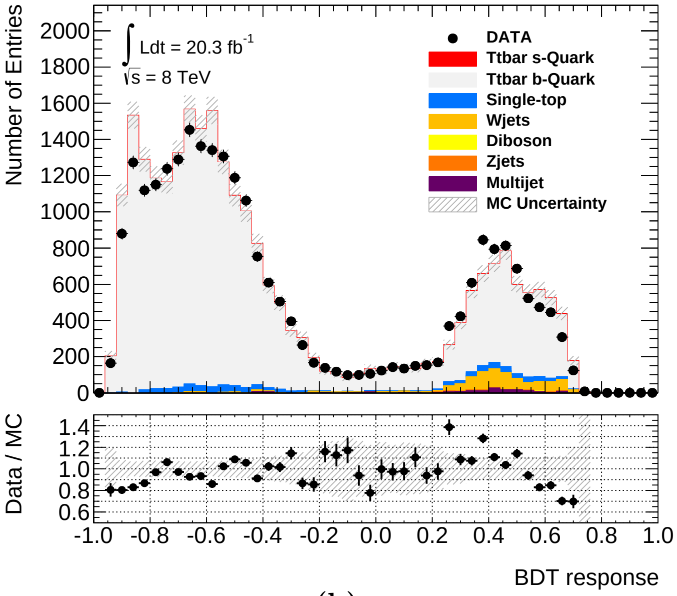


Figure A.11.: Data-MC comparison for V0 particles. Discrepancies of 30 - 50 % are visible.

A.6. BDT response based on `PROTOS` generated $t\bar{t}$ events



(a)



(b)

Figure A.12.: BDT response using `PROTOS` generated $t\bar{t}$ events with (a) no cut applied and (b) with at least one K^0_S candidate. As expected the agreements are quite loose, especially at high values which is caused by the missing NLO terms.

Bibliography

[1] K.A. Olive *et al.* (Particle Data Group), “The Review of particle physics,” *Chin. Phys. C*, vol. 38, p. 090001, 2014.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory.* Westview Press, 1995.

[3] D. Griffiths, *Introduction to Elementary Particles.* John Wiley & Sons, 1987.

[4] O. Biebel, “Studien der Quanten-Chromodynamik mit Collider-Experimenten,” 2004. Vorlesungsreihe im Wintersemester 2004/05 an der LMU München.

[5] L. Álvarez-Gaumé, J. Ellis, “Eyes on a prize particle,” *Nature Physics*, no. 1, p. 2–3, 2010.

[6] The ATLAS, CDF, CMS, D0 Collaborations, “First combination of Tevatron and LHC measurements of the top-quark mass,” *arXiv:1403.4427. ATLAS-CONF-2014-008. CDF-NOTE-11071. CMS-PAS-TOP-13-014. D0-NOTE-6416*, Mar 2014.

[7] S. W. Herb *et al.*, “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions,” *Phys. Rev. Lett.*, vol. 39, pp. 252–255, Aug 1977.

[8] F. Abe *et al.*, “Evidence for top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV,” *Phys. Rev. D*, vol. 50, pp. 2966–3026, Sep 1994.

[9] K. Hagiwara *et al.*, “Bound-state effects on top quark production at hadron colliders,” *Physics Letters B*, vol. 666, no. 1, pp. 71 – 76, 2008.

[10] N. Cabibbo, “Unitary symmetry and leptonic decays,” *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963.

[11] S. L. Glashow *et al.*, “Weak interactions with lepton-hadron symmetry,” *Phys. Rev. D*, vol. 2, pp. 1285–1292, Oct 1970.

[12] J. H. Christenson *et al.*, “Evidence for the 2π decay of the k_2^0 meson,” *Phys. Rev. Lett.*, vol. 13, pp. 138–140, Jul 1964.

[13] M. Kobayashi and T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction,” *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973.

[14] The LHCb Collaboration, “Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^- \pi^+$,” *New Journal of Physics*, vol. 15, no. 5, p. 053021, 2013.

