ETP-KA/2020-9

Karlsruhe Institute of Technology

Search for single top quark production in the s channel at 13 TeV with the CMS experiment

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

M.Sc. Denise Müller

aus Mosbach

Tag der mündlichen Prüfung:	5. Juni 2020
Referent:	Prof. Dr. Thomas Müller
Korreferent:	Prof. Dr. Ulrich Husemann

Erklärung der selbstständigen Anfertigung der Dissertationsschrift

Hiermit erkläre ich, dass ich die Dissertation mit dem Titel

Search for single top quark production in the s channel
at 13 TeV with the CMS experiment

selbstständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht habe, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde.

Ich versichere außerdem, dass ich die Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Karlsruhe, den 6. Mai 2020

. . . . . .
(Denise Müller)

Abstract

A search for s-channel single top quark production is presented, using 137 fb^{-1} of data recorded at a center-of-mass energy of 13 TeV by the CMS experiment at the CERN LHC. In the search, leptonically decaying top quarks are analyzed, resulting in a final state of the signal process consisting of one charged muon or electron and its corresponding neutrino, and two jets originating from hadronized bottom quarks. Depending on the jet and b jet multiplicities, signal and control event categories, accounting for different background processes, are defined. All signal and background processes are modeled via event simulation, except for the QCD multijet background contribution, which is modeled with a data-driven approach. A dedicated multivariate method is employed to enhance the separation between the signal process and the background processes, mainly top quark pair production. By performing a simultaneous fit on the multivariate discriminator in all event categories, the significance for observing s-channel single top quark production, as well as the inclusive cross section and the absolute value of the CKM matrix element V_{tb}, are extracted.

With an observed (expected) significance of 6.0 (4.7) standard deviations, this analysis is the first discovery for s-channel single top quark production using proton-proton collisions of the LHC. The measured cross section of this process is $\sigma_{s\text{-ch.}} = 14.65^{+2.48}_{-2.37}$ (stat. + syst.) pb and V_{tb} is measured to be $|f_{\text{LV}}V_{\text{tb}}| = 1.19 \pm 0.09\,(\text{exp.}) \pm 0.02\,(\text{theo.})$, with the form factor for potential anomalous left-handed vector boson couplings f_{LV} being 1 in the SM. These measured values are both within two standard deviations of the SM predictions.

Introduction

Since the postulation of the standard model (SM) of particle physics in the 1960s and 1970s, all predicted, yet unobserved, elementary particles have been discovered, with the last missing piece, the Higgs boson, being observed for the first time in 2012 at the Large Hadron Collider (LHC) near Geneva [1, 2]. Despite being now complete and demonstrated to be successful in numerous aspects, there are still open questions that cannot be answered by the SM: for instance, it does not describe gravity and does not provide any explanation for the existence of dark matter. As new particles explaining these phenomena have not been discovered so far in direct searches, indirect searches for new physics need to be performed in addition. Therefore, it is essential to measure all processes predicted by the SM, determine the properties of SM particles as accurate as possible, and search for potentially existing small deviations that can be a hint for physics beyond the SM.

As the top quark is the heaviest elementary particle of the SM, with its mass comparable to the mass of a gold atom, it may play a special role in the SM, in particular in the electroweak symmetry breaking mechanism. It was discovered in 1995 at the Tevatron [3, 4], a proton-antiproton collider situated at Fermilab, where only about 1000 top quark pairs were produced by that time. Nowadays, the LHC can be considered a top quark factory. During Run 2 of the LHC, proton-proton collisions were recorded from 2015 to 2018 at a center-of-mass energy of 13 TeV, in which more than 200 million top quark pairs and more than 80 million single top quarks were expected to be produced. Although being less often created than top quark pairs, single top quarks are of special interest for probing the electroweak sector of the SM. For example, single top quark processes enable direct measurements of the Cabibbo–Kobayashi–Maskawa matrix element $|V_{\mathrm{tb}}|$ and are suitable to search for possible anomalous Wtb couplings [5]. The production of single top quarks was first discovered at the Tevatron in 2009 [6, 7]. During Run 1 of the LHC, which comprises measured data from 2010 to 2012 at center-of-mass energies of 7 TeV and 8 TeV, and Run 2, two of the three main production modes for single top quarks, the t-channel process and the associated production with a W boson, were already successfully measured. However, the rarest of the three in proton-proton collisions, the s-channel process, has not been observed at the LHC so far. The aim of this thesis is therefore the observation of s-channel single top quark production by analyzing the 2016, 2017, and 2018 datasets recorded by the Compact Muon Solenoid (CMS) experiment. In particular, small deviations in the measured cross section of this process could be caused by contributions from new

physics, which could manifest in the detector with the same signature as s-channel single top quark production.

The structure of this thesis is as follows: In the first chapter, an overview of the theoretical description of the SM and of the top quark properties is given. The second chapter presents the design of the LHC and the CMS experiment. The simulation and reconstruction of events is described in the third chapter. All multivariate machine learning methods and statistical methods employed in this thesis are explained in the fourth chapter. The search for s-channel single top quark production is presented in the fifth chapter, where each step of the analysis is described in detail, starting with the search strategy and ending with the discussion of the measured results. A summary of the analysis, possible improvements for the search for s-channel single top quark production, and an outlook for future single top quark analyses at the LHC and at future electron-positron colliders are provided in the last chapter.

Contents

1. Theoretical motivation

The properties of all known elementary particles and their interactions are described by the standard model (SM) of particle physics, developed in the 1960s and 1970s. Many new particles described by this fundamental theory have been discovered afterwards, for instance the top quark in 1995 at the Tevatron [3, 4]. In 2012, the last missing elementary particle of the SM, the Higgs boson, has been observed at the Large Hadron Collider (LHC) [1, 2].

In this chapter, an overview of the fundamental forces and the elementary particles of the SM is given, with emphasis on the properties of the top quark. Natural units ($\hbar = c = 1$) are used for simplification.

1.1 Standard model of particle physics

The SM is a quantum field theory (QFT), combining the principles of special relativity and quantum mechanics. It unites two theories, quantum chromodynamics (QCD) [8, 9] describing the strong interaction, and the electroweak theory [10], describing the unification of electromagnetic and weak interaction. The SM is represented by the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ symmetry group. Noether's theorem [11] states that each continuous symmetry of a physical system leads to a conserved quantity. In case of the SM, the conserved quantum numbers, namely color charge (C), the weak isospin (L), and the weak hypercharge (Y) are assigned to each elementary particle. Various fields, corresponding to the forces and particles of the SM and whose excitations can be interpreted as particles observed in nature, are defined in the QFT by a Lagrangian density:

$$\mathcal{L}_\mathrm{SM} = \mathcal{L}_\mathrm{gauge} + \mathcal{L}_\mathrm{fermion} + \mathcal{L}_\mathrm{Higgs} + \mathcal{L}_\mathrm{Yukawa}. \tag{1.1}$$

The term $\mathcal{L}_\mathrm{gauge}$ describes the interaction of gauge bosons, while the dynamics of fermions are grouped into $\mathcal{L}_\mathrm{fermion}$. The SM Lagrangian density is completed by the Higgs sector

($\mathcal{L}_{\text{Higgs}}$) being correlated with the electroweak symmetry breaking and the fermion mass terms due to their Yukawa coupling ($\mathcal{L}_{\text{Yukawa}}$) to the Higgs field.

Despite the fact that all SM particles have been observed in nature and particles not included in the SM have not been discovered so far, there are still some phenomena which cannot be solved by the SM. For instance, atmospheric and solar neutrino oscillations have been observed [12, 13], proving that they are not massless as built into the SM by construction. The SM is furthermore incapable of explaining the existence of dark matter and neglects gravity, an important interaction experienced in everyday life. Thus, it is crucial to measure the properties of the SM particles as precise as possible to search for small deviations that may hint to new physics.

According to their spin, the elementary particles of the SM are classified into particles with integer spin and spin-1/2 particles, called bosons and fermions, respectively. In the following, these two particle groups are described in more detail, alongside with essential mechanisms of the SM.

1.1.1 Bosons and fermions

The SM comprises four types of gauge bosons, the Higgs boson, and six different flavors of quarks and leptons. An overview of the elementary particles of the SM is given in Fig. 1.1.

All gauge bosons are vector bosons with spin 1, whereas the Higgs boson is a scalar boson with spin 0. The strong, electromagnetic, and weak force between elementary particles are mediated by different gauge bosons listed in Table 1.1. The strong interaction is transmitted by eight massless gluons (g), which couple to color-charged particles. As each gluon carries one color and one anticolor charge, they can interact among themselves, thus limiting the range of the strong force. The massive $W^{\pm}$ bosons and the Z boson mediate the weak interaction. As a result of their masses of about 80 GeV and 91 GeV, respectively, the weak interaction has only a short range. The only interaction with an infinite range is thus the electromagnetic interaction mediated by the massless photon (γ), which carries no electric charge and which can only couple to electrically charged particles. The Higgs boson is the excitation of the Higgs field and couples to the massive gauge bosons through the Higgs mechanism, which is discussed in Section 1.1.3.

SM fermions are spin-1/2 particles, which can be separated into two groups based on their color charge: quarks and leptons. They are further divided into three generations, ordered by their masses. Each generation consists of two quarks and two leptons, which are distinguished by their weak isospin. The properties of all fermions are given in Table 1.2. For all fermions, an antifermion exists, with identical mass but different quantum numbers, such as electric charge and parity.

Six different flavors of quarks exist in the SM, which carry color and electric charge. Quarks carrying an electric charge of $+2/3\,e$ are referred to as up-type quarks, whereas quarks with an electric charge of $-1/3\,e$ are called down-type quarks. Each generation has one up- and one down-type quark. The first generation consists of the up (u) and down (d) quark, which are the lightest quarks. The charm (c) and strange (s) quark form the second generation,

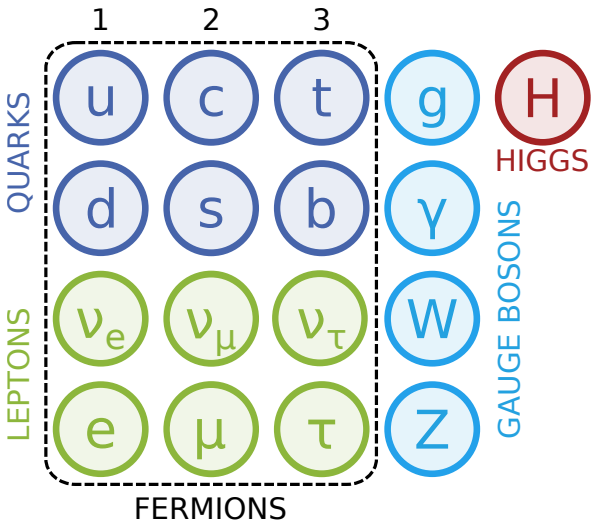

Figure 1.1: The standard model of particle physics. The SM contains fermions, divided into six different flavors of quarks and leptons and their corresponding antiparticles, and bosons, four different types of gauge bosons as force mediators, and the Higgs boson. The fermions are further split into three generations.

Table 1.1: The gauge bosons of the standard model. Gauge bosons are the mediators for three different types of particle interaction: strong, electromagnetic, and weak. The electric charge and the mass of each gauge boson are given. All values taken from Ref. [14].

Force	Gauge boson	charge	Mass (GeV)
Strong	Gluons (g)	0	0
Electromagnetic	Photon (γ)	0	0
Weak	W bosons ($W^{\pm}$)	± 1	80.379 ± 0.012
	Z boson (Z)	0	91.1876 ± 0.0020

Table 1.2: The fermions of the standard model. The fermions are divided into two groups, quarks and leptons, which are further split into three generations. In contrast to quarks, leptons do not carry color charge. The fermions are ordered by the third component of their weak isospin and by their generation. All values taken from Ref. [14].

Fermions	Generation			Electric	3^{rd} comp.
	1	2	3	charge (e)	of isospin
Quarks	u	c	t	$+\frac{2}{3}$	$+\frac{1}{2}$
	d	s	b	$-\frac{1}{3}$	$-\frac{1}{2}$
Leptons	ν_e	ν_μ	ν_τ	0	$+\frac{1}{2}$
	e	μ	τ	-1	$-\frac{1}{2}$

while the third generation includes the heaviest quark, the top (t) quark, and the bottom (b) quark. Quarks are the only SM particles that can take part in all three interactions. In particle detectors, isolated quarks cannot be observed. This is caused by the so-called confinement, where the self-interaction of gluons leads to a flux tube in the color field between the quarks when trying to separate them. With increasing distance, it is energetically more favorable for the quarks to create a new quark-antiquark pair. Consequently, quarks form color-neutral bound states called hadrons, which can be observed in nature and can be distinguished by their quark content. This mechanism is called hadronization. Hadrons consisting of a quark-antiquark pair are called mesons, while bound states formed by three quarks are referred to as baryons. As the heavier quarks are not stable, most hadrons are formed by first-generation quarks, e. g., protons (uud) and neutrons (udd). At high energies, corresponding to small distances, quarks behave like free particles, as the strength of their interactions decreases. This phenomenon is called asymptotic freedom and allows perturbative calculations of strong interactions, which are not possible at low energies.

In contrast to quarks, leptons do not carry color charge and thus do not couple to the strong force. They can be distinguished by their electric charge: charged leptons with an electric charge of $-e$ and electrically neutral neutrinos. Each generation consists of one charged lepton and one neutrino. The first generation is formed by the electron (e) and the electron neutrino (ν_e), the second generation by the muon (μ) and the muon neutrino (ν_μ), and the third generation by the tau lepton (τ) and the tau neutrino (ν_τ). Because of their electric charge, charged leptons couple to the electromagnetic and weak force, whereas neutrinos only interact weakly. According to the SM, neutrinos are massless, but is has been experimentally proven that they can change their flavor via oscillations, which is only possible if they are massive. Hence, neutrinos do have small masses, which will be measured at different experiments. For instance, the KATRIN experiment [15] is designed to measure the electron neutrino mass and has recently reported the best upper limit for direct measurement of the neutrino mass [16].

1.1.2 Electroweak interaction

The unification of the electromagnetic and weak interaction by Weinberg and Salam in the 1960s [10] was a great success of the SM, resolving divergences in the weak interaction by replacing the contact interactions of the Fermi theory [17] with the exchange of vector bosons. The simplest symmetry group for the combination of these two forces into the electroweak interaction is the $\mathrm{SU(2)_L} \times \mathrm{U(1)}_Y$ group, where $\mathrm{SU(2)_L}$ describes the weak isospin acting only on left-handed fermions because of parity violation in the weak interaction [18], and where $\mathrm{U(1)}_Y$ corresponds to the weak hypercharge Y acting on all particles.

According to the weak isospin quantum number I and the value of its third component I_3, fermions are grouped into left-handed weak isospin doublets L ($I = 1/2$, $I_3 = \pm 1/2$) and right-handed weak isospin singlets R ($I = I_3 = 0$):

$$L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \ldots, \begin{pmatrix} u \\ d \end{pmatrix}_L, \ldots, \tag{1.2}$$

$$R = e^-_R, \ldots, u_R, d_R, \ldots. \tag{1.3}$$

As right-handed neutrinos are not included the SM and do not couple to any known particle, they only occur in left-handed weak isospin doublets. The gauge transformation of the $\mathrm{SU(2)_L}$ symmetry group is given by $U(x) = \exp\left[\mathrm{i}(\sigma_k/2)\alpha^k(x)\right]$, with the local phase $\alpha^k(x)$ and the three generators $\sigma_k/2$, where σ_k ($k = 1, 2, 3$) can be represented by the Pauli matrices. Three gauge bosons W^k_μ are obtained in total, where combinations of W^1_μ and W^2_μ are directly related to the massive $\mathrm{W}^\pm$ bosons:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp \mathrm{i}W^2_\mu\right). \tag{1.4}$$

The gauge boson W^3_μ cannot be directly associated with the massive Z boson, as no coupling to right-handed fermions is allowed. The gauge transformation of the $\mathrm{U(1)}_Y$ symmetry group is given by $U(x) = \exp\left[\mathrm{i}\frac{Y}{2}\mathbb{1}_2\alpha(x)\right]$, with the generator $\mathbb{1}_2$ and the weak hypercharge Y, defined as a linear combination of electric charge Q and I_3:

$$Y = 2(Q - I_3), \tag{1.5}$$

and one resulting gauge boson B_μ.

For the left- and right-handed fermions, different covariant derivatives can be defined:

$$D_\mu L = \left(\partial_\mu - \mathrm{i}g\frac{\sigma_k}{2}W^k_\mu - \mathrm{i}g'\frac{Y}{2}\mathbb{1}_2 B_\mu\right)L, \tag{1.6}$$

$$D_\mu R = \left(\partial_\mu - \mathrm{i}g'\frac{Y}{2}B_\mu\right)R, \tag{1.7}$$

with the two coupling constants g and g' for the weak isospin and the weak hypercharge, respectively. There are no mass terms in the resulting Lagrangian density as these are not gauge invariant under local $\mathrm{SU(2)_L}$ transformations. On the other hand, the $\mathrm{W^{\pm}}$ and Z bosons are massive. This issue is resolved with the mechanism explained in the following section.

1.1.3 Electroweak symmetry breaking and Higgs mechanism

The Higgs mechanism [19, 20, 21] generates gauge boson mass terms in the Lagrangian density without violating gauge invariance through spontaneous electroweak symmetry breaking [22]. The spontaneous symmetry breaking is achieved by introducing a new scalar field Φ, called Higgs field, that is symmetric under the $\mathrm{SU(2)_L} \times \mathrm{U(1)}_Y$ group and leads to three massive gauge bosons and one massless photon:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.8}$$

The Higgs field corresponds to an $\mathrm{SU(2)_L}$ doublet consisting of two complex fields ϕ^+ and ϕ^0. Its quantum numbers are given by $I = 1/2$, $I_3 = \pm 1/2$, and $Y = 1$. The Lagrangian density of the Higgs field can be written as

$$\mathcal{L}_{\mathrm{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \tag{1.9}$$

with the covariant derivative

$$D_\mu = \left(\partial_\mu - \mathrm{i}g \frac{\sigma_k}{2} W_\mu^k - \mathrm{i}g' \frac{Y}{2} \mathbb{1}_2 B_\mu \right) \tag{1.10}$$

and the Higgs potential

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{1.11}$$

The form of the Higgs potential depends on the sign of μ^2. For $\mu^2 < 0$, only one global minimum exists at $\Phi = 0$. In case of $\mu^2 > 0$, which is illustrated in Fig. 1.2, infinitely many minima exist, located along a circle in the plane spanned by the real and complex components of the Higgs field with the radius

$$v = \sqrt{\frac{\mu^2}{2\lambda}}. \tag{1.12}$$

The radius v corresponds to the vacuum expectation value of the Higgs field and is calculated as

$$v = \left(\sqrt{2} G_{\mathrm{F}} \right)^{-1/2} = 246\,\mathrm{GeV}, \tag{1.13}$$

with the Fermi coupling constant G_F [14].

In order to conserve the symmetry of the electromagnetic U(1) subgroup, which is broken by the infinite number of minima of $V(\Phi)$, the charged component ϕ^+ of the Higgs field must be set to zero, resulting in a massless photon. The neutral component ϕ^0 can be rewritten such that the Higgs field is given by

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.14}$$

with the scalar field $h(x)$ representing the Higgs boson.

The physical fields of the gauge bosons, the photon field A_μ and the Z field Z_μ, are obtained by diagonalizing the matrix of the kinetic term of $\mathcal{L}_\mathrm{Higgs}$:

$$\begin{pmatrix} B_\mu \\ W^0_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{W} & -\sin\theta_\mathrm{W} \\ \sin\theta_\mathrm{W} & \cos\theta_\mathrm{W} \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}, \tag{1.15}$$

where θ_W is the weak mixing angle, also called Weinberg angle, and is defined as

$$\cos\theta_\mathrm{W} = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{1.16}$$

The Weinberg angle has been measured to be $\sin^2\theta_\mathrm{W} \approx 0.23$ [14]. The mass of the Z boson is then given by the relation:

$$m_\mathrm{Z} = \frac{m_\mathrm{W}}{\cos\theta_\mathrm{W}}. \tag{1.17}$$

Through the Higgs mechanism, the masses of the gauge bosons are obtained. To generate the fermion masses, a Yukawa coupling to the Higgs field is postulated, which is gauge invariant under $\mathrm{SU(2)_L} \times \mathrm{U(1)}_Y$ transformations. In case of electrons, the corresponding Yukawa term has the following structure:

$$-y_\mathrm{e}\left(\overline{L}\Phi R + \overline{R}\Phi^\dagger L\right). \tag{1.18}$$

The mass of an electron can then be expressed in terms of the Yukawa coupling constant of the electron field y_e and the vacuum expectation value v:

$$m_\mathrm{e} = \frac{y_\mathrm{e} v}{\sqrt{2}}. \tag{1.19}$$

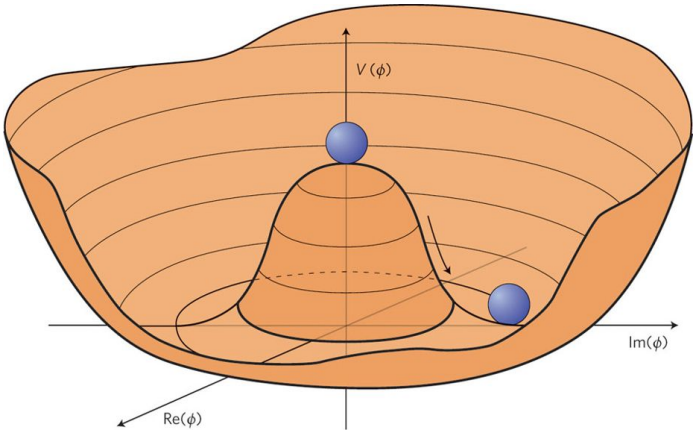

Figure 1.2: The Higgs potential. The real and imaginary parts of the Higgs field for the scenario $\mu^2 > 0$ are shown. An infinite number of minima located at a circle with radius $|\phi| = v$ exist, where v is the vacuum expectation value. Taken from Ref. [23].

1.1.4 Cabbibo–Kobayashi–Maskawa matrix

The quark masses can be derived in a similar way as the lepton masses through Yukawa couplings to the Higgs field. Through the exchange of a W boson, the flavor of a quark can be changed, leading to flavor mixing of the quarks: the mass eigenstates, which are physically observable particles, are not identical to the flavor eigenstates, describing the interactions with gauge bosons. This quark flavor mixing is expressed by the Cabbibo–Kobayashi–Maskawa (CKM) matrix V_{CKM} [24, 25]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.20}$$

with the flavor eigenstates d', s', b' and mass eigenstates d, s, b of the down-type quarks. As the CKM matrix is a unitary matrix, $V_{\mathrm{CKM}}^{\dagger} V_{\mathrm{CKM}} = \mathbb{1}_3$, various conditions for the CKM matrix elements need to be fulfilled. Using these conditions, the CKM matrix elements are precisely measured [14] to be

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}. \tag{1.21}$$

The magnitudes of the diagonal CKM matrix elements are close to 1, meaning that transitions within one generation are the most likely ones.

1.2 Top quark properties

The top quark is the heaviest particle of the SM and was discovered at the Tevatron in 1995 [3, 4]. In contrast to all other quarks, it decays before it forms bound states because of its high mass of $173.34 \pm 0.76\,\text{GeV}$ [26] and because of its possibility to decay into real W bosons. The top quark is of particular interest for the study of the Higgs boson couplings and has the highest Yukawa coupling strength. With the measured top quark mass given above and Eq. (1.19), it is determined as 0.997 ± 0.004, whereas the Yukawa coupling strength of the other SM quarks is of order 10^{-5} to 10^{-2}, depending on the quark mass.

1.2.1 Top quark pair production

The dominant top quark production mode is the production of top quark pairs ($\mathrm{t\bar{t}}$) via the strong interaction. The four relevant leading-order Feynman diagrams for $\mathrm{t\bar{t}}$ production are shown in Fig. 1.3. In the first three Feynman diagrams, initial-state gluons are involved in the $\mathrm{t\bar{t}}$ production. The production modes s and t channel directly refer to the Mandelstam variables [27], where s is defined as the square of the center-of-mass energy and t as the square of the momentum transfer. The fourth Feynman diagram shows the production of a top quark pair via quark-antiquark annihilation. As the initial-state antiquark can only arise from sea quarks, this production mode is the rarest one at the proton-proton collider LHC, but was the dominant one at the proton-antiproton collider Tevatron. The predicted cross section of $\mathrm{t\bar{t}}$ production at the LHC for a center-of-mass energy of 13 TeV is

$$\sigma_{\mathrm{t\bar{t}}} = 831.76^{+19.77}_{-29.20}\,(\text{scale}) \pm 35.06\,(\text{PDF}+\alpha_\text{s})^{+23.18}_{-22.45}\,(\text{mass})\,\text{pb}, \tag{1.22}$$

calculated for a top quark mass of 172.5 GeV [28, 29].

1.2.2 Single top quark production

Single top quark production is less frequent than $\mathrm{t\bar{t}}$ production, since single top quarks can only be produced via electroweak interactions. The single top quark process allows a direct measurement of the CKM matrix element V_tb, where small deviations could already be a hint for physics beyond the SM. The leading-order Feynman diagrams of the different single top quark production modes are shown in Fig. 1.4. Single top quarks can be produced via the t-channel process, in association with a W boson, or via the s-channel process.

In the t-channel process, which is the dominant single top quark production mode at the LHC, a light-flavored quark and a bottom quark exchange a virtual W boson in the t channel, causing both initial quarks to change their flavor. In this process, the light-flavored quark is preferably emitted in forward direction. The cross section at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ is calculated with next-to-leading-order (NLO) accuracy in QCD to be

$$\sigma_{t\text{-ch.}} = 216.99^{+6.62}_{-4.64}\,(\text{scale}) \pm 6.16\,(\text{PDF}+\alpha_\text{s})\,\text{pb}, \tag{1.23}$$

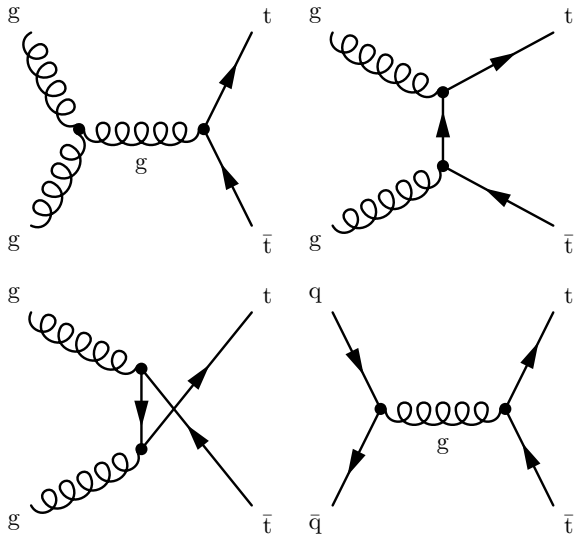

Figure 1.3: Top quark pair production. Top quark pairs are produced via the strong interaction. The dominant production mode at the LHC is gluon-gluon fusion, via the s (top left), t (top right), or u channel (bottom left). In addition, top quark pairs are also produced via quark-antiquark annihilation (bottom right).

assuming a top quark mass of 172.5 GeV [30, 31, 32]. There are higher-order cross section calculations available [33, 34], but they do not consider all uncertainty sources. For that reason, the cross section value given in Eq. (1.23) is used in ATLAS and CMS analyses.

Another possible single top quark production mode is the associated tW production, where a gluon and a bottom quark produce a top quark and a W boson. The predicted cross section at the LHC for $\sqrt{s} = 13$ TeV is [30, 35, 36]

$$\sigma_{\mathrm{tW}} = 71.7 \pm 1.80\,(\mathrm{scale}) \pm 3.40\,(\mathrm{PDF}{+}\alpha_{\mathrm{s}})\,\mathrm{pb}. \tag{1.24}$$

The third production mode is the single top quark s-channel process, where the top quark is produced together with a bottom quark through quark-antiquark pair annihilation into a virtual W boson. As an initial-state antiquark is required and the probability to observe an antiquark inside the proton with the relevant proton momentum fraction for single top quark production is small, s-channel production is the rarest single top quark process at the LHC. The process is predicted at 13 TeV with a small cross section [30, 31, 32] of

$$\sigma_{s\text{-ch.}} = 10.32^{+0.29}_{-0.24}\,(\mathrm{scale}) \pm 0.27\,(\mathrm{PDF}{+}\alpha_{\mathrm{s}})\,\mathrm{pb} \tag{1.25}$$

at NLO accuracy in QCD. Similar to the t-channel process, a higher-order cross section calculation is available [37], but does not consider all systematic uncertainties. Single top quark production in the s-channel has only been observed at the Tevatron [38], while there is only evidence so far at the LHC [39]. Small deviations in the measured s-channel

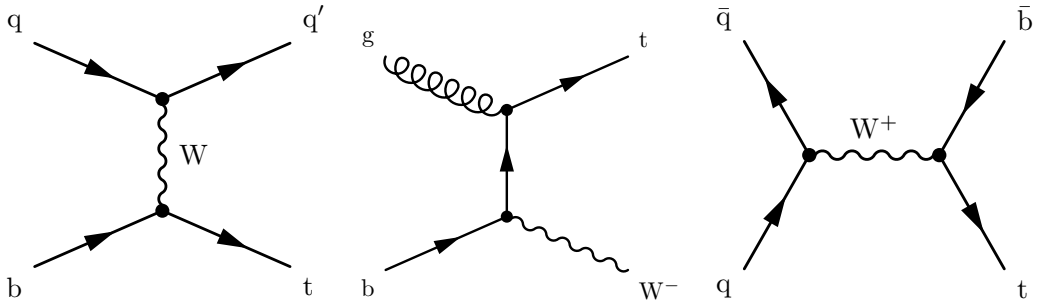

Figure 1.4: Single top quark production. Single top quarks are produced via electroweak interaction. The dominant production mode at the LHC is the *t*-channel process (left), where the top quark is produced together with a light-flavored quark in forward direction, followed by the production of a single top quark with a W boson (center), called tW process. The rarest production mode is the production of a top quark with a bottom quark via the *s* channel (right).

production could be caused by physics beyond the SM. Such new physics is predicted by models involving the exchange of a non-SM mediator, for example, a charged Higgs boson [40] or a W′ boson [41]. An overview of possible physics beyond the SM scenarios for *s*-channel single top quark production is given in Ref. [42].

1.2.3 Top quark decay

As the CKM matrix element V_{tb} is close to 1, the top quark decays almost exclusively into a bottom quark and a W boson, while the decays into other down-type flavors are suppressed. The subsequent decay of the W boson, which decays either hadronically into a quark-antiquark pair or leptonically into a charged lepton and the corresponding neutrino, characterizes the top quark decay. For the decay of a top quark pair, three different decay modes are possible, with the frequencies calculated from the W boson decay branching ratios. In case of the fully-hadronic final state (45.7% [14]), both W bosons decay into a quark-antiquark pair, whereas in case of the semileptonic (43.8% [14]) and dileptonic final states (10.5% [14]), one or both W bosons decay into a lepton-neutrino pair.

2. The Compact Muon Solenoid experiment at the Large Hadron Collider

High energies are needed to observe and study heavy SM particles like the Higgs boson or the top quark. Such a high amount of energy for producing these particles is provided by a suitable accelerator. Their decay products are measured by a dedicated detector system. In this chapter, the Large Hadron Collider and the Compact Muon Solenoid experiment are briefly described.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [43, 44, 45] is the largest particle accelerator ever built and is located around 100 m beneath the ground at the European Organization for Nuclear Research (CERN, Conseil Européen pour la Recherche Nucléaire) in Geneva, Switzerland. It is a ring containing superconducting magnets and cavities with a circumference of 27 km and is the last part of the CERN accelerator complex, illustrated in Fig. 2.1. The LHC has two different operation modes, accelerating either protons or heavy nuclei.

After being preaccelerated, two beams, each consisting of up to 2808 bunches with around 10^{11} protons per bunch, are injected into the LHC accelerator tunnel, which comprises eight arcs each containing 154 dipole magnets. These superconducting magnets are responsible for bending the beams and are cooled down by liquid helium to a temperature of 1.9 K, providing a maximum magnetic field of 8.33 T. The beams are focused by 392 quadrupole magnets in total. Between the arcs, eight straight sections are installed, which serve as insertions for experiments, beam injection, or acceleration utilities. Two cavity systems, one for each beam direction, are installed in the straight section at Point 4 of the LHC and are used to accelerate the protons until they reach high collision energies in the range of several TeV. The two proton beams are accelerated in opposite directions in two

different beam pipes situated next to each other and are brought to collision at four points in the LHC tunnel.

At the collision points, the four main experiments of the LHC are installed. The ATLAS and the Compact Muon Solenoid (CMS) experiments are general-purpose detectors designed for the search of the Higgs boson and for physics beyond the SM, as well as for precision measurements of SM processes. A Large Ion Collider Experiment (ALICE) is specialized in measuring heavy-ion collisions and studies the quark-gluon plasma state at high energy densities. The Large Hadron Collider beauty (LHCb) experiment is a forward detector studying the properties of hadrons that contain bottom quarks.

The first proton beam was successfully injected into the LHC on 10th September 2008. Due to an incident at the LHC around a week later, which resulted in several damaged magnets [47], the first data taking was delayed and started in 2010. In Run 1 of the LHC, data was recorded at two different center-of-mass energies: $\sqrt{s} = 7\,\text{TeV}$ (2010-2011) and 8 TeV (2012), corresponding to beam energies of 3.5 and 4 TeV, respectively. From 2013 to 2014, the LHC was shut down for upgrades of the accelerator and detectors to prepare for data taking at an increased center-of-mass energy of 13 TeV. Run 2 of the LHC started in 2015 and data was taken until the end of 2018. After another long shutdown, Run 3 will start in May 2021 and data will be taken until the end of 2024 [48]. It has not been decided yet, at which center-of-mass energy the collisions will take place, i. e., the LHC will be either operated at 13 TeV or at a slightly increased energy up to its design energy of 14 TeV.

The performance of an accelerator is measured by the instantaneous luminosity L. For collider rings, L is defined as

$$L = \frac{n_\text{b} N_1 N_2}{4\pi\sigma_x\sigma_y} \cdot f, \tag{2.1}$$

with the number of bunches n_b, the number of protons per bunch N_i in beam 1 and 2, the widths of Gaussian cross-section profiles σ_x and σ_y of the beams, and the beam revolution frequency f. The design luminosity of the LHC is $10^{34}\,\text{cm}^2\text{s}^{-1}$ and was first achieved in June 2016 [49]. The number of interactions for a specific process is calculated with L and the total cross section σ of the process:

$$N = \sigma \cdot \int L\,\mathrm{d}t = \sigma \cdot L_\text{int}. \tag{2.2}$$

The integrated luminosity L_int is a measure of the amount of collected data. During Run 1 and 2, a total integrated luminosity of $192.3\,\text{fb}^{-1}$ was delivered by the LHC [50].

2.2 The Compact Muon Solenoid experiment

The Compact Muon Solenoid (CMS) detector [51] is a multi-purpose particle detector and is located at Point 5 of the LHC near Cessy, France. The CMS detector has a length of 21 m, a

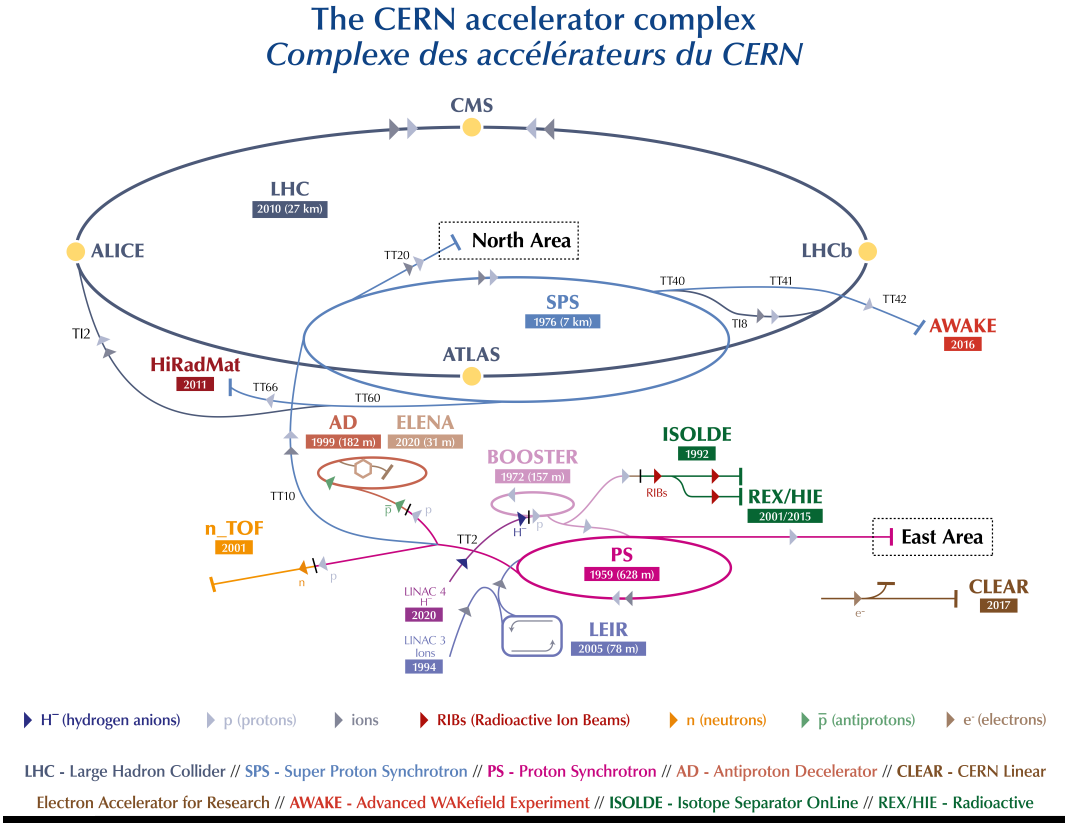

Figure 2.1: The CERN accelerator complex. Shown is the LHC with its four experiments ALICE, ATLAS, CMS, and LHCb, as well as other smaller experiments located at CERN. The acceleration of the protons starts at LINAC 2 and ends with the injection from the SPS into the LHC. Taken from Ref. [46].

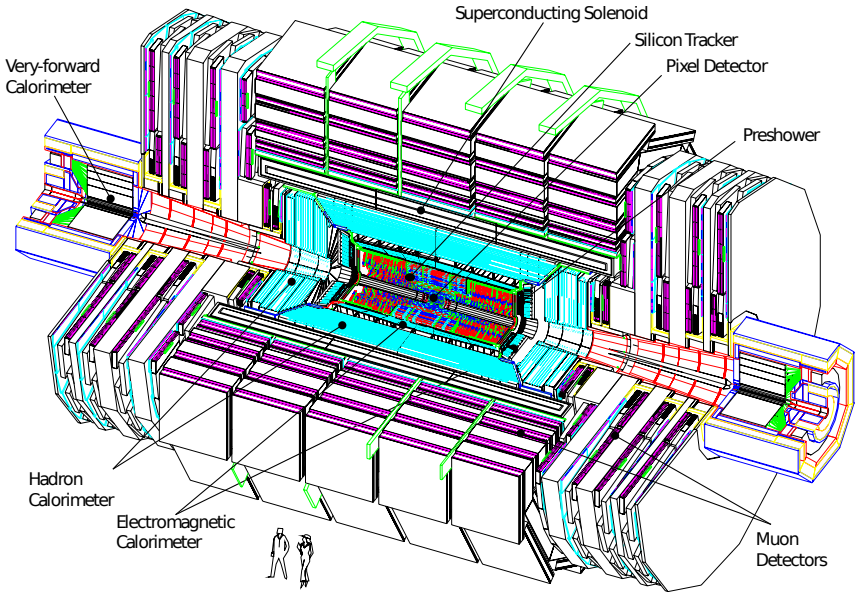

Figure 2.2: Overview of the CMS detector. The innermost part of the CMS detector consists of the pixel detector and the silicon tracker, followed by the preshower and the electromagnetic and hadron calorimeters. These components are surrounded by the superconducting solenoid. Different muon detectors form the muon system, which is the outermost part of the CMS detector. In addition, very-forward calorimeters are installed in the endcaps of the detector. Taken from Ref. [51].

diameter of 15 m, and a total weight of 14 000 t, with the steel return yoke alone weighing around 12 500 t [52]. The CMS experiment consists of several subdetectors, which are built around the interaction point of the two colliding beams. The silicon tracker and the electromagnetic and hadron calorimeters are located within a superconducting solenoid, whereas the muon system is located outside, being the outermost layer of the CMS detector. An overview of the CMS experiment is given in Fig. 2.2.

The coordinate system conventionally used for the CMS detector is shown in Fig. 2.3. Two angles are defined in this coordinate system: the azimuth angle ϕ in the x-y plane and the polar angle θ with regard to the beam direction. Another quantity frequently used in particle physics is the rapidity y, defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \tag{2.3}$$

Here, E corresponds to the energy and p_z to the momentum of a particle in z direction. The rapidity is a crucial observable in high-energy physics because differences in the rapidity are invariant with regard to Lorentz boosts along the z axis, resulting in a Lorentz

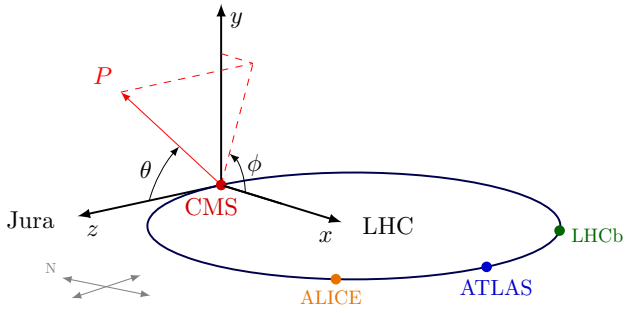

Figure 2.3: Coordinate system of the CMS detector. The x axis points radially inwards towards the center of the LHC, the y axis vertically upwards towards the surface. The z axis is aligned with the beam direction such that a right-handed coordinate system is formed. Taken from Ref. [53].

invariant angular separation of two particles under boosts in beam direction. As the energy and momentum of a relativistic particle are difficult to measure, a related quantity called pseudorapidity is defined, which only depends on the polar angle θ:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.4}$$

For massless particles, the two quantities y and η are identical. The pseudorapidity values range from 0 for particles perpendicular to the beam axis to (minus) infinity for particles (anti)parallel to the beam. The Lorentz-invariant angular separation of two particles is defined using the pseudorapidity and the azimuth angle ϕ:

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}. \tag{2.5}$$

Another important kinematic observable in particle physics is the transverse momentum

$$p_\mathrm{T} = \sqrt{p_x^2 + p_y^2} \tag{2.6}$$

describing the momentum perpendicular to the beam pipes.

In the following, all CMS detector components are described, from the innermost to the outermost layer of the CMS experiment. In addition, an overview of the trigger system and of the LHC computing infrastructure is given.

2.2.1 Silicon tracker

Of all CMS subdetectors, the tracker system [54, 55] is built closest around the beam interaction point. It measures the hits of electrically charged particles, which traverse the

tracker system. By fitting various hits, a track of the charged particle is obtained. As the tracker system is the innermost subdetector, it needs to be robust against radiation damage to guarantee a long life span. At the same time, an accurate measurement of all particle tracks is required to find the location of the hard interaction process. For these reasons, the whole tracker system consists of semiconducting silicon. Charged particles that traverse the silicon detector modules create electron-hole pairs, which are measured eventually as a current.

An overview of the tracker system is shown in Fig. 2.4. The tracker system consists of two parts, the inner silicon pixel detector and the outer silicon strip detector, which ensure a spatial resolution up to 10 µm for the hit position of a particle. The pixel detector [56] installed until 2016 covered a pseudorapidity area up to $|\eta| < 2.5$ and comprised three cylindrical layers forming a barrel and two disk layers on each side of the barrel, resulting in a total number of 1440 pixel detector modules. The barrel layers had radii of 4.3 cm, 7.3 cm, and 10.2 cm, and were 53.3 cm long. They contained 48 million pixels, which are mostly of size 100 µm × 150 µm. The disk layers consisted of 18 million pixels in total and they were located at a distance from the interaction point of ±34.5 cm and ±46.5 cm. Between 2016 and 2017, the pixel detector was replaced in the Phase 1 upgrade [57] to work efficiently at higher instantaneous luminosities than the LHC design value. In this upgrade, the number of barrel layers was increased to four, now having a length of 54.9 cm and diameters of 3.0 cm, 6.8 cm, 10.9 cm, and 16.0 cm. The new pixel detector comprises three instead of two endcaps per side, located ±29.1 cm, ±39.6 cm, and ±51.6 cm away from the collision point. In total, the number of pixels has been increased to 124 million.

The silicon strip detector consists of 15 148 detector modules, which contain 9.3 million strips in total and cover an area of 198 m^2. The strip detector modules are arranged in four subsystems: Tracker Inner Barrel (TIB) and Disk (TID), Tracker Outer Barrel (TOB), and Tracker EndCap (TEC). The TIB consists of four cylindrical layers enclosed by two TIDs, which each contain three wheels, and the TOB comprises six cylindrical layers. At each end of the silicon strip detector, a TEC is installed, which contains nine wheels. In the TIB, TID, and the four inner wheels of the TEC, silicon strips of thickness 320 µm are installed, whereas the outer TEC wheels and the TOB consist of 500 µm thick silicon strips. In total, the silicon strip detector has a length of 5.5 m and a diameter of 2.4 m [58], and covers a pseudorapidity area of up to $|\eta| < 2.5$.

2.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [59, 60] is built around the CMS tracker system. It is designed to measure the energy of electromagnetically interacting particles, i. e., electrons, positrons, and photons. If they pass the detector material, they create a cascade of electrons, positrons, and photons via bremsstrahlung, Compton scattering, and production of electron-positron pairs, which is referred to as electromagnetic shower. This cascade ends at the critical energy for electron-positron pair production. The particles are then absorbed by the detector material, which emits the absorbed energy in form of scintillation light. The energy deposit in the detector material is measured by transforming the scintillation light into an electrical signal with photodiodes.

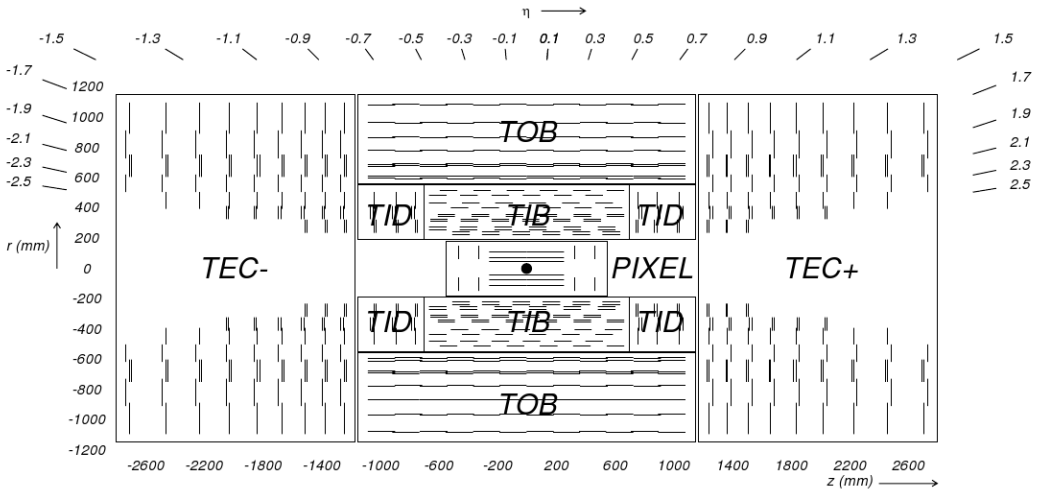

Figure 2.4: The CMS tracker system until 2016. The inner part of the CMS tracker system consists of three cylindrical pixel detector layers and two disks (PIXEL), while the silicon strip tracker forms the outer part, comprising three different subdetectors. Four layers and three wheels form the Tracker Inner Barrel (TIB) and Tracker Inner Disk (TID), respectively. The Tracker Outer Barrel (TOB) consists of six layers. The two Tracker EndCaps (TECs) each comprise nine wheels. Taken from Ref. [51].