[15] C. Elsasser, “Feynman Diagram Library.” `http://www.physik.uzh.ch/~che/FeynDiag/`, Feb. 2015.

[16] CERN, “About CERN.” `http://http://home.web.cern.ch/about`, August 2014.

[17] T. S. Pettersson and P. Lefèvre, “The Large Hadron Collider: conceptual design,” Tech. Rep. CERN-AC-95-05-LHC, CERN, Geneva, Oct 1995.

[18] F. Pauss, “The LHC: challenges on the way to the Higgs discovery.” `http://indico.cern.ch/event/277650/session/9/contribution/350`, Aug. 2014. 3rd International Conference on New Frontiers in Physics.

[19] C. Lefevre, “LHC: the guide (English version),” Nov. 2011. CERN-Brochure-2011-005-Eng.

[20] CERN, “CERN Document Server: Photos.” `http://cdsweb.cern.ch/collection/Photos`, Mai 2011.

[21] C. Schmitt, “Studien zum Underlying Event bei LHC Energien,” Diplomarbeit, Ludwig-Maximilians-Universität München, June 2011.

[22] *ATLAS detector and physics performance: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1999.

[23] *ATLAS central solenoid: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1997. Electronic version not available.

[24] J. P. Badiou *et al.*, *ATLAS barrel toroid: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1997. Electronic version not available.

[25] *ATLAS magnet system: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1997.

[26] *ATLAS inner detector: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1997.

[27] G. Aad *et al.*, “Atlas pixel detector electronics and sensors,” *Journal of Instrumentation*, vol. 3, no. 07, p. P07007, 2008.

[28] A. Vogel, “ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates,” Tech. Rep. ATL-INDET-PROC-2013-005, CERN, Geneva, Apr 2013.

[29] *ATLAS calorimeter performance: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1996.

[30] *ATLAS tile calorimeter: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1996.

[31] M. Leite, "Performance of the atlas zero degree calorimeter," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE*, pp. 1–5, Oct 2013.

[32] *ATLAS muon spectrometer: Technical Design Report.* Technical Design Report ATLAS, Geneva: CERN, 1997.

[33] J. Masik (ATLAS TDAQ), "ATLAS Trigger System in the first collisions," Oct. 2008. 11th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD08), Siena (Italy).

[34] I. Bird, "LHC Computing, CERN & Federated Identities." `http://indico.cern.ch/event/358127/contribution/1`, Feb 2015. FIM4R Workshop, CERN.

[35] J. A. Aguilar-Saavedra, "Protos - PROgram for TOp Simulations, MC Generator." `http://jaguilar.web.cern.ch/jaguilar/protos/`, Sept. 2014.

[36] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *Journal of High Energy Physics*, vol. 2002, no. 06, p. 029, 2002.

[37] H1 and ZEUS Collaborations, "Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA," 2010.

[38] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms," *Journal of High Energy Physics*, vol. 2004, no. 11, p. 040, 2004.

[39] M. L. Mangano *et al.*, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *Journal of High Energy Physics*, vol. 2003, no. 07, p. 001, 2003.

[40] G. Corcella *et al.*, "HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)," *Journal of High Energy Physics*, vol. 2001, no. 01, p. 010, 2001.

[41] The ATLAS Collaboration, " Search for pair production of new heavy quarks that decay to a Z boson and a third generation quark in pp collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector," ATLAS conference note, ATLAS-CONF-2013-056, CERN, Geneva, June 2013.

[42] The ATLAS Collaboration, "Measurement of the charge asymmetry in top quark pair production in pp collisions at $\sqrt{s}$ =7 tev using the atlas detector," *The European Physical Journal C*, vol. 72, no. 6, 2012.

[43] T. Sjostrand *et al.*, “PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026,” *hep-ph/0603175*, 2006.

[44] K. Nakamura *et al.* (Particle Data Group), “The Review of particle physics,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 37, p. 075021, 2010.