As detector material, lead tungstate ($PbWO_4$) crystals are used, acting as absorber and scintillator at the same time. They are highly transparent and are fast scintillators with a small Molière radius of 2.19 cm, which describes the transverse dimension of an electromagnetic shower. They enable a precise energy measurement. Because of their radiation length X_0 of 0.89 cm, defined as the mean distance after which an electron has $1/e$ of its initial energy left, and their high density of $8.28\,\mathrm{g/cm^3}$, the ECAL can be compactly built. An overview of the ECAL detector is given in Fig. 2.5. The ECAL Barrel (EB) covers a pseudorapidity area of up to $|\eta| < 1.48$ and consists of 61 200 $PbWO_4$ crystals, which are arranged in 36 supermodules, each containing 1700 crystals and weighing 3 t. The crystals in the EB have a size of $2.2\,\mathrm{cm} \times 2.2\,\mathrm{cm} \times 23\,\mathrm{cm}$, with the length corresponding to $25.8\,X_0$. Each ECAL Endcap (EE) is divided into two halves, called Dees, and covers a pseudorapidity area of $1.48 < |\eta| < 3.0$. The two EEs contain 14 648 $PbWO_4$ crystals in total, which are of size $2.9\,\mathrm{cm} \times 2.9\,\mathrm{cm} \times 22\,\mathrm{cm}$ and have a length of $24.7\,X_0$. In front of the EE, the Preshower detector (ES) is installed. The ES detector is a sampling calorimeter comprising one lead and one silicon strip layer and is introduced to distinguish single high-energy photons from pairs of low-energy photons stemming from neutral pion decays. With this setup, a precise energy measurement is guaranteed for a pseudorapidity area up to $|\eta| < 2.6$. The overall energy resolution of the CMS ECAL detector is given by

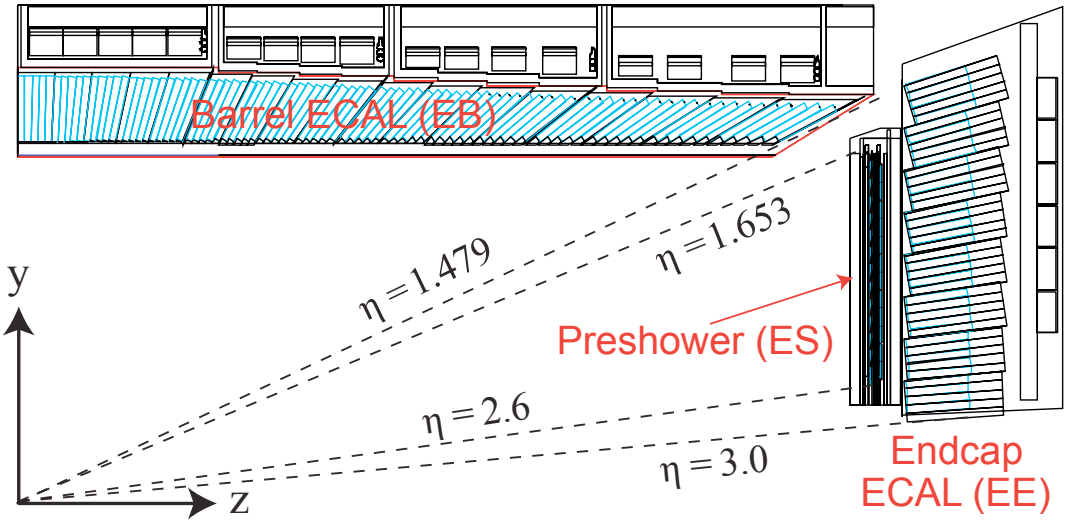

Figure 2.5: The CMS electromagnetic calorimeter. The ECAL Barrel (EB) comprises 61 200 lead tungstate crystals. Each ECAL Endcap (EE) is divided in two halves, called Dees, with each Dee holding 3662 crystals. The Preshower detector (ES) is a sampling calorimeter consisting of two layers, one layer of lead radiators and one layer of silicon strip sensors. Taken from Ref. [62].

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2. \tag{2.7}$$

The stochastic term S describes fluctuations of the electromagnetic shower and photon statistics, and the noise term N takes uncertainties in the energy measurement due to electronics noise into account. Any constant impact on the resolution, e. g., calibration effects, is summarized in the constant term C. It is crucial to keep the constant term as small as possible, as this term dominates the energy resolution at high energies. The energy resolution of the different terms has been measured with an electron test beam [61] to be

$$\left(\frac{\sigma}{E(\mathrm{GeV})}\right)^2 = \left(\frac{2.8\%}{\sqrt{E(\mathrm{GeV})}}\right)^2 + \left(\frac{12\%}{E(\mathrm{GeV})}\right)^2 + (0.3\%)^2\,. \tag{2.8}$$

2.2.3 Hadron calorimeter

The hadron calorimeter (HCAL) [63] measures in particular the energy of hadrons and neutral particles that have not been detected or absorbed by the inner layers of the CMS experiment yet. It is a sampling calorimeter consisting of alternating layers of absorber material (either brass or steel) and plastic scintillator layers. Incoming hadrons interact with the absorber material through inelastic scattering and produce secondary particles of lower energy that are detected by the scintillation layers. These secondary particles

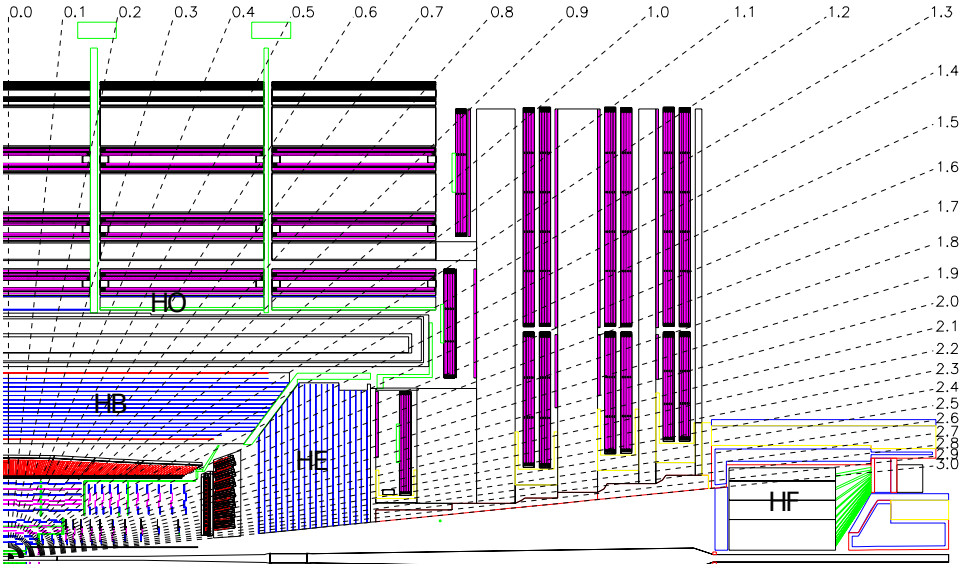

Figure 2.6: The CMS hadron calorimeter. The positions of the Hadron Barrel (HB), Hadron Endcap (HE), Hadron Outer (HO), and Hadron Forward (HF) calorimeters are shown. The dashed lines indicate fixed η values. Taken from Ref. [51].

also interact with the absorber layers, leading to a cascade of particles called hadronic shower. The scintillation light is collected by wavelength shifting fibers [64, 65] and is transformed into an electrical signal using photodiodes. The HCAL is divided into several subdetectors as shown in Fig. 2.6. The HCAL Barrel (HB) consists of 36 wedges and covers a pseudorapidity area of up to $|\eta| < 1.3$. It is restricted by the ECAL at a radial distance of 1.77 m from the interaction point and by the magnet coil at a radial distance of 2.95 m. At each end of the HB, an HCAL Endcap (HE) comprising 36 wedges is installed, covering the pseudorapidity area $1.3 < |\eta| < 3$. As a measurement of the full hadronic shower is not guaranteed in the barrel region due to its geometrical restrictions, a Hadron Outer (HO) calorimeter is installed outside of the solenoid, with the solenoid employed as an additional absorber. Hadron Forward (HF) calorimeters are located at a distance of 11.2 m from the collision point and extend the pseudorapidity coverage up to $|\eta| < 5.2$.

2.2.4 Superconducting solenoid

The momentum of charged particles is determined via the curvature of their tracks when passing a magnetic field. A magnetic field of high strength is required to bend high-energetic particles produced in proton-proton collisions and traversing the CMS tracker. For this purpose, a superconducting niobium-titanium solenoid is installed around the tracker and calorimeter layers of the CMS experiment. The solenoid has a diameter of 6.3 m and a length of 12.5 m and weighs around 220 t. It is designed to provide a magnetic field strength of 4 T, but is operated at 3.8 T to increase its life span [66]. Liquid helium is used to cool down the solenoid to a temperature of 4.45 K. The solenoid is surrounded

by a steel return yoke, which consists of eleven large elements: five barrel wheels and two endcaps, each consisting of three disks. The return yoke has a barrel length of 13 m and a diameter of 14 m. With a weight of 12 500 t, it is the heaviest part of the CMS detector.

2.2.5 Muon system

As muons produced at the LHC typically have energies of the order of GeV, they are minimum ionizing particles, meaning they pass through the inner CMS detector layers with only little interaction. Since all other particles, except neutrinos, which cannot be detected, are mostly absorbed by the inner CMS subdetectors, it is beneficial to reconstruct tracks of muons in the outermost layer of the CMS experiment, where no further distinction from other particle types is required. In Fig. 2.7, the subdetectors of the muon system [67] and their location within the CMS experiment are shown. The muon system is embedded in the return yoke and is divided into a cylindrical barrel section and two planar endcap regions, in which four different types of gaseous detectors are installed. Gaseous detectors are filled with inert gas, which gets ionized by charged particles traversing the detector volume. They contain one or multiple wire anodes, which are surrounded by a cathode cylinder or two cathode plates, respectively. High voltage between anode and cathode leads to drifting electrons and heavy ions. They are accelerated by the electric field and create additional free electrons, resulting in an avalanche of electrons that is measured as a current at the anode wire. In addition to the inert gas, a quench gas, typically CO_2, is used to absorb UV photons, which can be emitted when the ions are recombining and can cause an additional avalanche of electrons. As the magnetic field inside the return yoke has a strength of up to approximately 2 T [66], the embedded muon chambers are able to measure the transverse momentum of traversing muons independently of the tracker system, which is needed for the muon trigger system.

The employed muon chambers are required to be inexpensive, reliable and robust, as they cover a total detection plane area of 28 000 m^2. In the barrel region of the muon system, a low muon rate is expected and the magnetic field is homogeneous. Organized in four stations and covering a pseudorapidity area up to $|\eta| < 1.2$, 250 Drift Tube (DT) chambers are installed in the barrel, with the first three stations measuring the location of muons in the r-ϕ bending and the r-z longitudinal plane, and the fourth station measuring muons only in the r-ϕ bending plane. The endcaps of the muon system, where the magnetic field is non-uniform and higher background rates are expected compared to the barrel region, contain Cathode Strip Chambers (CSCs). With its fast response time, radiation resistance and fine segmentation, this kind of detector is suitable for the conditions in the endcap region. Initially, 468 CSCs distributed in four stations were installed in a pseudorapidity area of $0.9 < |\eta| < 2.4$, with the fourth station covering only a pseudorapidity of $|\eta| < 1.8$. During the first long shutdown of the LHC, the fourth endcap station has been extended by 72 additional CSCs, now covering a pseudorapidity area up to $|\eta| < 2.4$ [68]. Resistive Plate Chambers (RPCs) are installed in the barrel and the endcaps of the muon system to complement the DTs and CSCs. They serve as a muon trigger system, with 480 RPCs located in the barrel and 576 in the endcap region, organized in four stations in both regions. In the initial endcap design, only three RPC stations were installed. The fourth RPC station

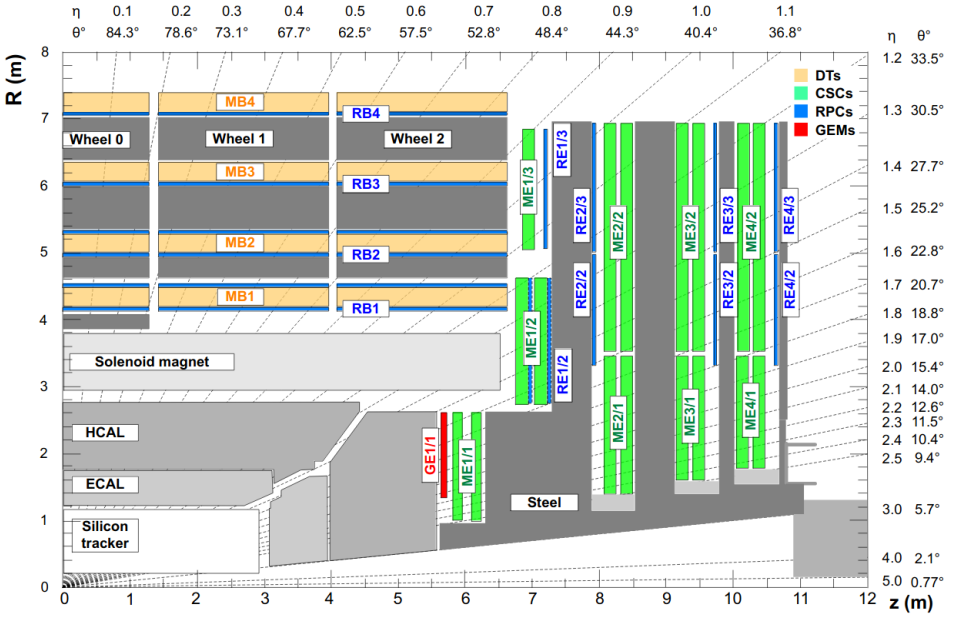

Figure 2.7: The CMS muon system. The outermost part of the CMS detector consists of four different detector types. The Drift Tube (DT) chambers are positioned in the barrel region and are divided into four stations (MB1 – MB4). In the endcap regions, Cathode Strip Chambers (CSCs), which are organized in four stations (ME1 – ME4), are used. A complementary detector type, Resistive Plate Chambers (RPCs), is installed both in the barrel and endcap regions, divided into eight stations in total (RB1 – RB4 and RE1 – RE4). A first station of Gas Electron Multiplier (GEM) chambers is installed (GE1), with an additional GEM station planned to be added during the third long shutdown of the LHC. Furthermore, the location of the inner components of the CMS detector is shown. Taken from Ref. [69].

covering $|\eta| < 1.9$ was installed during the first long shutdown [68]. In early 2017, first Gas Electron Multiplier (GEM) chambers were installed for demonstration in the forward region ($1.6 < |\eta| < 2.2$) of the muon system with the aim to complement the CSCs [69]. During the second long shutdown of the LHC, the installation first GEM station has been finalized [70]. Major modifications of the muon system, including the installation of a second layer of GEMs, are planned to be installed during the third long shutdown of the LHC starting in 2025. These upgrades will enable muon identification up to $|\eta| < 2.8$ [70].

2.2.6 Trigger system

With a data rate of the LHC of 40 MHz, corresponding to a bunch crossing interval of 25 ns, it is impossible to store all measured data of the CMS detector. Potentially interesting events need to be selected using a dedicated trigger system. The CMS trigger system [71, 72, 73, 74] consists of two parts, the hardware-based Level-1 (L1) and the software-based

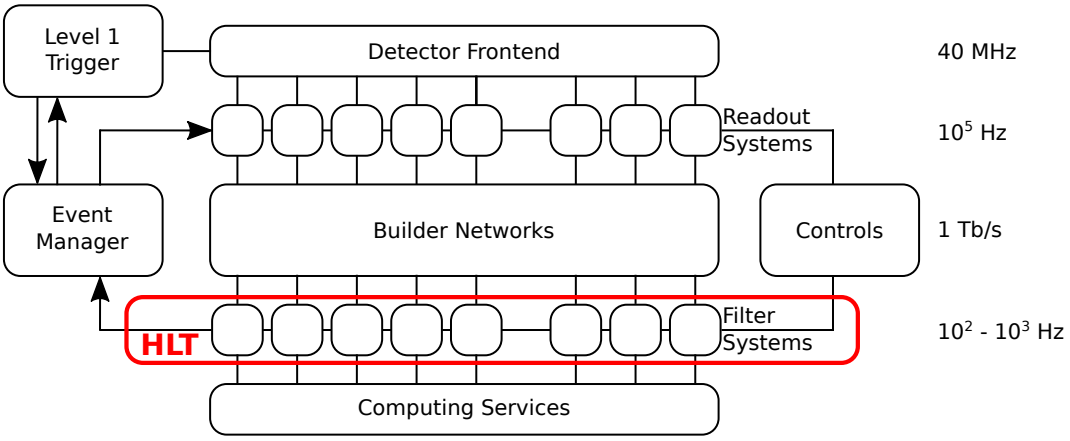

Figure 2.8: Overview of the CMS trigger system. The data rate of the LHC, which is 40 MHz due to the bunch crossing interval of 25 ns, is reduced via the hardware-based Level-1 (L1) and the software-based High-Level Trigger (HLT). The L1 trigger searches for events that contain data of interest based on the information from the muon system and calorimeters and reduces the data rate to 100 kHz. The HLT is performing the final decision of keeping an event using all available detector information, further reducing the data rate to the order of 10^3 Hz. Adapted from [75].

High-Level Trigger (HLT). The general structure of the CMS trigger system is shown in Fig. 2.8. Based on the information from the muon system and calorimeters, the L1 trigger searches for events of interest and reduces the data rate to 100 kHz. The HLT software is operated on standard computer farms and performs a first event reconstruction by using information of all detector components and employing loose requirements. It is therefore responsible for the final decision of storing the event. After applying the HLT, the averaged reduced data rate is of the order of 100 Hz to 1 kHz.

2.2.7 Computing

The data selected by the CMS trigger system is stored, preprocessed and made accessible to physics analyzers. For this purpose, a distributed computing infrastructure at the LHC is employed and commonly used by the four main experiments ATLAS, ALICE, CMS, and LHCb. The Worldwide LHC Computing Grid (WLCG) [76, 77] is hierarchically structured into four layers, the tiers 0, 1, 2, and 3. In Fig 2.9, the structure of the WLCG is depicted. The original, raw data measured by the experiments at the LHC is stored at Tier-0, a data center directly located at CERN. In this data center, a first data processing is performed and the reconstructed data are sent to Tier-1. The Tier-1 currently comprises 13 large computer centers at different locations around the world, with one of them located at Karlsruhe Institute of Technology (KIT). The main task of the Tier-1 is to fully reprocess and store event data. About 160 Tier-2 sites are distributed mainly at scientific institutes and universities. They do not have as much storage capacity as the Tier-1 computing

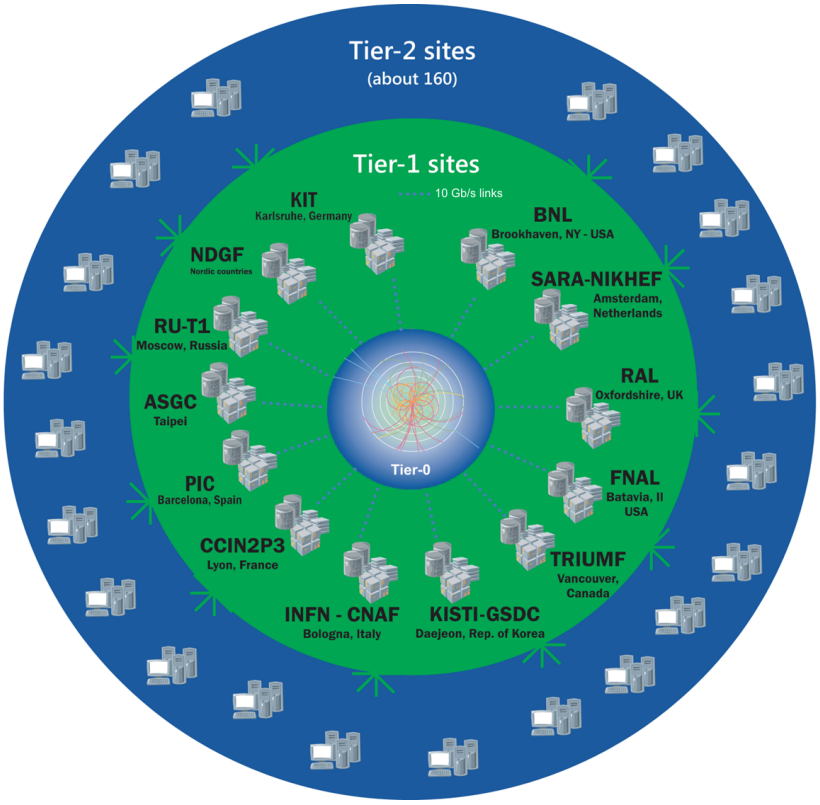

Figure 2.9: The Worldwide LHC Computing Grid (WLCG). The Tier-0 center is located at CERN. One of the 13 large computer centers (Tier-1) is located at Karlsruhe Institute of Technology (KIT). About 160 Tier-2 sites are distributed around the world. Not shown is the Tier-3, which corresponds to local computing resources. Taken from Ref. [78].

centers, but they provide a large fraction of CPU power. The Tier-2 is mostly employed for performing final physics analysis tasks. The final layer, Tier-3, corresponds to local computing resources and cloud services that are accessed individually and are indirectly related to the WLCG.

3. Event simulation and reconstruction

In particle physics, thorough simulation of the outcome of collision events is needed to understand the data composition and to compare measured data with theory predictions. This simulation provides the likelihood for specific physics processes and the probability density functions of their kinematic distributions and comprise multiple steps from the hard interaction of partons to the detector response of the CMS experiment. The same reconstruction chain is applied on the detector response of simulation and data in order to obtain physics objects like leptons from the detector signals. An overview of the structure of a proton-proton collision and tools used for event simulation is given in the first part of this chapter. The second part of this chapter focuses on the event reconstruction, and the final part of this chapter presents a preselection of physics objects used in the analysis later on.

3.1 Event simulation

Data measured by the CMS experiment is usually compared with simulation to verify or falsify theory predictions. For this simulation, a comprehensive understanding of the nature of a proton-proton collision and of the translation of collision products into detector signals are required. The general structure of a proton-proton collision is outlined in Fig. 3.1. Multiple software packages need to be employed to process the different simulation steps and to ensure a precise and accurate full event simulation.