[45] P. Z. Skands, “Tuning Monte Carlo generators: The Perugia tunes,” *Phys. Rev. D*, vol. 82, p. 074018, Oct 2010.

[46] J. Pumplin *et al.*, “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis,” *Journal of High Energy Physics*, vol. 2002, no. 07, p. 012, 2002.

[47] ATLAS Internal Documentation, “Common generator parameters in MC12.” `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/McProductionCommonParametersMC12`, Sept. 2014.

[48] S. Jadach *et al.*, “The τ decay library tauola, version 2.4,” *Computer Physics Communications*, vol. 76, no. 3, pp. 361 – 380, 1993.

[49] E. Barberio *et al.*, “Photos — A universal Monte Carlo for QED radiative corrections in decays,” *Computer Physics Communications*, vol. 66, no. 1, pp. 115 – 128, 1991.

[50] ATLAS Internal Documentation, “Top MC Validation.” `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMCValidation#ttbar_with_modified_BR_for_top_q`, Nov. 2014.

[51] C. Escobar. Private communication.

[52] ATLAS Internal Documentation, “D3PDMaker.” `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TruthD3PDMaker`, Sept. 2014.

[53] Wim Lavrijsen, “ATLAS offline software Athena-based tutorials.” `http://wlav.home.cern.ch/wlav/tutorials/`, Sept. 2014.

[54] S. Agostinelli *et al.*, “Geant4—a simulation toolkit,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003.

[55] Wolfgang Lukas (on behalf of the ATLAS Collaboration), “Fast simulation for atlas: Atlfast-ii and isf,” *Journal of Physics: Conference Series*, vol. 396, no. 2, p. 022031, 2012.

[56] M. zur Nedden. Private communication.

[57] ATLAS Internal Documentation, “Top group’s MC12 Samples For 2012 Data Analyses.” `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC12`, Sept. 2014.

[58] ATLAS Internal Documentation, "AtlasProductionGroupMC12a." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasProductionGroupMC12a`, Sept. 2014.

[59] The ATLAS Collaboration, "New ATLAS event generator tunes to 2010 data," Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011.

[60] H.-L. Lai *et al.*, "New parton distributions for collider physics," *Phys. Rev. D*, vol. 82, p. 074024, Oct 2010.

[61] ATLAS Internal Documentation, "Top MC12 SingleTop Samples." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC12SingleTopSamples`, Nov. 2014.

[62] ATLAS Internal Documentation, "Top MC12 Wjets Samples." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC12WjetsSamples`, Nov. 2014.

[63] ATLAS Internal Documentation, "Top MC12 Zjets Samples." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC12ZjetsSamples`, Nov. 2014.

[64] ATLAS Internal Documentation, "Top MC12 DiBoson Samples." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC12DiBosonSamples`, Nov. 2014.

[65] The 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 TeV with the ATLAS detector," ATLAS conference note, ATLAS-CONF-2014-058, CERN, Geneva, Sept. 2014.

[66] L. Heinrich, "Calibration of b-Tagging Algorithms via a Kinematic Fit of Top Quark Pair Decays," Master thesis, Humboldt-Universität zu Berlin, Aug. 2012.

[67] B. Acharya *et al.*, "Object selection and calibration, background estimations and MC samples for top quark analyses using the full 2012 data set," ATLAS Internal Communications: ATL-COM-PHYS-2013-1016, CERN, Geneva, Aug. 2013.

[68] ATLAS Internal Documentation, "ATLAS Luminosity Calculator ." `https://atlas-lumicalc.cern.ch/`, Oct. 2014.

[69] S. van der Meer, "Calibration of the effective beam height in the ISR," Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[70] The ATLAS collaboration, "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC," *Eur.Phys.J.*, vol. C73, p. 2518, 2013.