3.1.1 Hard scattering

The initial particles involved in the hard interaction process are the partons of the protons. These partons carry a certain proton momentum fraction x and can be either the valence quarks of the proton, gluons, or sea quarks. The parton distribution function (PDF) $f(x, \mu_F^2)$ expresses the probability to observe a specific parton with proton momentum fraction x

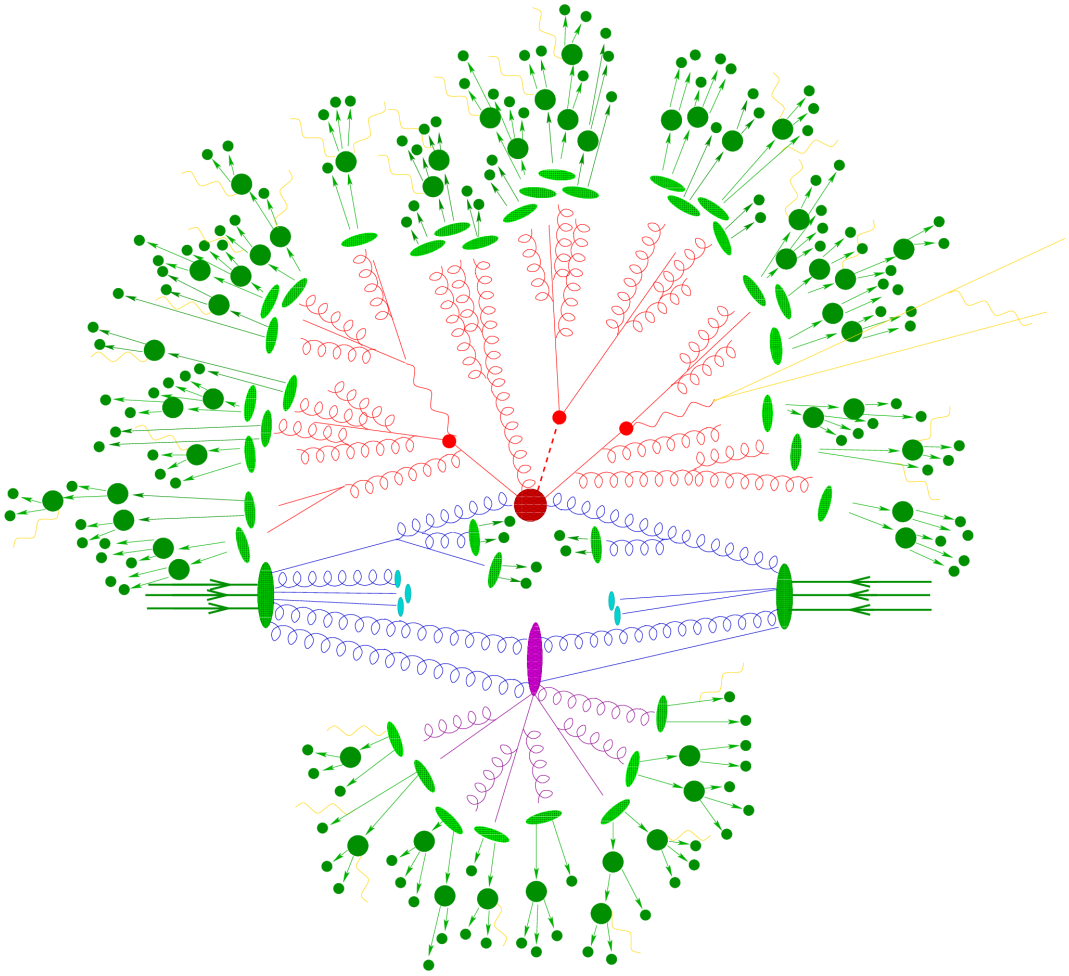

Figure 3.1: Sketch of a proton-proton collision. The red circle in the center of the sketch indicates the hard proton-proton scattering process. The hard interaction is followed by parton showers, initiated by gluon radiation and bremsstrahlung (red). The light and dark green structures show the hadronization process and the decay of hadrons. In addition, a second scattering process, called underlying event, is shown (purple) and the partons of the proton not involved in the scattering process are indicated as light blue ellipses. Taken from Ref. [79].

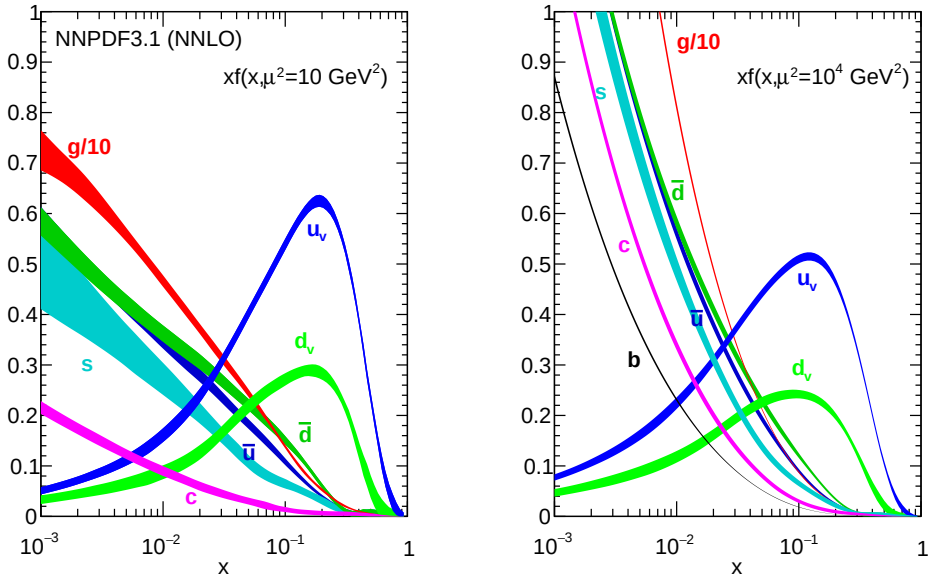

Figure 3.2: The Neural Network Parton Distribution Function (NNPDF) set. Shown are the NNPDF3.1 sets at next-to-next-to-leading order (NNLO) for valence quarks, sea quarks, and gluons for the momentum transfers 10 GeV^2 (left) and 10^4 GeV^2 (right). The gluon distribution, scaled by a factor of 1/10, is dominant for low proton momentum fraction values x. Taken from Ref. [84].

and is evaluated at a given energy scale, the factorization scale μ_F. As no perturbation calculation is available for the evolution of PDFs, one needs to apply the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [80, 81, 82]. Using the DGLAP equations, the PDF measured at a given scale can be evaluated at another scale. PDFs are derived from several measurements, e. g., deep-inelastic scattering processes, and are compiled by several collaborations. In this thesis, the NNPDF set [83, 84] is used, shown in Fig. 3.2 for two different scales. Processes involving a bottom quark in the initial state are special, as the bottom quark mass is higher than the minimum energy scale used for the evaluation of most PDFs. Such processes can be either described with the four-flavor scheme (4FS) or the five-flavor scheme (5FS). In the 4FS, it is assumed that bottom quarks cannot arise as sea quarks in the proton. Hence, they first need to be created in pairs through gluon splitting, with the second bottom quark being part of the final-state partons. In the 5FS, bottom quarks are considered to be massless, thus they can be constituents of the proton.

The total cross section for producing a particle X in the hard interaction of a proton-proton collision is calculated with the QCD factorization theorem [85, 86, 87] and is given by

$$\sigma\left(\mathrm{pp} \to \mathrm{X}\right) = \sum_{j,k} \iint \mathrm{d}x_1\, \mathrm{d}x_2 f_j(x_1, \mu_\mathrm{F}^2) f_k(x_2, \mu_\mathrm{F}^2) \hat{\sigma}_{jk\to\mathrm{X}}\left(x_1, x_2; \mu_\mathrm{F}^2, \mu_\mathrm{R}^2\right). \tag{3.1}$$

Here, $f_{j,k}$ and $x_{1,2}$ are the PDFs and momentum fractions of the initial partons j and k, while $\hat{\sigma}$ is the cross section for the partonic subprocess $jk \rightarrow \mathrm{X}$. The partonic subprocess cross section depends on the renormalization scale μ_R, which describes the scale dependence of the strong coupling constant α_s in perturbative calculations. For the calculation of $\hat{\sigma}$, in principle all existing Feynman diagrams need to be considered to determine the matrix element. As an infinite number of Feynman diagrams can be defined by adding radiation of gluons or by additional internal lines or loops, this is not possible. Instead, the cross section is perturbatively determined in a series expansion in α_s, which has small values for high momentum transfers:

$$\hat{\sigma} = \sigma_\mathrm{Born}\Big(1 + \underbrace{\sigma_1\left(\frac{\alpha_\mathrm{s}}{2\pi}\right)}_{\substack{\text{NLO}\\\text{corrections}}} + \underbrace{\sigma_2\left(\frac{\alpha_\mathrm{s}}{2\pi}\right)^2}_{\substack{\text{NNLO}\\\text{corrections}}} + \ldots\Big). \tag{3.2}$$

The leading-order (LO) cross section σ_Born includes only Feynman diagrams that have the smallest possible number of strong vertices, also referred to as Born-level Feynman diagrams. In the next-to-leading-order (NLO) cross section, the Born contribution and virtual loop corrections (σ_virt), which both lead to a final state of n partons, and corrections due to real emission of an additional parton (σ_real), resulting in a final state of $n + 1$ partons, are considered:

$$\sigma_\mathrm{NLO} = \int_n \mathrm{d}\sigma_\mathrm{Born} + \int_n \mathrm{d}\sigma_\mathrm{virt} + \int_{n+1} \mathrm{d}\sigma_\mathrm{real}. \tag{3.3}$$

By including Feynman diagrams with two additional real emissions, two virtual loops, and one real emission and one virtual loop, the cross section at next-to-next-to-leading order (NNLO) is calculated.

3.1.2 Parton shower

Higher-order real-emission corrections of lower energy compared to the hard scattering are approximated with parton showers. In the parton shower step, additional gluon radiation is simulated by adding gluon emissions either in the initial or in the final state of the Feynman diagrams, referred to as initial-state radiation (ISR) and final-state radiation (FSR). The added gluons can radiate further gluons and create quarks via gluon splitting, resulting into a cascade of partons, called parton shower. At lower energies, matrix element calculation cannot be used for parton shower simulation. Instead, the parton shower evolution is simulated using Sudakov form factors [88, 89] and the Altarelli–Parisi splitting functions [82]. For processes at NLO accuracy, the simulation of additional radiation can lead to double counting. For instance, the Born-level diagram of an NLO process can overlap with an LO process, where the first radiation was added by the parton shower. Matching and merging algorithms like MLM [90] and FxFx [91] are applied to solve this issue.

3.1.3 Hadronization

In the parton shower, free partons are simulated. As color confinement is dominant at lower energies, partons form colorless bound states. This process is called hadronization and is described by phenomenological models, as no perturbation theory is applicable because of low parton energies. One of these hadronization models is the Lund string model [92]. In this model, all quarks and gluons are considered as field lines. These field lines form a narrow tube as they are attracted to each other due to gluon self-interaction. As soon as the energy of the flux tube exceeds a critical value, a new quark-antiquark pair with new tubes is created, preventing the occurrence of color-charged particles. Most of the mesons and baryons produced in the hadronization step have a limited life span, hence they subsequently decay into stable particles.

3.1.4 Underlying event and pileup

Besides the partons taking part in the hard scattering process, the colored remnants of the protons can interact and cause additional hadronization. This is called underlying event and needs to be simulated as well. As bunches containing large numbers of protons are accelerated and brought to collision every 25 ns, multiple proton-proton collisions can overlap, referred to as pileup. The effect of pileup needs to be considered in simulation and is divided into two categories. In case of in-time pileup, multiple collisions occur in the same bunch crossing, whereas out-of-time pileup describes the leakage of collisions from different bunches, caused by the detector response time.

3.1.5 Monte Carlo event generators

The aforementioned simulation steps are computed by Monte Carlo (MC) event generators. Due to the statistical nature of quantum mechanics, a single collision cannot be predicted. Therefore, a large number of collisions needs to be simulated with the MC method [93] to obtain probability density functions for a given process. In this thesis, different MC event generators are used to determine the matrix elements of scattering processes at NLO precision and to compute the subsequent parton shower, the hadronization, the underlying event, and the pileup, although there are software packages available that can perform all steps.

MadGraph5_aMC@NLO

MadGraph5_aMC@NLO [94] combines the features of the two MC event generators MadGraph5 [95] and MC@NLO [96]. It provides simulated events and cross sections at LO and NLO precision and matches the calculated matrix elements with the parton shower simulation. The LO multi-leg ($2 \rightarrow n$) processes, with 2 initial-state and n final-state partons, are computed via MadGraph5, and the NLO diagrams are calculated with MC@NLO. At NLO, possible double counting of radiation processes is prevented by introducing negative weights for a certain fraction of events. The negative weights originate from subtraction terms, which depend on the subsequently used parton shower generator. For a given initial and final state, all possible diagrams are automatically calculated.

POWHEG

Another MC event generator used in this thesis is the positive weight hardest emission generator (POWHEG) [97, 98, 99]. It provides matrix elements of a given process at LO and NLO precision. In the matrix element calculations, the hardest radiation is determined first. As subsequent parton showers are not allowed to simulate harder emissions and are required to simulate p_T-ordered showers, no negative event weights need to be assigned to avoid double counting. This also means that the POWHEG generator is independent of the choice of the parton shower generator due to the absence of subtraction terms. In contrast to MADGRAPH5_AMC@NLO, every single process needs to be implemented by the authors, thus only predefined processes can be generated.

PYTHIA

PYTHIA [100, 101] is an event generator that provides, in addition to matrix element calculations, the computation of parton shower and hadronization as well as underlying event simulation. As the matrix elements are only computed at LO precision, PYTHIA is often used for the parton shower and the subsequent simulation steps only. It is interfaced with NLO MC event generators to obtain a full event simulation. The generated emissions of the parton shower are p_T-ordered and the Lund string model is applied for the hadronization step.

3.1.6 Detector simulation

Before the generated events can be compared with measured data, the detector response needs to be simulated. The software package GEANT4 [102, 103, 104] provides a full simulation of the interaction of particles with different subdetectors of the CMS detector. In addition, the electronic signals of the sensors inside the detector are simulated, such that a comparison with experimental data is possible.

3.2 Event reconstruction

The measured or simulated signals of each CMS subdetector need to be interpreted correctly and combined to reconstruct different physics objects, such as muons and electrons. In the following, the algorithms and programs used for the reconstruction of all relevant physics objects are explained.

3.2.1 Particle Flow algorithm

The Particle Flow (PF) algorithm [105] is designed to reconstruct most of the detectable particles produced within a proton-proton collision, i. e., electrons, muons, photons and hadrons, by an optimal combination of the CMS subdetector information. This approach requires fine granularity of the different detector components and efficient tracking to distinguish individual particles. In a first step, the PF algorithm creates *PF elements*, which comprise charged-particle tracks and calorimeter clusters. These elements are combined

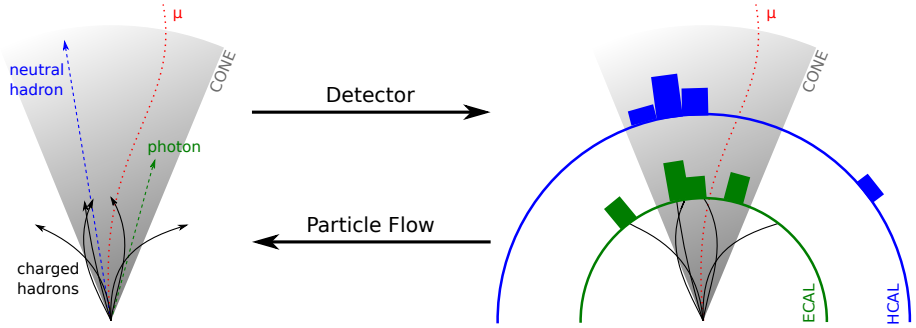

Figure 3.3: The Particle Flow algorithm. The algorithm combines the information collected by the ECAL, HCAL, and the silicon tracker and reconstructs physics objects like neutral and charged hadrons or photons. Adapted from [106].

to *PF blocks*, using a link algorithm that finds the matching track for an energy deposit. In a last step, an identification and reconstruction sequence is applied on these blocks to interpret them as physics objects. This sequence starts with the identification of muons, as they leave the simplest and most characteristic signature in the CMS detector, and ends with the reconstruction of hadrons. The working principle of the PF algorithm is sketched in Fig. 3.3. Subsequent algorithms are applied to build jets and to determine the missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ using all the individual particles reconstructed by the PF algorithm.

3.2.2 Tracks and vertices

Trajectories of charged particles in the silicon pixel and strip detectors, called tracks, are reconstructed by combining hits in different layers of the tracking detectors. Hits are signals induced by charged particles traversing the detector material. Tracks are reconstructed from hits using the Combinatorial Track Finder (CTF) tracking software [107], which is based on Kalman filters [108, 109, 110]. An iterative approach is used, where tracks with transverse momenta of $p_{\mathrm{T}} > 0.8\,\mathrm{GeV}$ and nearby the interaction region are reconstructed first and their associated hits are then removed before the next track is fitted. Each iteration comprises four steps: First, initial track candidates with only two or three hits are found. Kalman filters are applied on these trajectories as the second step to find additional hits that can be associated with the tracks. In the third step, a dedicated track-fitting module based on Kalman filters is employed to find the best possible track for the given hits. Tracks that fail specific quality criteria are then discarded in the final step.

The reconstructed tracks are used to find the vertices of a proton-proton collision, in particular the primary vertex, which is the location of the hard interaction process. Further vertices are caused by pileup and by subsequent particle decays, hence a reconstruction algorithm is required that not only finds all vertices, but also determines the primary vertex of a collision. In principle, the vertex reconstruction consists of three steps: selection of high-quality tracks, clustering of tracks that are candidates for originating from the same

vertices, and fitting of the vertex position using the associated tracks. The deterministic annealing (DA) algorithm [111] is used to cluster tracks according to their z coordinates relative to the beam spot center. This algorithm ensures that as many vertices as possible are resolved, without accidentally splitting a vertex into more than one cluster of tracks. The position of a vertex is determined by applying the adaptive vertex fitter (AVF) algorithm [112], which finds the best possible vertex position using vertex candidates with at least two associated tracks. The primary vertex of a collision is identified to be the vertex with the highest p_{T}^2 sum of the physics objects, which are assigned to tracks associated with the vertex.

3.2.3 Muons

Tracks of muons are reconstructed using the information of two subdetectors of the CMS experiment, the tracker and the muon system. Muon tracks are first reconstructed independently in the inner tracker (tracker track) and in the muon system (standalone-muon track). These two track types can be combined in two different ways using dedicated tracking algorithms [68]. By employing Kalman filters, information of CSCs, DTs, and RPCs is used to determine standalone-muon tracks. Tracker muon tracks are reconstructed with the *inside-out* approach, where tracker tracks with transverse momentum $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and total momentum $p > 2.5\,\mathrm{GeV}$ are propagated to the muon system. Extrapolated tracker tracks that can be matched with at least one muon segment, which is built up from reconstructed hits in the CSCs or DTs, are then considered as tracker muon tracks. A third type of muon tracks, called global muon tracks, are built *outside-in* by connecting standalone-muon tracks with tracker tracks using a Kalman filter. About 99% of all reconstructible muons are either reconstructed as tracker tracks, global tracks, or both. Tracker muon reconstruction is highly efficient for muons with low p_{T}, whereas global muon reconstruction provides high efficiency for muons with $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, which penetrate multiple stations of the muon system. Due to worse momentum resolution and higher cosmic muon background, standalone muon reconstruction is least efficient.

3.2.4 Electrons

As electrons have a significantly lower mass than muons, the most challenging aspect of electron reconstruction [113, 114] is bremsstrahlung caused in the tracker material. Due to bremsstrahlung, electrons do not traverse as many subdetectors as muons and they leave a shower with a large spread in azimuthal direction in the ECAL. Thus, elaborate techniques and the combination of information from the tracker and ECAL are needed for the difficult electron reconstruction. Clustering algorithms collect bremsstrahlung photons and construct *superclusters* using the energy deposits of ECAL crystals (clusters) within a narrow region of the ECAL. These superclusters consist of multiple clusters and are the starting point for extrapolating electron tracks towards the collision point. Electron tracks are reconstructed using a Gaussian Sum Filter (GSF) [115, 116], which employs a dedicated energy loss modeling. With the algorithms described above, electrons with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ can be reconstructed. In case of low-energetic electrons with $p_{\mathrm{T}} < 5\,\mathrm{GeV}$, tracks are first

reconstructed with Kalman filters and a GSF and are then matched to the energy deposits of the ECAL.

3.2.5 Photons and hadrons

After identifying muons and electrons, the remaining PF blocks are assigned to hadrons and photons. The reconstruction of isolated photons is performed in a similar way as for electrons, but without assigning the ECAL superclusters to a track [117]. Non-isolated photons, charged hadrons, and neutral hadrons originate from jet hadronization processes. Within the tracker acceptance ($|\eta| < 2.5$), remaining ECAL and HCAL clusters that are not linked to any track are reconstructed as non-isolated photons and neutral hadrons, respectively. As roughly 25% of the hadronic jet energy is carried by photons and as only about 3% of the energy deposit in the ECAL is caused by neutral hadrons, ECAL clusters with no tracks are always associated to photons. Remaining ECAL and HCAL clusters that are connected to a track are classified as charged hadrons. Outside of the tracker acceptance, no track information is available to distinguish charged and neutral hadrons. As hadrons deposit approximately 25% of the jet energy in the ECAL in the region $|\eta| > 2.5$, the choice of reconstructing all ECAL clusters as photons is no longer justified. Instead, only ECAL clusters without a link to HCAL clusters are associated with photons, whereas connected ECAL and HCAL clusters are classified as charged or neutral hadrons originating from the same hadron shower.

3.2.6 Jets

After all particles have been reconstructed with the PF algorithm, it needs to be determined if they are constituents of a particle shower caused by hadronization of partons. This abstract physics object is called a jet, defined as a cascade of hadrons and their decay products within a narrow cone. As some types of hadrons decay via weak and electromagnetic interaction, jets can contain non-isolated leptons and photons.

Different algorithms are available to reconstruct a jet by clustering particles. Cone-based algorithms, which aim to cluster all particles within a cone of specified size, and sequential recombination algorithms, which subsequently combine particles to a jet, are the two main jet algorithm types. All jet algorithms should be collinear and infrared safe, meaning that they are insensitive to the collinear splitting of a hadron and to the soft emission of a gluon. The most commonly used jet clustering algorithm in high-energy physics are sequential recombination algorithms. They are collinear- and infrared-safe and are based on two distance measures:

$$d_{ij} = \min\left(k_{\mathrm{T},i}^{2n}, k_{\mathrm{T},j}^{2n}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{3.4}$$

$$d_{i\mathrm{B}} = k_{\mathrm{T},i}^{2n}. \tag{3.5}$$

The distance d_{ij} between two particles i and j and the distance $d_{i\mathrm{B}}$ between a particle and the beam both depend on the transverse particle momentum k_{T}. The difference between

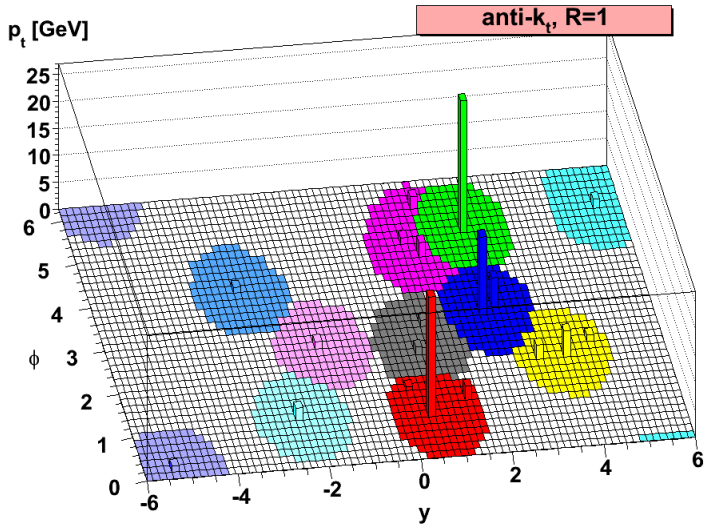

Figure 3.4: The anti-k_T jet clustering algorithm. The transverse momenta of partons of a simulated event in the y-ϕ plane are shown. The anti-k_T algorithm reconstructs circular jets, each with a radius of $R = 1$. Taken from Ref. [118].

two particles in the y-ϕ plane is defined as $\Delta_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ and R is the fixed radius parameter. Depending on the parameter n, a different type of sequential clustering algorithm is used. In case of the anti-k_T algorithm [118], the value is set to -1. A value of $n = 0$ refers to the Cambridge/Aachen algorithm [119, 120], whereas n is set to 1 for the usage of the k_T algorithm [121, 122]. The most common algorithm used for analyses at the CMS experiment is the anti-k_T algorithm, with the radius parameter set to $R = 0.4$. In sequential recombination algorithms, particles are iteratively clustered into jets. First, all distances d_{ij} and d_{iB} are calculated and the particle pair with the smallest distance d_{ij} is grouped into a new object. In the next iteration step, all distances are recalculated and further objects are combined. This procedure is repeated until all particles are clustered into jets. As soon as the smallest distance corresponds to d_{iB}, the object is removed from the set of particles and is considered as a jet. An example of jet reconstruction with the anti-k_T algorithm is shown in Fig. 3.4. As many particles are produced within a proton-proton collision, the calculation of all particle distances is very computing intensive. Using the nearest neighbor approach [123] of the FastJet software package [124], the number of iteration steps for jet clustering is significantly reduced.