[71] ATLAS Luminosity Group, "Preliminary Luminosity Determination in pp Collisions at $\sqrt{s} = 8$ TeV using the ATLAS Detector in 2012," ATLAS conference note, ATL-COM-LUM-2012-013, CERN, Geneva, Nov 2012.

[72] The ATLAS collaboration, "Monitoring and data quality assessment of the atlas liquid argon calorimeter," *Journal of Instrumentation*, vol. 9, no. 07, p. P07024, 2014.

[73] ATLAS Public Results, "COMA Period Documentation." `https://atlas-tagservices.cern.ch/tagservices/RunBrowser/runBrowserReport/rBR_Period_Report.php`, Oct. 2014.

[74] ATLAS Public Results, "Luminosity Public Results." `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`, Oct. 2014.

[75] The ATLAS collaboration, "Search for s-channel single top-quark production in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *arXiv:1410.0647*, Oct. 2014.

[76] B. Acharya *et al.*, "Object selection and calibration, background estimations and MC samples for the Winter 2013 Top Quark analyses with 2012 data," ATLAS Internal Communications: ATL-COM-PHYS-2013-088, CERN, Geneva, July 2013.

[77] ATLAS Internal Documentation, "Physics Analysis Tools - User Analysis Tools." `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PATUserAnalysisTools#egamma_AnalysisUtils_EisoTool_Ei`, Oct. 2014.

[78] T. Lagouri *et al.*, "A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN (ATL-CONF-2003-011)," *IEEE Transactions on Nuclear Science,*, vol. 51, pp. 3030–3033, Dec 2004.

[79] ATLAS Internal Documentation, "ATLAS Muon Combined Performance Guidelines for Analyses of 2012 Data ." `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesData2012`, Nov. 2014.

[80] ATLAS Internal Documentation, "Cosmic Tagging Tool - Mauro Iodice." `https://indico.cern.ch/event/141181/`, May 2011.

[81] The ATLAS Collaboration, "dE/dx measurement in the ATLAS Pixel Detector and its use for particle identification," ATLAS conference note, ATLAS-CONF-2011-016, CERN, Geneva, March 2011.

[82] G. R. Farrar and P. Fayet, "Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry," *Phys.Lett.*, vol. B76, pp. 575–579, 1978.

[83] ATLAS Internal Documentation, "Missing transverse energy." `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EtMiss`, Oct. 2014.

[84] M. Cacciari *et al.*, "The anti-k_t jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.

[85] The ATLAS collaboration, "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV," *The European Physical Journal C*, vol. 73, no. 3, 2013.

[86] The ATLAS Collaboration, "Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} =$ TeV with ATLAS 2011 data," ATLAS conference note, ATLAS-CONF-2013-004, CERN, Geneva, Jan 2013.

[87] A. B. Galtieri *et al.*, "Precision measurements of the top quark mass from the Tevatron in the pre-LHC era," *Reports on Progress in Physics*, vol. 75, no. 5, p. 056201, 2012.

[88] S. Argyropoulos *et al.*, "Pile-up subtraction and suppression for jets in ATLAS," ATLAS Internal Communications: ATL-COM-PHYS-2013-251, CERN, Geneva, Mar 2013.

[89] I. Riu, "The Readout System of the ATLAS Liquid Argon Calorimeters," September 2002. 8th Workshop on Electronics for LHC Experiments (Colmar, France).

[90] The ATLAS Collaboration, "Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data," ATLAS conference note, ATLAS-CONF-2011-102, CERN, Geneva, July 2011.

[91] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 262, no. 2–3, pp. 444 – 450, 1987.

[92] The ATLAS Collaboration, "Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment," ATLAS conference note, ATLAS-CONF-2014-004, CERN, Geneva, Feb 2014.

[93] The ATLAS Collaboration, "Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data," ATLAS conference note, ATLAS-CONF-2014-046, CERN, Geneva, July 2014.

[94] The D0 Collaboration, "b-Jet identification in the D0 experiment," *Nuclear Instruments and Methods in Physics A*, vol. 620, no. 2–3, pp. 490 – 517, 2010.