The energy of a reconstructed jet needs to be corrected both in simulation and data to account for the finite detector resolution, inhomogeneous detector response and effects caused by pileup interactions. A factorized approach is employed for applying jet energy corrections (JEC) [125, 126, 127] in a fixed order, with each of these corrections taking care of a different effect. The first correction applied on data and simulation is the level 1 (L1) pileup correction, which removes energy stemming from pileup. By comparing the

simulation of QCD dijet events processed with and without pileup overlay, the pileup contribution is estimated. The pileup corrections are determined as functions of the energy density, area, transverse momentum, and pseudorapidity of the jet. Residual differences between data and simulation are calculated in dependence of jet η with the random cone method applied on events that do not originate from any hard interaction processes. In this method, many jets are reconstructed in each event by clustering particles in randomly placed cones. By measuring the average transverse momentum of these jets, the average energy offset caused by pileup is determined. The L2L3 MC truth correction takes care of effects due to the simulated jet response. The response corrections are derived by comparing the reconstructed jet p_T with the p_T at particle level and depend on jet p_T and η. Last remaining jet response differences between data and simulation are corrected with the L2L3 residual corrections, which are only applied on measured data. The L2 residual corrections are η-dependent corrections estimated from dijet events, where both jets have similar p_T, but one of them is required to originate from the barrel region. The absolute jet energy scale is corrected by applying the L3 residuals, which are determined as a function of jet p_T from Z+jets, γ+jets and multijets processes. The L5 flavor correction accounts for differences in the response for different jet flavors and is applied on simulation. This correction level is not included in this thesis.

3.2.7 b jets

With the exception of top quarks, all partons of the SM created in a proton-proton collision hadronize and form jets. In top quark analyses, it is essential to reliably identify jets that originate from a bottom quark, as the top quark decays in almost every case to a bottom quark and a W boson. The CMS Collaboration has developed various algorithms for identifying these b jets, which are commonly referred to as *b tagging* algorithms [128, 129]. These algorithms make use of a special property, the secondary vertex, which is only present in heavy-flavored jets, but not in light-flavored jets stemming from u, s, and d quarks or gluons. The secondary vertex of b jets is caused by the decay of hadronized bottom quarks, called B mesons. B mesons have a long lifetime of around 1.6 ps [14], as the bottom quark cannot decay to a top quark and as decays to up or charm quarks are suppressed by the small values of the CKM matrix elements V_{ub} and V_{cb}. Therefore, B mesons travel, depending on their momentum, a distance of a few mm to 1 cm before decaying to further particles. This creates a secondary vertex, which is displaced compared to the primary vertex of the hard scattering process. The decay of a B meson and the creation of a secondary vertex is sketched in Fig. 3.5. Similarly, c jets can be identified, which is more challenging because of the shorter lifetime of D mesons of 1 ps or less. In addition to the lifetime, heavy-flavored jets can be identified by the presence of non-isolated *soft leptons* stemming from semileptonic B and D meson decays, and by the *hard fragmentation* of these jets caused by the mass of b and c quarks. This means that decay products of heavy-flavor hadrons tend to have higher p_T relative to the jet axis compared to the ones in light-flavored jets.

In this thesis, the latest b tagging algorithm, DeepJet [130, 131], is used to identify b jets. The DeepJet algorithm is based on a deep neural network (see also Section 4.1), which

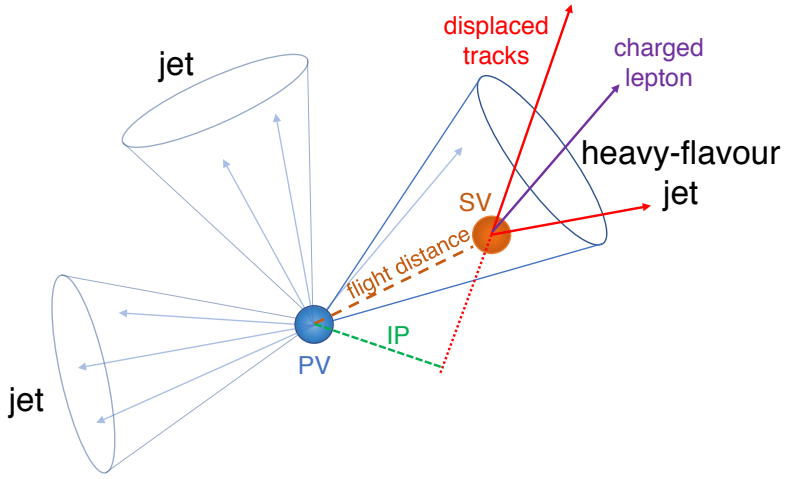

Figure 3.5: Identification of b jets. At the primary vertex (PV), a B meson is produced. Due to its long lifetime, it travels a macroscopic distance before decaying to different particles at the secondary vertex (SV). The decay products cause charged-particle tracks that are displaced with regard to the PV and thus have a large impact parameter (IP) value. Taken from Ref. [129].

employs advanced multivariate techniques like convolutional layers [132] and long short-term memory architectures (LSTMs) [133] to distinguish b jets from other jets. For this complex deep neural network structure, 16 properties of up to 25 charged PF jet constituents, six different kinematic variables of 25 neutral PF jet constituents, 17 properties from up to four secondary vertices associated with the jet, and global variables are used as input information. The DeepJet algorithm takes the phase-1 upgrade of the silicon pixel detector into account, where an additional layer of pixel modules was added in early 2017. The output of this algorithm are the probabilities for the identification of jets stemming from different flavors, which are classified into six categories [134]: jets containing exactly one b hadron, at least two b hadrons, one leptonically decaying b hadron, at least one c hadron but no b hadrons, light-quark jets, and gluon-jets. By calculating the sum of the first three categories, the b jet identification probability is determined. Figure 3.6 displays receiver operating characteristic (ROC) curves, which are later explained in Section 4.1, and shows the efficiency of the DeepJet algorithm and of its predecessor, DeepCSV [129], evaluated for jets from $t\bar{t}$ events. In general, three different working points are defined for b tagging algorithms, based on the probability of misidentifying light-flavored jets stemming from u, s, and d quarks or gluons as b jets: loose (10%), medium (1%), and tight (0,1%). At the medium working point, which is used in this thesis, the b tagging efficiency of the DeepJet algorithm is around 83%, therefore outperforming the DeepCSV algorithm, which only provides an efficiency of roughly 76% with the phase-1 pixel detector.

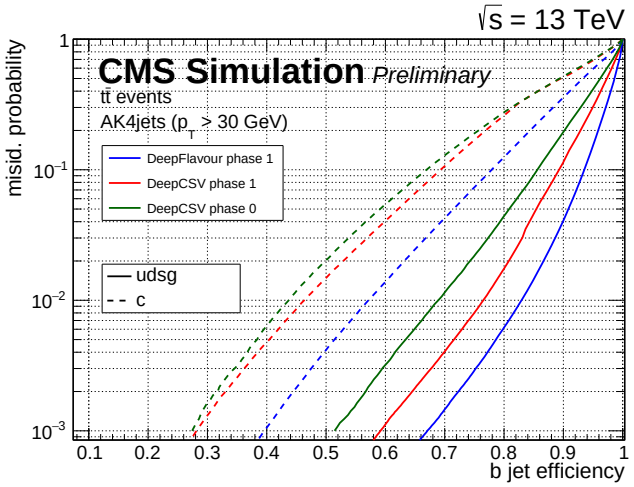

Figure 3.6: b tagging efficiency. The probability for misidentifying light-flavored jets (solid line) and c jets (dashed line) as b jets is shown in dependence of the b tagging efficiency. The performance of the DeepJet algorithm (blue; called DeepFlavour in this figure) is compared for simulated $\mathrm{t\bar{t}}$ events with its predecessor DeepCSV, for the upgraded phase-1 pixel detector (red) and for the phase-0 pixel detector before the upgrade in early 2017 (green). Taken from Ref. [131].

3.2.8 Missing transverse momentum

The initial protons taking part in a proton-proton collision only carry longitudinal momenta. Therefore, the total transverse momentum of all particles produced within the collision is zero as a result of momentum conservation. As neutrinos only interact weakly, they cannot be directly measured with the CMS detector. Instead, their presence leads to an imbalance in the transverse momentum measurement, called missing transverse momentum. In addition to neutrinos, potentially existing particles caused by physics beyond the SM could be indirectly detected as missing energy. The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is defined as the negative vectorial sum of the transverse momenta of all reconstructed PF particles [135]:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = -\sum_{i=1}^{N_{\mathrm{particles}}} \vec{p}_{\mathrm{T},i}. \tag{3.6}$$

Its magnitude is referred to as missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$. As the reconstruction of PF particles is smeared by all the applied jet corrections as described in Section 3.2.6, the missing transverse momentum needs to be calibrated to account for the detector response. This is done by propagating the corrected jet p_{T} to the missing transverse momentum:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss,\,corr}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} - \sum_{j=1}^{N_{\mathrm{PF\,jets}}} \left(\vec{p}_{\mathrm{T},j}^{\,\mathrm{corr}} - \vec{p}_{\mathrm{T},j} \right). \tag{3.7}$$

3.3 Preselection of physics objects

In this section, dedicated initial selection requirements for the reconstructed primary vertices, leptons, jets, and missing transverse momentum are presented. By applying these criteria, which slightly vary in each year because of different data-taking conditions, high-quality physics objects are obtained in all analyzed datasets.

3.3.1 Primary vertices

The primary vertex is required to be within a cylinder of radius 2 cm around the beam axis. Its z coordinate needs to satisfy $|z| < 24\,\mathrm{cm}$, with $z = 0\,\mathrm{cm}$ corresponding to the detector center. The primary vertex needs to have at least five degrees of freedom, meaning that at least four tracks are associated with this vertex.

3.3.2 Muons

Muon candidates are defined according to the muon identification criteria provided by the CMS Collaboration [136]. The aim of these identification requirements, which are summarized in muon identification (ID) flags, is to suppress cosmic muons and muons stemming from kaon and pion decays. If a muon candidate does not fulfill all criteria of a given ID flag, they are rejected and not further considered in the analysis. In this thesis, the tight and loose muon ID criteria are applied on the muon candidates. The loose muon ID provides an efficiency of $\geq$ 99%, whereas the tight muon ID has an efficiency of 95% to 99%, depending on the η range. The selection requirements of these two muon IDs are listed in Table 3.1. In addition, the relative muon isolation $I_{\mathrm{rel},\mu}$ is defined, which describes the amount of deposited energy of other particles within a cone around the muon:

$$I_{\mathrm{rel},\mu} = \frac{1}{p_{\mathrm{T},\mu}} \left[\sum p_{\mathrm{T}}^{\mathrm{ch.\,had.}} + \max\left(0, \sum p_{\mathrm{T}}^{\mathrm{neutr.\,had.}} + \sum p_{\mathrm{T}}^{\gamma} - \Delta\beta \sum p_{\mathrm{T}}^{\mathrm{ch.\,had.,\,PU}}\right)\right]. \tag{3.8}$$

Here, $p_{\mathrm{T},\mu}$ corresponds to the muon transverse momentum, and $p_{\mathrm{T}}^{\mathrm{ch.\,had.}}$, $p_{\mathrm{T}}^{\mathrm{neutr.\,had.}}$, p_{T}^{γ}, and $p_{\mathrm{T}}^{\mathrm{ch.\,had.,\,PU}}$ are the transverse momenta of charged hadrons from the primary vertex, neutral hadrons, photons, and charged hadrons from pileup, respectively. The cone around the muon has a radius of $R = 0.4$ and $\Delta\beta$ is a correction factor that is applied to estimate the contribution of neutral hadrons originating from pileup. This correction is necessary, as it is not possible to determine if a neutral hadron is produced at the primary vertex or at a pileup vertex. The fraction of neutral to charged hadrons from pileup is estimated to be $\Delta\beta = 0.5$ [137]. Muons that stem from semileptonic decays within jets, called non-prompt muons, are rejected by applying a selection requirement on the relative muon isolation.

Table 3.1: Criteria for identifying tight and loose muons. A loose muon is only required to be a PF muon (see Section 3.2.3) reconstructed as either a global or tracker muon. Tight muons need to be reconstructed as global muons and fulfill stringent requirements on the fitted tracks.

Criterion	tight muon ID	loose muon ID
Global muon	yes	–
Global or tracker muon	–	yes
PF muon	yes	yes
χ^2/ndof of global-muon track fit	< 10	–
Number of muon chamber hits	> 0	–
Number of muon segments in muon stations	> 1	–
Transverse impact parameter d_{xy} wrt. primary vertex	$< 2\,\text{mm}$	–
Longitudinal distance d_z wrt. primary vertex	$< 5\,\text{mm}$	–
Number of pixel hits	> 0	–
Number of tracker layers with hits	> 5	–

3.3.3 Electrons

The CMS Collaboration provides recommendations for the identification of electrons [138] to ensure a high quality of the reconstructed electron candidates. Similar to muons, the selection criteria are summarized in electron ID flags, with an identification efficiency of around 70% and 95% for the tight and veto ID, respectively. These two electrons IDs are applied in this thesis and their selection requirements are summarized in Table 3.2. As the background conditions in the ECAL endcap are different as in the barrel and as the tracker coverage is restricted to $|\eta| < 2.5$, different criteria need to be employed in the forward region. Analogously to muons, the energy of other particles deposited within a cone around the electron is expressed by the relative electron isolation $I_{\text{rel,e}}$, which is essential for electron identification and is therefore already part of the electron ID flags:

$$I_{\text{rel,e}} = \frac{1}{p_{\text{T,e}}} \left[\sum p_{\text{T}}^{\text{ch. had.}} + \max\left(0, \sum p_{\text{T}}^{\text{neutr. had.}} + \sum p_{\text{T}}^{\gamma} - \rho A_{\text{eff}}\right)\right]. \tag{3.9}$$

In case of electrons, the cone has a radius of $R = 0.3$ and the pileup contribution is determined by the average transverse momentum density ρ and the effective area A_{eff} of the cone accounting for pileup arising from neutral hadrons. In addition to the ID flag, further requirements on the impact parameters with regard to the primary vertex, d_{xy} and d_z, have to be applied on the electron candidates. These values are determined to be $d_{xy} = 0.05\,\text{cm}$ ($0.10\,\text{cm}$) and $d_z = 0.10\,\text{cm}$ ($0.20\,\text{cm}$) in the ECAL barrel (endcap) region.

3.3.4 Jets

Depending on the year of data taking, different selection criteria for jet identification are recommended by the CMS Collaboration [139]. These criteria depend on the jet pseudorapidity and comprise information about the number of charged and neutral jet constituents

Table 3.2: Identification of tight and veto electrons. The selection criteria for electron identification are different for the ECAL barrel ($|\eta_{SC}| \leq 1.479$) and ECAL endcap ($|\eta_{SC}| > 1.479$) regions, which are defined by the supercluster (SC) pseudorapidity η_{SC}. All momenta and energies given in this table are in units of GeV. Requirements on the impact parameters d_{xy} and d_z need to be applied in addition as they are not part of the electron IDs.

$\lvert\eta_{SC}\rvert \leq 1.479$	tight electron ID	veto electron ID
SC shower shape	< 0.0104	< 0.0126
$\lvert\Delta\eta(\text{SC, track})\rvert$	< 0.00255	< 0.00463
$\lvert\Delta\phi(\text{SC, track})\rvert$	< 0.022	< 0.148
Hadronic energy/EM energy	$< 0.026 + 1.15/E_{SC} + 0.0324\rho/E_{SC}$	$< 0.05 + 1.16/E_{SC} + 0.0324\rho/E_{SC}$
Electron isolation	$< 0.0287 + 0.506/p_T$	$< 0.198 + 0.506/p_T$
$\lvert 1/E_{SC} - 1/p_{track}\rvert$	< 0.159	< 0.209
Expected missing inner hits	≤ 1	≤ 2
Pass conversion veto	yes	yes

$\lvert\eta_{SC}\rvert > 1.479$	tight electron ID	veto electron ID
SC shower shape	< 0.0353	< 0.0457
$\lvert\Delta\eta(\text{SC, track})\rvert$	< 0.00501	< 0.00814
$\lvert\Delta\phi(\text{SC, track})\rvert$	< 0.0236	< 0.19
Hadronic energy/EM energy	$< 0.0188 + 2.06/E_{SC} + 0.183\rho/E_{SC}$	$< 0.05 + 2.54/E_{SC} + 0.183\rho/E_{SC}$
Electron isolation	$< 0.0445 + 0.963/p_T$	$< 0.203 + 0.963/p_T$
$\lvert 1/E_{SC} - 1/p_{track}\rvert$	< 0.0197	< 0.132
Expected missing inner hits	≤ 1	≤ 3
Pass conversion veto	yes	yes

Table 3.3: Identification of jets. Depending on the year of data taking, different selection criteria are applied for jet identification. The requirements depend on the jet pseudorapidity and comprise information from neutral and charged constituents of the reconstructed jet, and are summarized in the tight ID flag.

2016 analysis	$\lvert\eta\rvert \leq 2.4$	$2.4 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	> 1	–	–
Neutral hadron fraction	< 0.90	< 0.90	< 0.98	–
Neutral EM fraction	< 0.90	< 0.90	> 0.01	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	> 0	–	–	–
Charged multiplicity	< 0.99	–	–	–

2017 analysis	$\lvert\eta\rvert \leq 2.4$	$2.4 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	> 1	–	–
Neutral hadron fraction	< 0.90	< 0.90	–	> 0.02
Neutral EM fraction	< 0.90	< 0.90	> 0.02 and < 0.99	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	–	–	–	–
Charged multiplicity	> 0	–	–	–

2018 analysis	$\lvert\eta\rvert \leq 2.6$	$2.6 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	–	–	–
Neutral hadron fraction	< 0.90	< 0.90	–	> 0.02
Neutral EM fraction	< 0.90	< 0.99	> 0.02 and < 0.99	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	–	–	–	–
Charged multiplicity	> 0	> 0	–	–

and about the fraction of electromagnetic and hadronic energy within a jet. By applying the selection requirements on the reconstructed PF jet candidates, misidentified jets, jets with low reconstruction quality, and jets stemming from detector noise are rejected, while obtaining an efficiency for high-quality jets of 98% to 99%. Similar to muons and electrons, the recommended quality requirements are summarized in jet ID flags. In this thesis, the tight jet ID, listed in Table 3.3, is applied for all three years. For b tagging, the medium working point of the DeepJet algorithm, which is described in detail in Section 3.2.7, is chosen to ensure a balance between event selection and b jet identification efficiency. The medium working point is calculated for each year [140] and is given by DeepJet discriminant values of 0.3093, 0.3033, 0.2770 for 2016, 2017, and 2018 data, respectively.

3.3.5 Missing transverse momentum

Several filters are available to reject reconstructed missing transverse momentum that stems from detector noise in the HCAL barrel and endcap, low-quality crystals in the ECAL endcap, as well as beam halo effects and low-quality reconstructed PF objects [141]. These filters are applied on measured data and simulation in all years of data taking. As they are especially designed for searches for beyond the SM physics with high missing transverse momentum in the final state, their impact on this analysis is small.

4. Multivariate and statistical methods

Elaborate tools need to be employed for a successful signal extraction, especially if the signal process is extremely rare compared to the overwhelming background. For this purpose, multivariate methods are applied in the event classification. These methods combine the information of several kinematic variables and are therefore a powerful tool for distinguishing signal from background events. In the first part of this chapter, the multivariate method used in this thesis, a deep neural network, is presented. The second part of this chapter focuses on statistical methods used in this thesis for signal extraction, maximum likelihood estimation and significance calculation, and on the treatment of systematic uncertainties within a fit.

4.1 Deep neural networks

An artificial neural network is a machine learning method, which is inspired by the interconnection of biological neurons and which is used to learn and find common features in a given set of data. In the following, the discussion is restricted to supervised learning. This means that labeled data with known classification, i. e., simulated signal and background events, is used for the learning process. Artificial neural networks are in general organized in multiple layers comprising a different number of neurons. In case of fully connected neural networks, all neurons of a layer are connected with all neurons of the previous and next layer, and each connection is assigned with a weight, which is modified during the learning process, also called training. The first layer of a neural network is referred to as input layer, as each neuron of this layer corresponds to an input variable. Starting from a single neuron, the weighted average of n inputs $x_1, x_2 \ldots, x_n$ of one event is calculated, using the current weight vector $\vec{w}$ and bias b, which is an optional pre-assumption of the data modeling and changes during the training:

$$z = \vec{w}^{\mathrm{T}}\vec{x} + b. \tag{4.1}$$

The result z is then passed to a non-linear activation function f to obtain the predicted value of the neuron, $a = f(z)$. The activation function decides if a feature of a dataset is important enough for the learning process. Different activations are available, for instance the Rectified Linear Unit (ReLU), which applies the function $f_{\mathrm{ReLU}}(v) = \max(0, v)$, or the sigmoid function $f_{\mathrm{sig}}(v) = 1/(1 + \mathrm{e}^{-v})$. The latter one is in particular chosen for the output layer in case of a binary classification problem, where values between 0 and 1 should be returned, whereas the former one is one of the most common activation functions nowadays used for hidden layers, i. e., the layers between the input and output layers. The weights change the steepness of the activation function, and the bias shifts the entire function by an offset.

An artificial neural network usually contains only one hidden layer between the input and output layer, whereas deep neural networks (DNNs) have more than one hidden layer and can be combined with low-level input. In Fig. 4.1, an example architecture for a DNN is shown. For a DNN with multiple hidden layers, each neuron i of layer l calculates a value $z_i^{[l]}$, with the activation $\vec{a}$ of the previous layer, which corresponds to the input $\vec{x}$ for the input layer:

$$z_i^{[l]} = \vec{w}_i^{\mathrm{T}} \vec{a}^{[l-1]} + b_i. \tag{4.2}$$

The activation of a neuron i in layer l is then given by $a_i^{[l]} = f^{[l]}(z_i^{[l]})$. All neurons of layer l can be expressed with the weight matrix $W^{[l]}$ and bias vector $\vec{b}^{[l]}$, defined as

$$W^{[l]} = \begin{pmatrix} \vec{w}_1^{[l]\mathrm{T}} \\ \vec{w}_2^{[l]\mathrm{T}} \\ \vdots \\ \vec{w}_{N_l}^{[l]\mathrm{T}} \end{pmatrix}, \ \vec{b}^{[l]} = \begin{pmatrix} b_1^{[l]} \\ b_2^{[l]} \\ \vdots \\ b_{N_l}^{[l]} \end{pmatrix}, \tag{4.3}$$

with the total number of neurons N_l of layer l.

The results of all neurons of layer l are then calculated with

$$\vec{z} = W^{[l]} \vec{a}^{[l-1]} + b^{[l]}, \tag{4.4}$$

where the activation of layer l is given by $\vec{a}^{[l]} = f^{[l]}(\vec{z}^{[l]})$. The activation of the last layer, the output layer, is the final predicted value y of the DNN.

The learning process is monitored by a loss function, which expresses the difference between the value y predicted by the DNN and the true value $\hat{y}$ from data and should therefore be as small as possible. In general, cross entropy is a suitable loss function, as it is defined as the average amount of information obtained by measuring a random variable. For a classification into signal and background events, the binary cross entropy is employed as loss function. For m events, it is defined as:

$$L(y, \hat{y}) = -\frac{1}{m} \sum_{i=1}^{m} \left[\hat{y}_i \ln y_i + (1 - \hat{y}_i) \ln(1 - y_i) \right]. \tag{4.5}$$

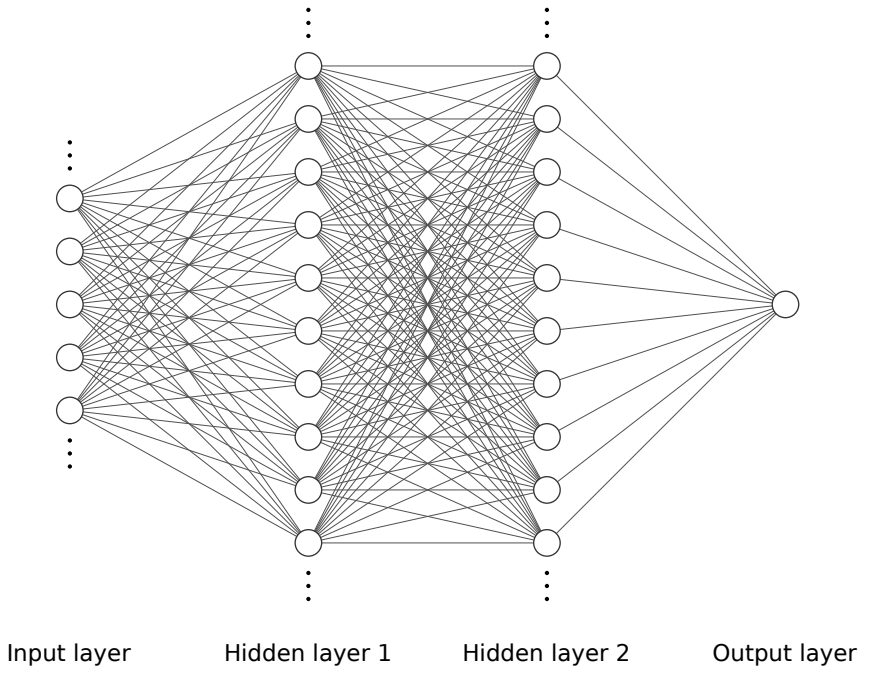

Figure 4.1: Example of a deep neural network. A deep neural network has in general at least two hidden layers, several input nodes and at least one output node. In this example, a fully connected deep neural network is shown, i.e., all nodes of neighboring layers are connected with weights. For illustration purpose, the number of nodes per layer is significantly reduced compared to the number of nodes actually used in deep neural networks, indicated by the dots. Generated with [142].

In the training of a DNN, the loss function L needs to be minimized, which is realized with the gradient descent method implemented in a backpropagation algorithm [143]. During backpropagation, the weights and bias are adjusted to minimize the loss function, starting from the last layer until the first layer:

$$W^{[l]\prime} = W^{[l]} - \alpha \frac{\partial L}{\partial W^{[l]}}, \tag{4.6}$$

$$\vec{b}^{[l]\prime} = \vec{b}^{[l]} - \alpha \frac{\partial L}{\partial \vec{b}^{[l]}}. \tag{4.7}$$

If the value chosen for the learning rate α is too small, the learning process is too slow and local minima are potentially falsely identified as the global minimum, whereas high learning rates lead to the risk that the global minimum cannot be found. In this thesis, the Adam algorithm [144] is used for optimizing the loss function. It determines learning rates individually for different parameters through adaptive moment estimation. This means that estimations of the first and second moment of the gradient are employed to adapt the learning rate for each weight of the DNN.

In this thesis, a DNN is employed for classification of events into signal and background, implemented with the machine-learning tools Keras [145] and TensorFlow [146]. The training of the DNN is performed on a subset of simulated events, which is later removed from the further analysis to avoid biasing the final result of the event classification. If the structure of the DNN is too complex or if the number of events for training is too small, the risk of overtraining is increased. Overtraining means that the DNN is not generalizing features to separate signal from background events and is learning statistical fluctuations of the training dataset instead. This leads to a worse performance of the event classification when applied to a different, independent subset of events. By comparing the DNN output distribution of the training subset with the outcome of an independent subset, potential overtraining is detected. If no overtraining occurred, these two subsets have a similar distribution of the DNN output values, which can be determined by the Kolmogorov–Smirnov (KS) test [147, 148]. For a fixed size of a training dataset, overtraining is in general avoided by reducing the complexity of a neural network, e. g., fewer layers or fewer neurons per layer. In case of a DNN, the dropout method [149, 150] is used to avoid overtraining, allowing to maintain the complex structure. In this method, a defined fraction of random neurons in one or more hidden layers is deactivated during the training to avoid too large influence of single neurons.

The performance of the event classification is measured with the receiver operating characteristic (ROC) curve. It is a probability curve, where the true positive rate is shown against the false positive rate of a machine-learning classifier. The ROC curve is obtained by scanning the DNN output distribution from low to high values, determining the true positive and false positive rates for each value. The true positive rate is the fraction of signal events in the dataset that is correctly classified as signal by the classifier, whereas the false positive rate corresponds to the fraction of background events wrongly classified

as signal events. An event classifier is required to have a large true positive rate and a low false positive rate. The area under the ROC curve (AUC) is taken as measure of the performance of a classifier. If the AUC equals 1, the true positive rate is 1 and the false positive rate is 0, meaning that the classifier perfectly separates signal and background events. An AUC value of 0.5 corresponds to a random decision for classifying an event as signal or background. Therefore, good event classifiers are required to have AUC values as close to 1 as possible.

4.2 Statistical methods

The statistical methods introduced in this section are implemented with the COMBINE software package [151, 152], which employs the RooFit library [153].

4.2.1 Maximum likelihood estimation

For a measurement x, where the functional form of its probability density function f is known, but one or more of its parameters are unknown, the maximum likelihood method is used to estimate these parameter values, such that f describes the distribution of a dataset the best. If a variable x has been measured n times, resulting in values $x_1, x_2, \ldots x_n$, the m parameters $a = \{a_1, a_2, \ldots, a_m\}$ are computed by maximizing the likelihood function

$$L(a) = \prod_{i=1}^{n} f(x_i|a), \tag{4.8}$$

i. e., finding the best parameter set $\hat{a}$, for which $L(\hat{a}) \geq L(a)$ is fulfilled for all possible parameter sets a. The likelihood function is often transformed with the negative natural logarithm, resulting in a function containing sums instead of products, which is less computationally intensive and which is minimized instead of maximized:

$$-\ln L(a) = -\sum_{i=1}^{n} \ln f(x_i|a). \tag{4.9}$$

Therefore, two conditions need to be satisfied by the best parameters $\hat{a}$:

$$\left.\frac{\partial(-\ln L)}{\partial a_j}\right|_{a=\hat{a}} \stackrel{!}{=} 0, \tag{4.10}$$

$$\left.\frac{\partial^2(-\ln L)}{\partial a_i \partial a_j}\right|_{a=\hat{a}} \text{ is negative definite.} \tag{4.11}$$

In this thesis, binned distributions in histograms are obtained from counting experiments, which means that the probability density function follows a Poisson distribution. For a histogram with k bins, the mean value of each bin i is defined as

$$\nu_i = n \int_{x_{i-1}}^{x_i} f(x|a)\,\mathrm{d}x, \tag{4.12}$$

with the number of total events n of the histogram. The likelihood function is then rewritten as

$$L(a) = \prod_{i=1}^{k} \frac{\nu_i^{n_i}}{n_i!} \mathrm{e}^{-\nu_i}. \tag{4.13}$$

Each bin value ν_i consists of a predicted number of signal events s_i and background events b_i, which both depend on nuisance parameters θ that describe systematic uncertainties of the measurement. With the observed number of events per bin n_i, the likelihood function is given by

$$L(\mathrm{data}|\mu,\theta) = \prod_{i=1}^{k} \frac{(\mu s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} \mathrm{e}^{-(\mu s_i(\theta)+b_i(\theta))}, \tag{4.14}$$

where $\mu = \sigma/\sigma_{\mathrm{SM}}$ is a signal strength modifier, usually defined as the ratio of observed to predicted signal production cross section, and is the parameter of interest that is adjusted in the fit. A value of $\mu = 1$ means an observation as predicted by the SM.

4.2.2 Nuisance parameters

Each experimental and theoretical uncertainty that impacts an analysis is considered as a nuisance parameter in the maximum likelihood estimation. There are two types of systematic uncertainty sources, rate and shape uncertainties. A rate uncertainty does not change the shape, but the normalization of a predicted histogram, called template, by shifting all bins of a template by the same value in the same direction. A shape uncertainty alters the shape of a template, either by a correlated shift of the bins, e. g., in case of a shifted energy scale, or by an uncorrelated shift of the bins, referred to as bin-by-bin uncertainties. Rate uncertainties are implemented in the fit by extending the likelihood function with a log-normal distribution:

$$\pi(n) = \frac{1}{n\sigma_n\sqrt{2\pi}} \exp\left[-\frac{(\ln(n) - \ln(\overline{n}))^2}{2\sigma_n^2}\right], \tag{4.15}$$

with the number of events n and the corresponding median $\overline{n}$ and uncertainty σ_n. The advantages of this distribution compared to a Gaussian distribution are that unphysical values $n < 0$ are avoided and that multiplicative uncertainties like scale factors and efficiencies can be described. For each rate uncertainty, the likelihood function is multiplied with a log-normal distribution. In case of shape uncertainties, a more complex approach needs to be used to incorporate them as nuisance parameters, as they affect each bin of a

template differently. This kind of uncertainty is usually provided by two additional templates that correspond to an up and down variation of the nominal template by one standard deviation of the uncertainty. As these three templates represent discrete values, a template morphing method [154, 155] is applied to access templates at values that lie in between or beyond the three values.

4.2.3 Significance

For the search of a new signal process, two different hypotheses are introduced. In the background-only hypothesis, also called null hypothesis, the signal strength modifier μ is set to 0, whereas in the signal-plus-background hypothesis, μ is larger than 0. In the special case that a signal is measured with the same cross section as predicted, μ is set to 1. The signal-plus-background hypothesis is tested against the null hypothesis using a dedicated test statistic q, which decides if the null hypothesis is rejected in favor of the signal-plus-background hypothesis. This is the case if the measured data is very unlikely to be reproduced only by background processes. According to the Neyman–Pearson lemma [156], the ideal test statistic is the likelihood ratio of the two hypotheses. In the COMBINE framework, a slightly modified test statistic [157] is chosen:

$$q_\mu = -2 \ln \left(\frac{L(\text{data}|\mu = 0, \hat{\hat{\theta}}_0)}{L(\text{data}|\mu = \hat{\mu}, \hat{\theta})} \right), \tag{4.16}$$

with $\hat{\hat{\theta}}_0$ maximizing the likelihood function under the assumption of the null hypothesis, and $\hat{\mu}$ and $\hat{\theta}$ maximizing the likelihood function without any condition other than $\hat{\mu} \geq 0$. The measured test statistic q_μ^{obs} is used to determine the p-value:

$$p = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu|\mu, \theta) \, \mathrm{d}q_\mu, \tag{4.17}$$

with the probability density function of the test statistic for a given signal strength μ, $f(q_\mu|\mu, \theta)$, which is computed with MC toy experiments. The p-value expresses the probability to measure a value $q_\mu \geq q_\mu^{\text{obs}}$ for a given μ under the assumption that the signal process exists. In particle physics, the statistical significance of a signal process is quoted as Z score instead of a p-value. The Z score is computed by converting the p-value with the quantile function of the Gaussian distribution Φ^{-1}:

$$Z = \Phi^{-1}(1 - p). \tag{4.18}$$

Therefore, Z corresponds to the distance from the mean in standard deviations if the test statistic q_μ follows a Gaussian distribution. Arbitrary values have been set for the significance for constituting evidence or observation of a new signal process. In case of evidence, a value of $Z \geq 3$ is required, corresponding to a p-value of 0.0013. A discovery is declared if the significance is $Z \geq 5$, i. e., the probability that a data excess originates from

a background-only prediction is less than $2.87 \cdot 10^{-7}$. In general, two different numbers are quoted for the significance of a signal process: the expected significance is computed based on the prediction of the simulation for a given dataset, whereas the observed significance is based on the measured data.

5. Search for s-channel single top quark production at $\sqrt{s}$ = 13 TeV

Of the three main production modes for single top quark production, the t-channel process and the associated production with a W boson has been already precisely measured at the LHC [158, 159], whereas single top quark production in the s channel is the rarest production mode and has not been observed yet in proton-proton collisions.

Single top quark production in general was first observed by the CDF and D0 Collaborations at the Tevatron in 2009 [6, 7], with no distinction made with regard to the production mode. As the Tevatron was a proton-antiproton collider and therefore provided valence antiquarks inside the antiprotons to annihilate with the valence quarks inside the protons, the s-channel process was the dominant production mode. In 2013, first evidence solely for s-channel single top quark production was announced by the D0 Collaboration [160], and the process was observed one year later by a combined search of the CDF and D0 Collaborations [38].

Due to the missing valence antiquark in the colliding protons, the s-channel process is strongly suppressed at the LHC, as it is far less likely to observe sea quarks within the proton momentum fraction range relevant for single top quark production. During Run 1, searches for this particular production mode were performed by the ATLAS and CMS Collaborations. The ATLAS Collaboration announced evidence for s-channel single top quark production with an observed significance of 3.2 standard deviations in 2016, based on the matrix element method [161] and using datasets recorded at center-of-mass energies of $\sqrt{s}$ = 7 and 8 TeV [39]. The CMS Collaboration also performed a search for s-channel single top quark production, which employed boosted decision trees and used 7 and 8 TeV data from Run 1 of the LHC, resulting in an observed significance of 2.5 standard deviations [162].

No analyses focusing on s-channel single top quark production were published yet using Run 2 data at a center-of-mass energy of 13 TeV. Compared to the dominant background

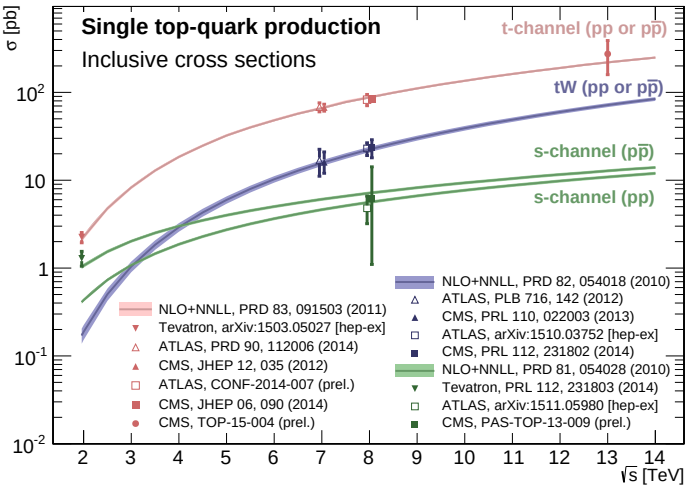

Figure 5.1: Cross sections of single top quark production modes measured at different center-of-mass energies at the Tevatron and the LHC. The measured cross sections for the t-channel process, associated production with a W boson, and s-channel process are shown. In addition, theory predictions for proton-proton and proton-antiproton collisions are shown. Taken from Ref. [163].

process, $t\bar{t}$ production, and to all other single top quark production modes, the predicted cross section of the s-channel process increases less strongly as a function of the center-of-mass energy, as shown in Fig. 5.1. This makes it even more challenging to observe this production mode at higher center-of-mass energies. Nevertheless, as Run 2 has delivered about five times more data compared to Run 1 and as more elaborate analysis tools are available for signal extraction, it is possible to obtain at least evidence.

In this chapter, the search for s-channel single top quark production using the 2016, 2017, and 2018 datasets is described in detail. The general outline of the analysis and the event topology of the signal process and of the relevant background processes are presented in Sections 5.1 and 5.2. Sections 5.3 and 5.4 describe the simulation and selection of events, and the reconstruction of the top quark. Necessary corrections of the simulated events and the data-driven estimation of the QCD multijet background contribution are explained in Sections 5.5 and 5.6. Sections 5.7 and 5.8 are devoted to the event classification using a DNN and to the signal extraction procedure. In Section 5.9, all systematic uncertainty sources considered are discussed. The expected and observed significances for s-channel single top quark production and the measured production cross section are presented in Section 5.10.

5.1 Search strategy

The goal of this analysis is to observe s-channel single top quark production at the LHC with a significance of at least five standard deviations. For this purpose, datasets recorded by the CMS detector in 2016, 2017, and 2018 at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ are analyzed. These datasets correspond to integrated luminosities of $35.9\,\text{fb}^{-1}$, $41.5\,\text{fb}^{-1}$, and $59.7\,\text{fb}^{-1}$, respectively, adding up to a total integrated luminosity of $137.1\,\text{fb}^{-1}$. Only leptonic decay modes of the top quark are considered to suppress the background contribution from QCD multijet processes. Events with one muon or electron in the final state are selected, including tau leptons stemming from W boson decays that further decay into muons or electrons. Depending on the number of jets and b-tagged jets, events are classified into four independent event categories: one signal category and three different control categories that account for the modeling of the most important background processes. The remaining QCD multijet background contribution is estimated with a dedicated data-driven approach, whereas all other processes are modeled using MC simulation. The top quark is reconstructed from its decay products, assuming that the missing transverse momentum is entirely caused by the neutrino from the top quark decay. Due to the overwhelming background, the signal-to-background ratio is only around 1%. Therefore, deep neural networks are employed to classify events as signal- or background-like. In the final step of the analysis, a maximum likelihood fit of templates in the multivariate classifier distribution is performed simultaneously in all years of data taking and in all event categories to compare the background-only with the signal-plus-background hypothesis. If the background-only hypothesis can be rejected, the cross section for s-channel single top quark production is measured by extracting the signal strength from the maximum likelihood fit.

5.2 Event topology

It is crucial to understand the composition of background events that contribute to the event categories chosen in the search for s-channel single top quark production for a successful signal extraction. Several background processes need to be considered in this analysis, as they can mimic the final state of the signal process. In the following, the final-state signatures of the s-channel single top quark process and of its main backgrounds are presented.

5.2.1 Signal process

In s-channel single top quark production, a top quark is produced together with a bottom quark. In Fig. 5.2, the LO Feynman diagram for this production mode is shown, together with the subsequent leptonic decay of the top quark. Therefore, the final state of the signal process comprises a charged lepton and its corresponding neutrino, measured as missing transverse momentum in the detector, as well as two bottom quarks that hadronize within the detector and are identified as b-tagged jets.

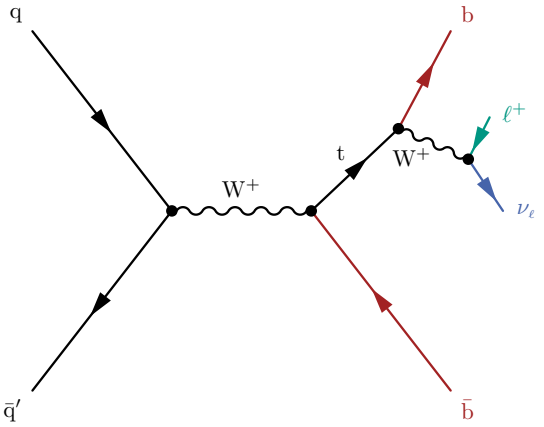

Figure 5.2: The *s*-channel single top quark signal process. The LO Feynman diagram for single top quark production in the *s* channel and the subsequent decay of the top quark are shown. Leptonically decaying top quarks are considered, where the W boson stemming from the top quark decays into a lepton-neutrino pair.

5.2.2 Background processes

All relevant background processes in the search for *s*-channel single top quark production are described in the following. In Fig. 5.3, one LO Feynman diagram, including the final-state particles, is shown for each background contribution.

t-channel single top quark production

The *t*-channel single top quark process is the dominant production mode for single top quarks at the LHC, with its predicted cross section being a factor of around 20 higher than the predicted *s*-channel cross section at $\sqrt{s} = 13$ TeV. The final state of leptonically decaying *t*-channel single top quark events comprises a lepton-neutrino pair and one bottom quark from the top quark decay. In the 4FS, a second bottom quark is produced through gluon splitting in the initial state, which is shown in Fig. 5.3, and often fails the detector acceptance due to its low transverse momentum. In contrast to the signal process, a light-flavored quark, which is preferably emitted in forward direction, is produced together with the top quark. If, for instance, the second bottom quark is not detected and if the jet stemming from the light-flavored quark is falsely reconstructed as a b-tagged jet, the final state of the signal process is obtained.

Single top quark production in association with a W boson

The production of a top quark in association with a W boson (tW) is the second most common production mode for single top quarks. In the 5FS, its final state consists of one bottom quark and two W bosons, which subsequently decay into further particles, as shown in Fig. 5.3. The final-state signature of the signal process can be mimicked by the tW process

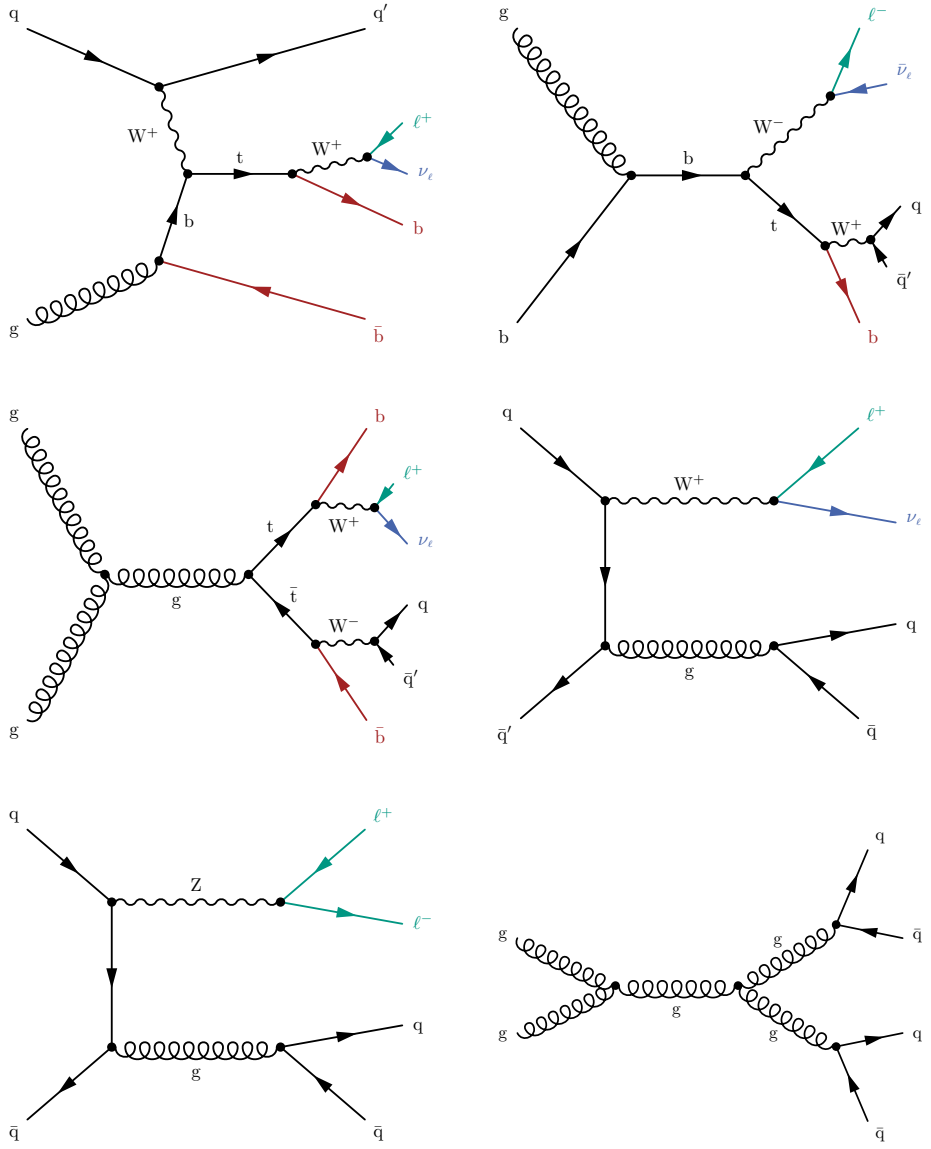

Figure 5.3: Background processes for *s*-channel single top quark production. LO Feynman diagrams for all relevant background processes that can mimic the final state of the signal process are shown: the *t*-channel and tW single top quark production modes (top left and top right), $t\bar{t}$ production, which is the dominant background (center left). Minor background contributions are the production of W and Z bosons in association with jets (center right and bottom left), and QCD multijet production (bottom right).

if one of the W bosons decays leptonically and the other hadronically, meaning that the final state of the tW process comprises a lepton-neutrino pair and two quarks of preferentially light flavor from the W boson decays, as well as the aforementioned bottom quark. If one of the light-flavored quarks is falsely identified as a bottom quark, the same final state is found as for the *s*-channel process.