[95] ATLAS Internal Documentation, "MV3 Status Update - K. Mochizuki, M. Ughetto, L. Vacavant." `https://indico.cern.ch/event/237798/`, Feb. 2013.

[96] The CMS collaboration, "Measurement of the ratio $B(t \to Wb)/B(t \to Wq)$ in pp collisions at $\sqrt{s} = 8$ tev," *Physics Letters B*, vol. 736, no. 0, pp. 33 – 57, 2014.

[97] The ATLAS Collaboration, "Kinematic Fitting of ATLAS Data Using the KinFitter Package," ATLAS Internal Communications: ATL-COM-PHYS-2012-1554, CERN, Geneva, Oct 2012.

[98] R. Herrberg-Schubert, *Cross-Section Measurement of Single-Top t-Channel Production at ATLAS.* Dissertation, Humboldt-Universität zu Berlin, April 2014.

[99] L. Lyons, *Statistics for Nuclear and Particle Physicists.* Cambridge University Press, 1989.

[100] The ATLAS Collaboration, "KinFitter – A Kinematic Fit with Constraints," Sept. 2009.

[101] P. Rieck, "Entwicklung eines kinematischen Fits zur Untersuchung elektroschwacher Top-Quark-Produktion bei ATLAS," Master thesis, Humboldt-Universität zu Berlin, Dec. 2010.

[102] O. M. Kind *et al.*, "A++ framework." `http://ms2.physik.hu-berlin.de/~atlas/aplusplus/htmldoc/index.html`, Jan. 2015.

[103] J. Beringer *et al.* (Particle Data Group), "The Review of particle physics," *Physical Review D*, vol. 86, p. 010001, 2012.

[104] The ATLAS Collaboration, "Measuring the b-tag efficiency in a $t\bar{t}$ sample with 4.7 fb^{-1} of data from the ATLAS detector," ATLAS conference note, ATLAS-CONF-2012-097, CERN, Geneva, July 2012.

[105] F. Thomas, "Study of the Properties of Additional Jets in Top Pair Events with the ATLAS Detector," Diplomarbeit, Technische Universität Dresden, June 2011.

[106] O. Biebel, *Untersuchungen zur Unabhängigkeit der starken Wechselwirkung von den Flavour-Quantenzahlen mit Bottom-, Charm-, Strange- und leichten Quarks.* Dissertation, Universität Bonn, November 1993.

[107] P. Igo-Kemenes *et al.*, "Z physics at LEP 1: Heavy Flavours at LEP," *CERN-89-08 Vol. 1*, p. 284, 1989.

[108] F. Halzen and A. D. Martin, *Quarks and Leptons.* John Wiley & Sons, 1984.

[109] P. Mättig, "The structure of jets in $\mathrm{e}^+\mathrm{e}^-$ collisions," *Physics Reports*, vol. 177, no. 3–4, pp. 141 – 317, 1989.

[110] B. Ioffe, "Associated production of gluonic jets and heavy mesons in e^+e^- annihilation," *Physics Letters B*, vol. 78, no. 2–3, pp. 277 – 280, 1978.

[111] A. Ballestrero *et al.*, "Heavy quark production at e^+e^- colliders in three- and four-jet events," *Physics Letters B*, vol. 294, no. 3–4, pp. 425 – 430, 1992.

[112] OPAL Collaboration, "A study of K_S^0 production in Z^0 decays," *Physics Letters B*, vol. 264, pp. 467 – 475, 1991.

[113] J. Binnewies *et al.*, "Neutral kaon production in e^+e^-, ep , and $pp^{\,-}$ collisions at next-to-leading order," *Phys. Rev. D*, vol. 53, pp. 3573–3581, Apr 1996.

[114] J. Binnewies, “Fragmentation Functions in Next-To-Leading Order QCD,” *arXiv:hep-ph/9707269*, July 1997.