Top quark pair production

Top quark pair production is the dominant production mode for top quarks, with a cross section 80 times higher than for the *s*-channel single top quark process in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$. Different decay modes exist for $t\bar{t}$ production, of which the semileptonic decay, shown in Fig. 5.3, contributes the most to the background, followed by the dileptonic decay. The final state of a semileptonically decaying top quark pair consists of a lepton-neutrino pair, two bottom quarks, and two quarks from the decay of one of the W bosons, which are preferably light-flavored. Thus, up to four jets are reconstructed for semileptonic $t\bar{t}$ processes. This background contribution can be partially rejected by requiring exactly two jets in the final state. In case of dileptonically decaying top quark pairs, the final state comprises two leptons and the corresponding neutrinos, as well as two bottom quarks, resulting in two jets measured in the detector. These events can be suppressed by requiring exactly one lepton. With the two aforementioned selection requirements, the third $t\bar{t}$ decay mode, fully hadronic, is almost entirely rejected and therefore not considered as a background process in the search for single top quark production in the *s* channel.

Production of W and Z bosons in association with jets

The production of W and Z bosons in association with jets is referred to as W+jets and Z+jets production, of which the latter is a minor background contribution. The vector bosons can be produced together with a gluon that further splits into quarks, which hadronize and form jets. These two processes can mimic the final-state signature of the signal process if the produced quarks are identified as bottom quarks and if the vector bosons decay leptonically. In case of a W boson, this leads to a lepton-neutrino pair, whereas two charged leptons are produced in the Z boson decay. If one of these two leptons is not reconstructed, the signature of the Z+jets process corresponds to the signal final state.

QCD multijet production

Another minor background contribution is the production of multiple jets through QCD interactions. Despite the fact that the final state does not contain any isolated lepton in the hard scattering with high transverse momentum, QCD multijet processes still contribute to the background due to their high production cross section of around 70 mb. This means that a small, but non-negligible number of QCD multijet events passes the selection criteria for the signal process. Jets falsely identified as leptons, or leptons produced within a jet and falsely reconstructed as isolated leptons can lead to the same final state as *s*-channel single top quark production.

5.3 Event simulation and selection

The signal process and the background contributions need to be precisely modeled to analyze the composition of the measured data. For all processes but QCD multijet production, dedicated simulation samples are used, which are described in the first part of this section. A data-driven approach is employed for the modeling of the QCD multijet background, which is described in detail in Section 5.6. In the second part of this section, the selection criteria applied to enrich signal events and to define control categories, accounting for the main background contributions, are presented.

5.3.1 Event simulation

Depending on the year of data taking and on the simulated process, different PDF sets, generator versions for MC event generation and parton shower simulation, as well as different tunes to model the underlying event and multiparton interactions are chosen, summarized in Table 5.1. All processes are generated at NLO accuracy. The *s*-channel signal process is generated in the 4FS using the MC event generator MadGraph5_aMC@NLO, whereas the other two single top quark production modes, *t* channel and tW, are simulated with powheg [164, 165, 166], applying 4FS and 5FS, respectively. The 4FS is applied to improve the simulation for *t*-channel production with regard to the second bottom quark stemming from gluon splitting. In case of the tW process, the 5FS is chosen to avoid interference at LO with $\mathrm{t\bar{t}}$ production, which is also generated with powheg [167]. For all processes involving top quarks, the value of the top quark mass is set to $m_\mathrm{t} = 172.5\,\mathrm{GeV}$. The production of W and Z bosons in association with jets is simulated with MadGraph5_aMC@NLO. For these two processes, the FxFx merging scheme is applied to account for additional partons in the matrix element calculation. The parton shower and hadronization steps are modeled with pythia for all simulated samples. In 2016, the underlying event was initially modeled with tune CUETP8M1 [168, 169], whereas tune CP5 [170] is employed for simulation samples generated in 2017 and 2018. By the time of writing this thesis, most of the simulated samples used in 2016 have been regenerated with tune CP5 to achieve a coherent description of the underlying event throughout all years of data taking. In case of the W/Z+jets processes, these samples were not available. Similarly, the 2016 samples were initially produced using the NNPDF 3.0 NLO PDF set [83], whereas the NNPDF 3.1 NNLO PDF set [84] is employed in 2017 and 2018. With the exception of W/Z+jets production, all simulated samples for the 2016 analysis are now generated with the NNPDF 3.1 NNLO PDF set. For all simulation samples, the number of pileup interactions expected to be found in data are included, and the detector response is simulated using Geant4. A full list of all MC simulation samples is provided in Appendix A.

5.3.2 Event selection

The final state of *s*-channel single top quark production comprises a bottom quark and a top quark, for which leptonic decays are considered in this thesis. Therefore, exactly one isolated muon or electron, an imbalance in the total transverse momentum caused by the neutrino, and two b-tagged jets are measured in the detector. By applying a dedicated

Table 5.1: Summary of settings for MC simulation. The different versions for PDF set, MC event generators, and parton shower and underlying event modeling used in the simulation of the signal and background processes are listed. Depending on the simulated process, differences between 2016 and 2017/2018 exist.

Simulation step	2016	2017/2018
PDF set		
NNPDF	3.1 NNLO 3.0 NLO (W/Z+jets)	3.1 NNLO
MC event generator		
MadGraph5_aMC@NLO	2.4.2 2.2.2 (W/Z+jets)	2.4.2 2.6.0 (W+jets)
powheg	v2	v2
Parton shower and underlying event		
pythia	8.226 8.212 (W/Z+jets)	8.230
Tune	CP5 CUETP8M1 (W/Z+jets)	CP5

event selection, a phase-space region is defined, for which the signal-to-background ratio is increased compared to the entire phase space, referred to as the signal event category. In addition, multiple phase-space regions are defined to serve as control event categories, which are employed to check the modeling and provide additional constraints for specific background processes.

In a first step, events of interest are preselected by passing the requirements of one or more HLT paths. Different HLT paths are applied for events with muons and electrons, which are listed in Table 5.2. Each HLT path comprises a threshold on the transverse momentum of the lepton, which is desired to be small for single top quark analyses. As the increased luminosity and the computation power of the CMS trigger system limit the available bandwidth for a single trigger, the p_{T} threshold of the chosen trigger paths differ per year of data taking. In addition, some HLT paths have a requirement on the maximum absolute pseudorapidity. From 2017 onwards, new HLT paths have been developed, which store events with one electron and one additional central jet, allowing to decrease the p_{T} threshold of the electron compared to single-electron trigger paths.

Each reconstructed lepton must pass a slightly higher p_{T} threshold than required by the chosen HLT paths to avoid the kinematic region affected by turn-on effects of the HLT paths, where the trigger efficiency may be different in data compared to simulation. For each lepton flavor and year of data taking, these p_{T} requirements are summarized in Table 5.3. As defined in Section 3.3.2, the reconstructed muon is required to have a maximum absolute pseudorapidity of $|\eta| < 2.4$ and needs to fulfill the quality criteria of the

Table 5.2: High-level trigger paths used for event selection. Different high-level trigger paths are applied per year of data taking and per lepton flavor. In cases with multiple trigger paths, events must fulfill the requirements of at least one of these paths. From 2017 onwards, dedicated trigger paths are in addition available, which select events with one electron and one jet.

Year	lepton flavor	trigger path
2016	μ	HLT_IsoTkMu24
		HLT_IsoMu24
	e	HLT_Ele32_eta2p1_WPTight_Gsf
2017	μ	HLT_IsoMu27
	e	HLT_Ele35_WPTight_Gsf
		HLT_Ele30_eta2p1_WPTight_Gsf_CentralPFJet35_EleCleaned
2018	μ	HLT_IsoMu24
	e	HLT_Ele32_WPTight_Gsf
		HLT_Ele30_eta2p1_WPTight_Gsf_CentralPFJet35_EleCleaned

tight ID working point. In addition, it must have a relative isolation of $I_{\text{rel}} < 0.06$. Events with one electron must satisfy an electron pseudorapidity requirement of $|\eta| < 2.1$ and need to pass the tight electron ID working point as defined in Section 3.3.3, where the criterion on the relative isolation is already included. Furthermore, electrons that are reconstructed from superclusters of the transition region between ECAL barrel and endcap, i. e., $1.442 < |\eta_{\text{SC}}| < 1.566$, are rejected. This definition of tight leptons suppresses most QCD multijet background events, which mainly comprise non-isolated leptons with lower transverse momenta. Background contributions from leptonically decaying Z bosons are suppressed by removing events that contain more than one lepton candidate, where the additional leptons fulfill loose selection criteria. Loose muons need to pass the loose ID working point and need to satisfy $p_{\text{T}} > 10\,\text{GeV}$, $|\eta| < 2.4$, and $I_{\text{rel}} < 0.2$, whereas loose electrons are required to pass the veto ID working point and to fulfill $p_{\text{T}} > 15\,\text{GeV}$ and $|\eta| < 2.5$. Jets that satisfy $p_{\text{T}} > 40\,\text{GeV}$, $|\eta| < 4.7$, and pass the tight jet ID working point as described in Section 3.3.4 are selected. For the analysis of 2017 and 2018 data, the p_{T} threshold is increased to 60 GeV for forward jets with $|\eta| > 2.4$ to suppress detector noise effects in the ECAL and HCAL that occurred during data taking [171, 172]. Furthermore, jets are neglected if their spatial distance to tight leptons in the η-ϕ plane is $\Delta R < 0.4$. In addition to the aforementioned jet selection criteria, candidates for b tagging must pass the medium working point of the DeepJet algorithm and must fulfill $|\eta| \leq 2.4$. The medium working point is chosen to ensure a high selection efficiency while having a moderate mistagging efficiency. The tight working point provides a lower mistagging efficiency, but also leads to a smaller selection efficiency for possible signal events, which results in a lower measured significance of the signal process. In order to further suppress the QCD multijet background, events are required to have a missing transverse momentum of $p_{\text{T}}^{\text{miss}} > 30\,\text{GeV}$. In addition, this requirement accounts for the neutrino that originates from the leptonically decaying W boson.

Table 5.3: Lepton p_T thresholds. Depending on the year of data taking and the lepton flavor, different p_T requirements are set according to the applied HLT paths.

Year	lepton flavor	p_T (GeV)
2016	μ	> 26
	e	> 35
2017	μ	> 30
	e	> 32
2018	μ	> 26
	e	> 32

Depending on the number of selected jets and b-tagged jets, four different event categories are defined. As the final state of *s*-channel single top quark production comprises two bottom quarks, the signal category is required to have two selected jets, which are both b-tagged. Therefore, the signal category is referred to as 2-jet 2-tag (2j2t) category. Three different control categories are defined to examine the modeling and to constrain specific background processes in the final fit for signal extraction. The 2-jet 1-tag (2j1t) category requires two jets, of which exactly one jet is b-tagged, and accounts for *t*-channel single top quark and W+jets production. In addition, this event category is employed to validate the data-driven approach for QCD multijet background estimation. The 3-jet 1-tag (3j1t) and 3-jet 2-tag (3j2t) categories require three jets, of which one or two are identified as b jets, respectively. These two event categories account for the dominant background process in this analysis, $t\bar{t}$ production. All event categories share the same trigger paths and the same lepton and p_T^{miss} requirements, and contain events with either one isolated muon or electron in the final state. A summary of the selection criteria for all event categories is given in Table 5.4. The expected event yields for the signal process and all background processes and of measured data in the signal and control categories are shown in Tables 5.5 and 5.6.

Table 5.4: List of event selection criteria. The 2-jet 2-tag category is the signal category. Three control categories are defined to account for the t-channel single top quark and W+jets backgrounds (2-jet 1-tag), and for the $t\bar{t}$ background (3-jet 1-tag and 3-jet 2-tag). Differences in the definitions of the event categories are highlighted.

Requirement	2-jet 1-tag (2j1t)	2-jet 2-tag (2j2t)	3-jet 1-tag (3j1t)	3-jet 2-tag (3j2t)
Trigger paths	see Table 5.2	see Table 5.2	see Table 5.2	see Table 5.2
Tight leptons	= 1	= 1	= 1	= 1
Add. loose leptons	= 0	= 0	= 0	= 0
Selected jets	***= 2***	***= 2***	***= 3***	***= 3***
b-tagged jets	***= 1***	***= 2***	***= 1***	***= 2***
$p_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV	> 30 GeV	> 30 GeV	> 30 GeV

Table 5.5: Event yields of simulated processes and measured data in the signal category. For each year of data taking, the event yields of all considered processes are listed. The yield uncertainties of all simulated processes include statistical and systematic uncertainties included in the analysis.

2j2t category	2016	2017	2018
t channel	11 500 ± 1100	17 000 ± 2100	25 900 ± 2300
tW	5340 ± 680	6700 ± 1000	11 900 ± 1400
$t\bar{t}$	136 000 ± 18 000	171 000 ± 25 000	297 000 ± 40 000
W+jets	16 100 ± 3200	20 200 ± 3800	36 400 ± 6300
Z+jets	2860 ± 550	3410 ± 550	6030 ± 870
QCD multijet	9650 ± 150	16 470 ± 230	22 420 ± 240
Sum of bkgs.	181 000 ± 18 000	235 000 ± 25 000	400 000 ± 41 000
s channel	2460 ± 190	2890 ± 230	5120 ± 370
Data	169 832	259 545	424 282

Table 5.6: Event yields of simulated processes and measured data in the control categories. For each year of data taking and each control category, the event yields of all considered processes are listed. The yield uncertainties of all simulated processes include statistical and systematic uncertainties included in the analysis.

2j1t category	2016	2017	2018
t channel	112 300 ± 9700	112 000 ± 11 000	177 000 ± 13 000
tW	59 500 ± 6200	70 600 ± 7400	115 000 ± 11 000
$t\bar{t}$	472 000 ± 63 000	562 000 ± 75 000	930 000 ± 120 000
W+jets	461 000 ± 66 000	528 000 ± 79 000	890 000 ± 130 000
Z+jets	47 100 ± 7200	56 200 ± 9300	98 000 ± 19 000
QCD multijet	145 460 ± 390	243 890 ± 670	386 400 ± 810
Sum of bkgs.	1 297 000 ± 92 000	1 570 000 ± 110 000	2 600 000 ± 180 000
s channel	3290 ± 140	3150 ± 190	5420 ± 300
Data	1 225 890	1 687 290	2 709 810

3j1t category	2016	2017	2018
t channel	29 800 ± 3500	24 700 ± 3800	38 800 ± 5700
tW	42 100 ± 4700	45 300 ± 5200	74 800 ± 8500
$t\bar{t}$	511 000 ± 72 000	533 000 ± 72 000	890 000 ± 120 000
W+jets	158 000 ± 30 000	144 000 ± 27 000	227 000 ± 42 000
Z+jets	17 700 ± 3500	17 900 ± 4200	30 200 ± 7300
QCD multijet	60 310 ± 310	97 840 ± 550	142 380 ± 610
Sum of bkgs.	819 000 ± 78 000	863 000 ± 77 000	1 400 000 ± 130 000
s channel	964 ± 81	860 ± 100	1490 ± 180
Data	794 537	957 258	1 471 660

3j2t category	2016	2017	2018
t channel	18 100 ± 1900	18 400 ± 2400	28 500 ± 2600
tW	10 100 ± 1300	11 900 ± 1800	20 900 ± 2600
$t\bar{t}$	305 000 ± 43 000	357 000 ± 50 000	618 000 ± 82 000
W+jets	13 000 ± 2200	12 800 ± 2300	19 300 ± 3800
Z+jets	2130 ± 580	2350 ± 490	4000 ± 830
QCD multijet	12 590 ± 390	10 180 ± 350	7750 ± 200
Sum of bkgs.	361 000 ± 43 000	413 000 ± 50 000	698 000 ± 82 000
s channel	883 ± 73	960 ± 110	1670 ± 170
Data	328 502	455 725	716 013

5.4 Top quark reconstruction

As top quarks decay before they reach the detector, they need to be reconstructed from their decay products. In this thesis, leptonically decaying top quarks are considered, where the decay products are a bottom quark and a W boson, which subsequently decays into a charged lepton and its corresponding neutrino. Hadronically decaying top quarks, where the W boson decays into a quark-antiquark pair, are not considered to suppress background events from QCD multijet processes. The lepton considered in the top quark reconstruction is a muon or electron that either directly originates from the W boson decay or is created through the decay of a tau lepton stemming from the W boson. In the first step of the top quark reconstruction, the four-momentum of the W boson is determined. As the neutrino cannot be directly measured in the detector, its four-momentum is inferred from the missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$. The information on the longitudinal neutrino momentum is missing, making the full reconstruction of the W boson four-momentum challenging. Assuming that the x and y components of $p_{\mathrm{T}}^{\mathrm{miss}}$ correspond to the ones of the neutrino and applying a constraint on the W boson mass by requiring the literature value of $m_{\mathrm{W}} = 80.385\,\mathrm{GeV}$ [14], one obtains for the W boson mass:

$$m_{\mathrm{W}}^2 = 2 \cdot \left(E_\ell \sqrt{{p_{\mathrm{T}}^{\mathrm{miss}}}^2 + p_{z,\nu}^2} - p_{\mathrm{T},\ell} p_{\mathrm{T}}^{\mathrm{miss}} \cos\Delta\phi - p_{z,\ell} p_{z,\nu} \right) = (80.385\,\mathrm{GeV})^2 \,. \tag{5.1}$$

The energy of the lepton is defined as $E_\ell = \sqrt{p_{\mathrm{T},\ell}^2 + p_{z,\ell}^2}$ and $\Delta\phi$ is the azimuthal angle between the charged lepton $p_{\mathrm{T},\ell}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$. The z component of the neutrino momentum is then determined as the solution of a quadratic equation:

$$p_{z,\nu} = \frac{\Lambda p_{z,\ell}}{p_{\mathrm{T},\ell}^2} \pm \sqrt{\frac{\Lambda^2 p_{z,\ell}^2}{p_{\mathrm{T},\ell}^4} - \frac{E_\ell^2 \left(p_{\mathrm{T}}^{\mathrm{miss}}\right)^2 - \Lambda^2}{p_{\mathrm{T},\ell}^2}}, \tag{5.2}$$

with

$$\Lambda = \frac{m_{\mathrm{W}}^2}{2} + p_{\mathrm{T},\ell} p_{\mathrm{T}}^{\mathrm{miss}} \cos\Delta\phi. \tag{5.3}$$

Depending on the value of the discriminant in Eq. (5.2), different solutions for $p_{z,\nu}$ are obtained. If the discriminant is positive, two solutions exist, of which the one with the smaller absolute value is chosen. Due to finite detector resolution, the $p_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is imperfect and can lead to a negative discriminant resulting in complex $p_{z,\nu}$ solutions. In this case, which occurs in approximately one third of all s-channel single top quark events, the x and y components of $p_{\mathrm{T}}^{\mathrm{miss}}$ are slightly modified without violating the requirement $m_{\mathrm{W}} = 80.385\,\mathrm{GeV}$, such that the discriminant becomes zero [173]. With the solution for $p_{z,\nu}$, the neutrino four-momentum is determined and used together with the lepton four-momentum to reconstruct the W boson. The top quark is then reconstructed with the

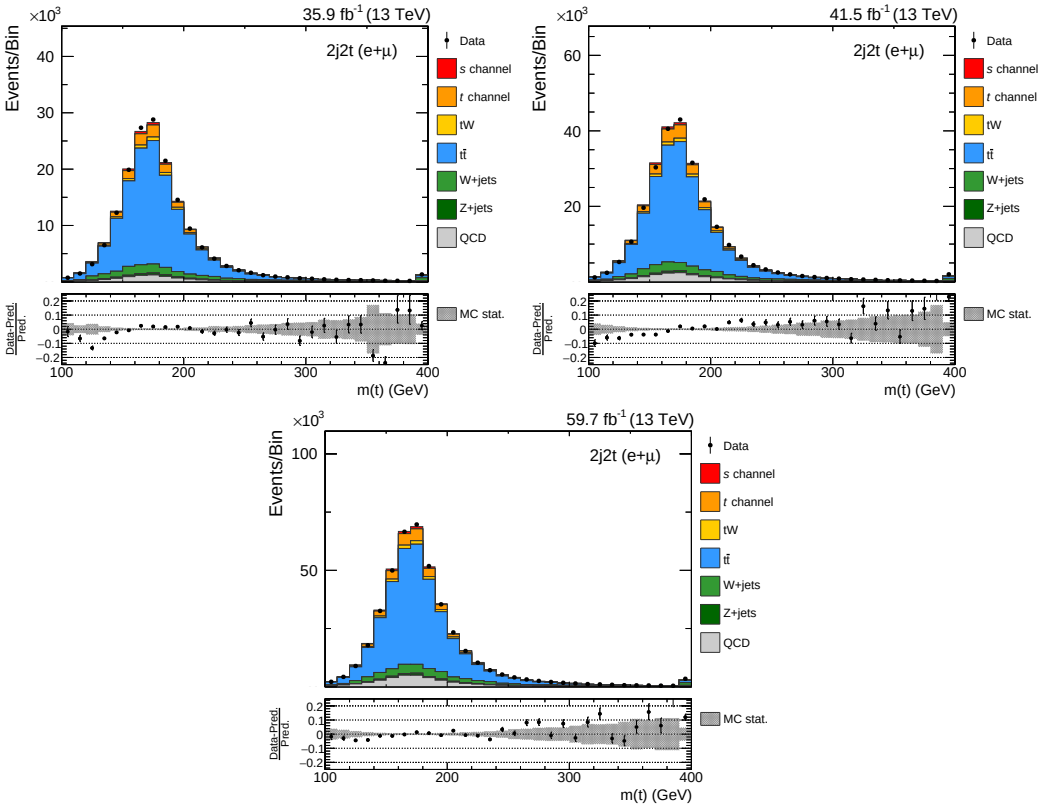

Figure 5.4: Reconstructed top quark mass in the 2j2t signal category. The distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) are shown. The prediction is scaled to the number of measured events and all corrections later described in Section 5.5 are applied on the simulated events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

four-momenta of the W boson and the b jet candidate stemming from the top quark decay. In case of more than one possible b jet candidate, different top quark hypotheses are determined by calculating the invariant top quark mass for each candidate. The b jet candidate leading to an invariant top quark mass closest to the value used in simulation, $m_t = 172.5\,\text{GeV}$, is chosen for the top quark reconstruction. In Fig. 5.4, the reconstructed top quark mass is shown for the 2j2t signal category for each year of data taking. For all other event categories, the reconstructed top quark mass is shown in Appendix B.