[115] M. Hauschild *et al.*, “Particle identification with the OPAL jet chamber,” *Nucl.Instrum.Meth.*, vol. A314, pp. 74–85, 1992.

[116] C. G. Ivan, *Open charm analysis with the ALICE detector in pp collisions at LHC.* Dissertation, Universiteit Utrecht, May 2010. arXiv:1005.4954.

[117] ZEUS Heavy Flavour Group, “Heavy Flavour Event Displays.” `http://www-zeus.desy.de/physics/hfla/public/displays/`, Dec. 2014.

[118] A. Hoecker *et al.*, “TMVA: Toolkit for Multivariate Data Analysis,” *PoS*, vol. ACAT, p. 040, 2007. arXiv:physics/0703039.

[119] Y. Freund, “Boosting a weak learning algorithm by majority,” *Information and Computation*, vol. 121, no. 2, pp. 256 – 285, 1995.

[120] I. Chakravarti *et al.*, *Handbook of Methods of Applied Statistics*, vol. I of *Wiley series in probability and mathematical statistics.* John Wiley and Sons, 1967. pp. 392-394.

[121] B. Alvarez Gonzalez *et al.*, “Measurements of s-channel single top quark production in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” ATLAS Internal Communications: ATL-COM-PHYS-2014-044, CERN, Geneva, Jan 2014.

[122] Nikolaos Kidonakis, “Differential and total cross sections for top pair and single top production,” *arXiv:1205.3453*, May 2012.

[123] J. Butterworth *et al.*, “Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV,” ATLAS Internal Communications: ATL-COM-PHYS-2010-695, Aug. 2010.

[124] B. Acharya *et al.*, “Estimation of the W^{+}Jets Background for Top Quark Re-Discovery in the Single Lepton+Jets Channel,” ATLAS Internal Communications: ATL-COM-PHYS-2010-834, Oct. 2010.

[125] ATLAS Internal Documentation, “Top Systematic Uncertainties for 8 TeV 2012 analyses.” `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopSystematicUncertainties`, Feb. 2015.

[126] H. Esch, *Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the t-channel using flavour-tagging and a neural network with ATLAS data at $\sqrt{s} = 8$ TeV.* PhD thesis, Dortmund U., Jan. 2015.

[127] The ATLAS collaboration, “Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV,” ATLAS conference note, ATLAS-CONF-2010-054, CERN, Geneva, Jul 2010.

[128] The ATLAS collaboration, “Jet energy resolution in proton-proton collisions at $\sqrt{s}$ =7 TeV recorded in 2010 with the ATLAS detector,” *The European Physical Journal C*, vol. 73, no. 3, 2013.

[129] The ATLAS collaboration, “Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,” *The European Physical Journal C*, vol. 75, no. 1, 2015.

[130] K. Cranmer *et al.*, “HistFactory: A tool for creating statistical models for use with RooFit and RooStats,” *CERN-OPEN-2012-016*, June 2012.

[131] L. Moneta *et al.*, “The RooStats Project,” *PoS*, vol. ACAT2010, p. 057, 2010. arXiv:1009.1003.

[132] G. Cowan *et al.*, “Asymptotic formulae for likelihood-based tests of new physics,” *The European Physical Journal C*, vol. 71, no. 2, 2011.

[133] A. L. Read, “Presentation of search results: the CL_s technique,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, no. 10, p. 2693, 2002.

[134] LHC Physics Web, “NNLO+NNLL top-quark-pair cross sections.” `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO`, Feb 2015.

[135] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders,” *Computer Physics Communications*, vol. 185, no. 11, pp. 2930 – 2938, 2014.

[136] G. Cowan *et al.*, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur.Phys.J.*, vol. C71, p. 1554, 2011.

[137] The CDF collaboration, “First Measurement of the Ratio $B(t \to Wb)/B(t \to Wq)$ and Associated Limit on the Cabibbo-Kobayashi-Maskawa Element $|V_{tb}|$,” vol. 86, pp. 3233–3238, Apr 2001.

[138] The CMS collaboration, “Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV,” *Journal of High Energy Physics*, vol. 2014, no. 6, 2014.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich während meiner Promotion und der Anfertigung dieser Arbeit großartig unterstützt haben.
Vielen herzlichen Dank an ...

- Prof. Dr. Otmar Biebel für die gute Betreuung meiner Arbeit und das spannende Thema zur Bestimmung des CKM Matrix-Elements $|V_{ts}|^2$. Auch wenn er mit vielen andere Aufgaben beschäftigt war, so stand seine Tür stets offen.
- Prof. Dr. Christian Kiesling für die Erstellung des Zweitgutachtens und für die prompte Zusage in meiner Promotionskommission mitzuwirken.
- die weiteren Mitglieder meiner Promotionskommission, Prof. Dr. Bender, Prof. Dr. Buchalla, PD Dr. Thirolf, PD Dr. Elmsheuser, für das Mitwirken zum Erlangen meines Doktorgrades.
- Prof. Dr. Dorothee Schaile für die Unterstützung, das gute Arbeitsklima am Lehrstuhl und die tollen Erfahrungen, die sie mir durch Konferenzbesuche und den Aufenthalt am CERN ermöglicht hat.
- Sören Stamm der HU Berlin für die großartige Unterstützung, ohne die meine Arbeit nicht zu dem geworden wäre, was sie ist. Danke für die tolle Zeit in Berlin.
- Dr. Martin zur Nedden der HU Berlin für die ausgiebigen Diskussionen und die wöchentlichen Skype-Meetings, während denen wir viele neue Ideen für die weitere Analyse entwickelten.
- Dr. Oliver Maria Kind der HU Berlin für die großartige Zusammenarbeit, Unterstützung und die Integration in die Berliner Arbeitsgruppe. Dabei durfte ich viel Neues über physikalische und programmiertechnische Methoden lernen.
- Patrick Rieck der HU Berlin für die ausführlichen Antworten bei Fragen.
- David Salek und Maurice Iodice für die Betreuung bei der Entwicklung einer neuen Identifikationsmethode zur Rekonstruktion kosmischer Myonen mit dem ATLAS Detektor.
- Michael Bender für das ausgiebige und zügige Korrekturlesen dieser Arbeit, auch während seines Urlaubs! Darüber hinaus natürlich auch für die tolle Zeit in Garching und in der Freizeit.
- Dr. Josephine Wittkowski für das gute Büroklima, die ausgiebigen Gespräche und die Zeit außerhalb der Büros.
- Friedrich Hönig für die Hilfe bei einigen Code-Problemen und die vielen kurzen Pausen in Michis Büro - und natürlich für die *Fun Facts.*

- Clara Hofmann, Pauline Drews und Franz Fuchs für die schöne gemeinsame Zeit in Thoiry während meines CERN-Aufenthalts.

- Philippe Calfayan und Bonnie Chow für die spannende Zeit am CERN und die gemeinsamen Kochabende.

- Elke Grimm-Zeidler und ihre Vorgängerin Herta Franz für die Hilfe bei administrativen Problemen, bei denen sie mir immer freundlich mit Rat und Tat behilflich waren.

- Felix Rauscher für die stets schnellen und guten Antworten bei Problemen mit Computer und Netzwerk.

- alle IPP-"Mitläufer", die sich mit mir das Essen in der Kantine schmecken lassen, für die spannenden und lustigen Gespräche.

- alle aktiven und ehemalige Mitglieder des Lehrstuhls Schaile für die tolle Zeit und die gute Atmosphäre.

- meine Familie, Ursula und Georg Schmitt, sowie Katja und Stefan Obermeier, dass sie für mich da sind und mich immer unterstützen.

- meine Frau Pia für das intensive Korrekturlesen dieser Arbeit. Danke außerdem für deine stete moralische Unterstützung, besonders während der Endphase der Promotion, und vor allem für deine Liebe! Danke, dass du immer für mich da bist!