5.5 Event correction

Last remaining differences between simulation and data are resolved by calculating efficiency corrections and applying dedicated event weights. For most efficiencies, official scale factors are provided by the physics object groups (POGs) of the CMS Collaboration.

5.5.1 Number of pileup vertices

Depending on the year of data taking, conditions like the instantaneous luminosity slightly change, leading to a different expected average number of pileup interactions for each year. As the measured maximum instantaneous luminosity in 2017 and 2018 is approximately $2.1 \cdot 10^{32}\,\mathrm{cm^{-2}s^{-1}}$ and therefore higher than the measured maximum instantaneous luminosity of around $1.5 \cdot 10^{32}\,\mathrm{cm^{-2}s^{-1}}$ in 2016, the average number of primary vertices per event in these two years is increased to 32 compared to 23 in 2016 [174]. Due to the time-consuming production of the simulation samples, they are already produced before and during the data-taking period. This requires a preliminary pileup simulation, in which the pileup distribution is modeled from a Poisson distribution with the mean set to a rough estimate of the average number of pileup interactions [175], which can lead to a significant discrepancy compared to measured data. Therefore, each simulated event is reweighted such that the simulated distribution agrees with the number of pileup interactions in data, which is computed with a total inelastic proton-proton cross section of 69.2 mb [176, 177, 178]. In Fig. 5.5, the effect of pileup reweighting is shown for each year of data taking. After applying pileup reweighting, the agreement between simulation and measured data is only slightly improved and mismodeling is still visible. By choosing an alternative inelastic proton-proton cross section of 80.0 mb, as predicted by PYTHIA, in the pileup reweighting, significantly improved agreement can be obtained. As the total inelastic proton-proton cross section is accurately determined at the CMS experiment and as all relevant kinematic distributions are not effected by the reweighting, this alternative approach is not employed nevertheless.

5.5.2 Lepton efficiencies

Remaining differences in the reconstruction efficiencies of leptons between data and simulation are corrected using dedicated scale factors, which depend on lepton p_T and η and which take different effects of the lepton reconstruction into account.

For muons, the overall efficiency is computed as

$$\epsilon = \epsilon_\mathrm{ID} \cdot \epsilon_{\mathrm{isolation|ID}} \cdot \epsilon_{\mathrm{trigger|isolation}}, \tag{5.4}$$

with identification (ID), isolation, and trigger efficiencies. As indicated in Eq. (5.4), the isolation efficiency depends on the calculated ID efficiency and the trigger efficiency on the isolation efficiency. Each efficiency correction is determined with a tag-and-probe method using J/ψ meson or Z boson resonances [179, 180] and is provided by the MUO POG for each year of data taking [181, 182, 183]. As a very strict muon isolation requirement of $I_\mathrm{rel} < 0.06$ is imposed to suppress QCD multijet production, no official isolation scale factors are available. Therefore, privately derived trigger and isolation scale factors, which are described in detail in Ref. [184], are employed.

For electrons, efficiency corrections for identification and reconstruction are computed with a tag-and-probe method using $\mathrm{Z} \to \mathrm{e^+e^-}$ events [185, 186] and are centrally provided

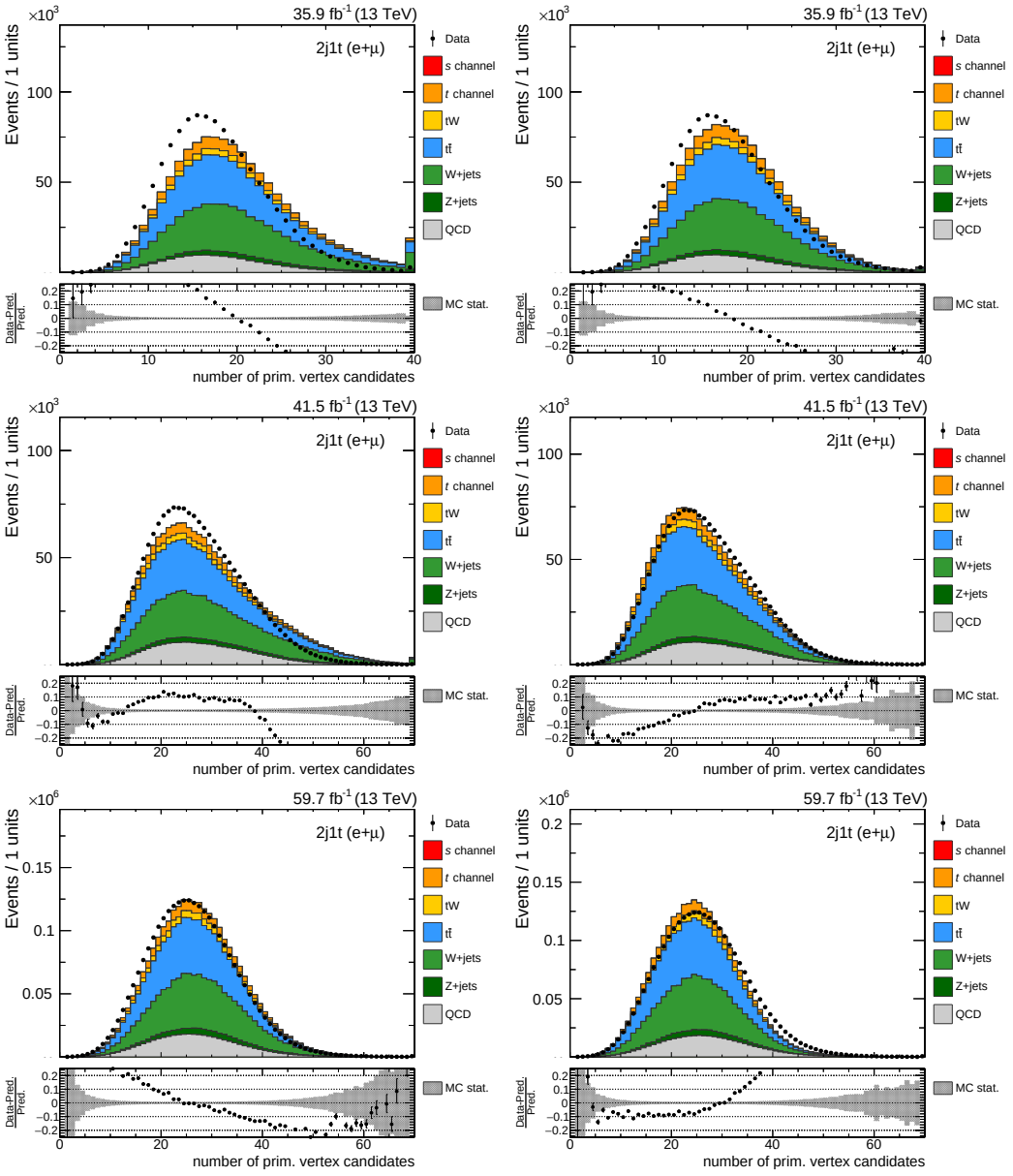

Figure 5.5: Application of pileup reweighting on the distribution of the number of primary vertices. The distributions are shown in the 2j1t control category before (left) and after (right) pileup reweighting, for the 2016 (top), 2017 (center), and 2018 (bottom) data. Before and after pileup reweighting, all other event weights considered in this thesis are applied. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

by the EGamma POG [187]. As the electron trigger efficiency scale factors are not centrally available, custom scale factors need to be computed. Depending on the chosen electron trigger, two different methods are employed. In case of the single electron trigger of the 2016 analysis, *HLT_Ele32_eta2p1_WPTight_Gsf*, dedicated scale factors are calculated with the tag-and-probe method, as described in Ref. [188]. The trigger *HLT_Ele30_eta2p1_WP Tight_Gsf_CentralPFJet35_EleCleaned* combines requirements on the electron and one jet and is used together with a single electron trigger path in the analysis of 2017 and 2018 data. For this trigger, a different approach [189] needs to be employed to determine the scale factors. In this method, data samples containing single muon events and simulated dileptonically decaying $\mathrm{t\bar{t}}$ events are used to compute the trigger efficiencies. Both data and simulation need to pass a loosened selection for *s*-channel single top quark events, which is summarized in Table 5.7. The efficiency of the chosen electron trigger paths as listed in Table 5.2 is defined by

$$\epsilon_{\text{trigger}} = \frac{N_{\text{e-triggers+}\mu\text{-trigger}}}{N_{\mu\text{-trigger}}}, \tag{5.5}$$

with the number of events $N_{\mu\text{-trigger}}$ passing the aforementioned event selection and the single muon trigger path as given in Table 5.2 and the number of events $N_{\text{e-triggers+}\mu\text{-trigger}}$ passing in addition the electron trigger paths. For this method, uncorrelated electron and muon trigger paths need to be employed. The correlation is checked on the dileptonic $\mathrm{t\bar{t}}$ simulation sample by computing

$$\alpha = \frac{\epsilon^{\text{MC}}_{\mu\text{-trigger}} \cdot \epsilon^{\text{MC}}_{\text{e-triggers}}}{\epsilon^{\text{MC}}_{\text{e-triggers+}\mu\text{-trigger}}}. \tag{5.6}$$

A value of $\alpha = 1$ means that the electron trigger paths and the single muon trigger path are uncorrelated. In 2017 and 2018, the values for α were determined as 0.998. The trigger efficiency is evaluated in different p_{T} and $|\eta|$ bins of the electron and of each jet, which are chosen such that the number of events is approximately the same in each bin. The trigger scale factors are calculated as

$$\text{SF}_{\text{trigger}} = \frac{\epsilon^{\text{Data}}_{\text{trigger}}}{\epsilon^{\text{MC}}_{\text{trigger}}}, \tag{5.7}$$

for three different cases: electron p_{T} and $|\eta|$, electron p_{T} and leading jet p_{T}, and leading jet p_{T} and $|\eta|$. The corresponding scale factors are shown in Appendix C, separately for 2017 and 2018. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied on the calculated scale factors to account for a possible performance difference of the triggers in data and simulation. As the trigger efficiencies depend mostly on electron p_{T} and $|\eta|$, the scale factors calculated in different electron p_{T} and $|\eta|$ bins are chosen to be applied in this analysis.

Table 5.7: Event selection criteria used for calculation of electron trigger efficiencies. The event selection is based on the 2j2t signal category (see Table 5.4), but requires loosened selection criteria to ensure a significantly higher number of events for efficiency computation.

Physics object	number
Tight electrons	= 1
Add. loose electrons	= 0
Jets ($p_T > 40$ GeV, $\|\eta\| < 2.4$, no b tagging)	= 2

In Fig. 5.6, the effect of lepton efficiency scale factors is shown for the analysis of 2016, 2017, and 2018 data. The difference in the shape of the reconstructed lepton p_T distribution caused by the applied weights is negligible. In addition, the overall normalization is changed by approximately 4%.

5.5.3 b tagging efficiency

The simulated events are reweighted using b tagging efficiency scale factors to predict the corrected event yield as observed in data. A dedicated method [190] is applied, for which only simulated events passing the analysis-specific selection need to be considered in the final event weight calculation. This means that no events are migrated to a different b-tagged jet multiplicity. By conserving the multiplicity of b-tagged jets, the definition of variables that rely on a specific number of b-tagged jets is maintained, as these variables are otherwise ill-defined if, for instance, one b-tagged jet is missing in an event due to the jet-multiplicity migration.

For each simulated process, efficiencies for identifying jets of true flavor f as b jets via b tagging need to be determined before the event selection is applied. The efficiencies are determined in dependence of specific jet p_T and η bins and are given by

$$\epsilon_f(p_T, \eta) = \frac{N_f^{\text{b-tagged}}(p_T, \eta)}{N_f^{\text{total}}(p_T, \eta)}, \tag{5.8}$$

with the total number of jets N_f^{total} and b-tagged jets $N_f^{\text{b-tagged}}$ with true flavor f. For b-flavored jets, ϵ_b corresponds to the b tagging efficiency, i. e., the fraction of b jets that is correctly identified by the b tagging algorithm. In case of c-flavored and light-flavored jets, ϵ_c and ϵ_{udsg} are mistagging efficiencies for falsely identifying c jets and jets from u, d, s quarks and gluons (udsg) as b jets, respectively.

The probabilities for correctly identifying all jets in simulation and data are determined for a given number of jets and b-tagged jets in the event selection and are calculated as

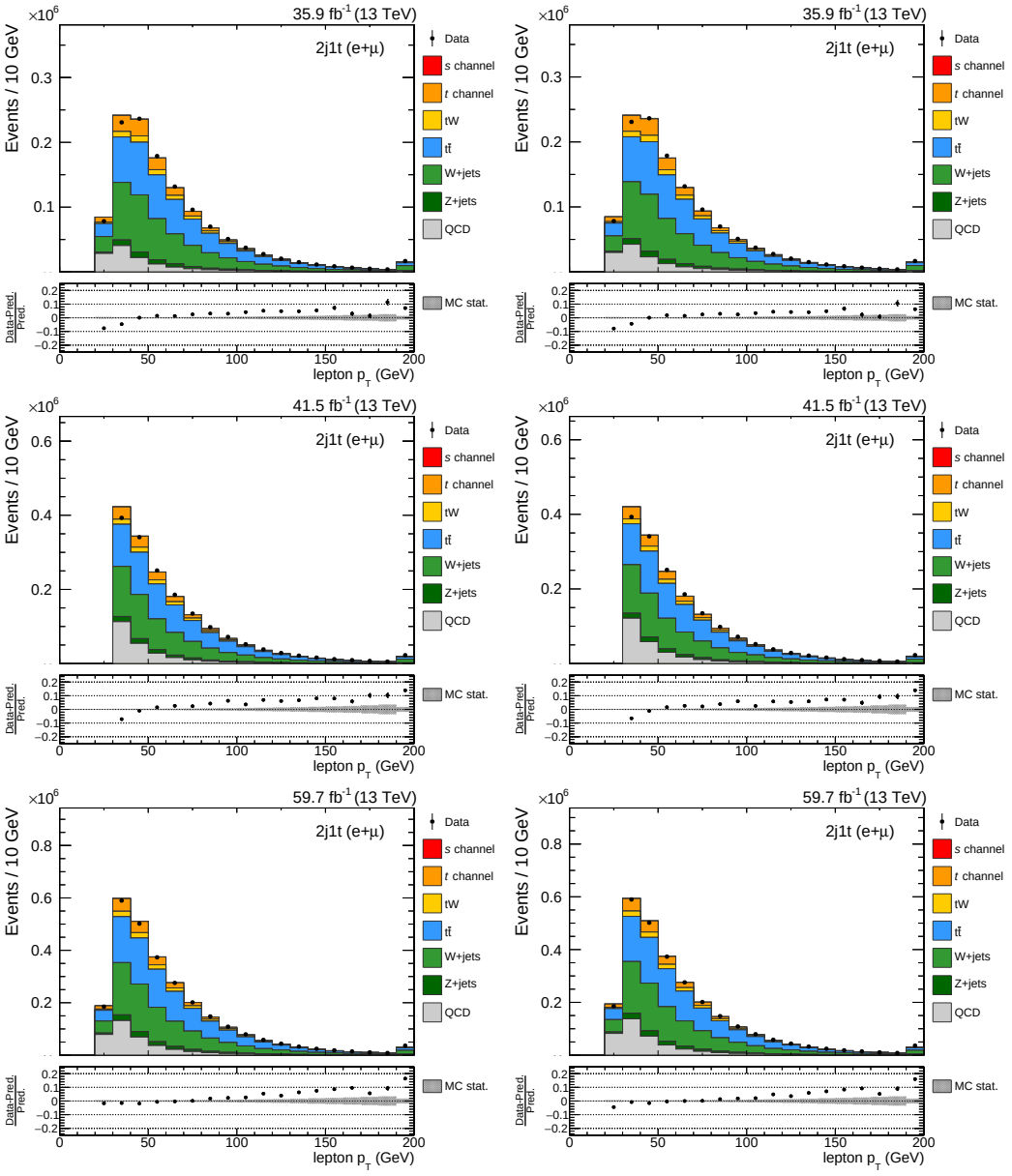

Figure 5.6: Application of lepton efficiency scale factors on the p_T distribution of the reconstructed isolated lepton. The distributions are shown in the 2j1t control category before (left) and after (right) applying lepton efficiency scale factors, for the 2016 (top), 2017 (center), and 2018 (bottom) data. Before and after lepton efficiency correction, all other event weights considered in this thesis are applied. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

$$P(\text{MC}) = \prod_{i=\text{tagged}} \epsilon_{\text{b}}^{i} \prod_{j=\text{untagged}} \left(1 - \epsilon_{\text{c, udsg}}^{j}\right), \tag{5.9}$$

$$P(\text{Data}) = \prod_{i=\text{tagged}} \text{SF}_{\text{b}}^{i} \epsilon_{\text{b}}^{i} \prod_{j=\text{untagged}} \left(1 - \text{SF}_{\text{c, udsg}}^{j} \epsilon_{\text{c, udsg}}^{j}\right). \tag{5.10}$$

Here, ϵ_{b}^{i} and $\epsilon_{\text{c, udsg}}^{j}$ are the b-tagging and mistagging efficiencies in simulation for each b-tagged jet i and untagged jet j, and SF_{b}^{i} and $\text{SF}_{\text{c, udsg}}^{j}$ are the scale factors of the DeepJet algorithm for b jets and jets of flavor c or udsg, respectively. These scale factors express the difference in the b tagging and mistagging efficiencies between data and simulation and are defined as $\text{SF} = \epsilon_{\text{Data}}/\epsilon_{\text{MC}}$. They are centrally provided for the CMS Collaboration by the b Tag & Vertexing POG [191, 192, 193].

The final event weight is then computed as

$$w = \frac{P(\text{Data})}{P(\text{MC})}. \tag{5.11}$$

The effect of b tagging weights on the shape of the distribution of the reconstructed b-tagged jet with highest p_{T} is shown in Fig. 5.7 for each year of data taking. Differences in the shape are mainly observed for a low p_{T} range of around 40 GeV to 60 GeV. The event yield is changed by around 3% when applying b tagging weights.

5.5.4 L1 ECAL prefiring

During the 2016 and 2017 data-taking periods, a high fraction of high-η L1 trigger primitives [71] (TPs), which are electronic signals generated by the ECAL employed in the L1 trigger, was erroneously associated to the previous bunch crossing. This issue was caused by the gradual timing shift of the ECAL readout electronics not being propagated properly to the L1 TPs. This leads to events that can veto themselves, as two consecutive bunch crossings are forbidden to fire the L1 trigger according to L1 rules. This veto occurs in case of a significant ECAL energy deposit in the region $2.0 < |\eta| < 3.0$ and is not modeled in simulation. Therefore, weights are computed based on the probability of an event not to prefire and applied to the simulated events [194]. These weights are defined as

$$w = 1 - P(\text{prefire}) = \prod_{\substack{i=\text{jets,} \\ \text{photons}}} \left(1 - \epsilon_{\text{prefire}}^{i}(\eta, p_{\text{T}})\right), \tag{5.12}$$

with the prefiring efficiency $\epsilon_{\text{prefire}}^{i}$ of a jet or photon i and the transverse momentum p_{T} of a photon or jet. In Fig. 5.8, the prefiring efficiencies for 2016 and 2017 data are shown in dependence of jet p_{T} and η.

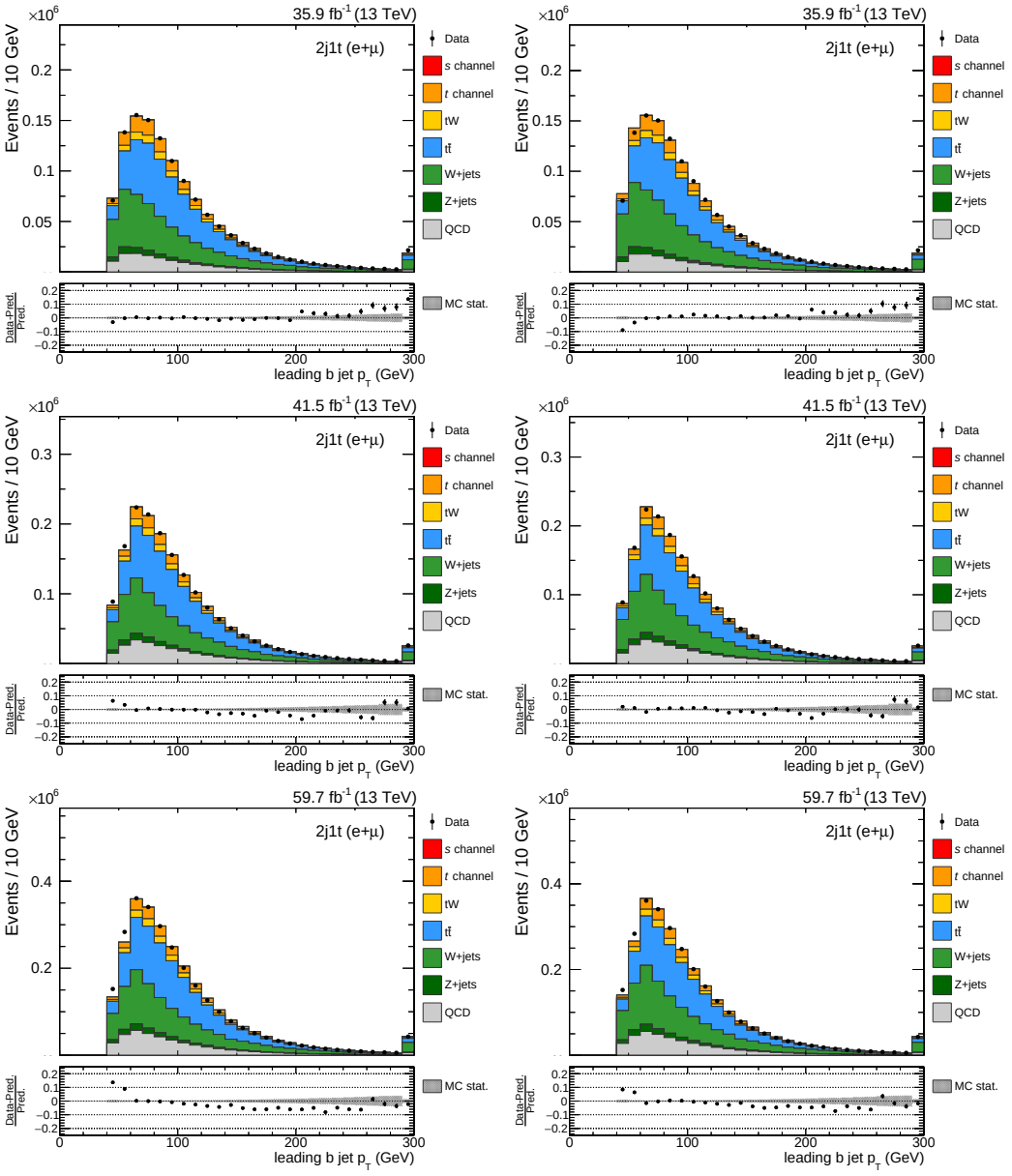

Figure 5.7: Application of b tagging weights on the distribution of the reconstructed b-tagged jet with highest p_T. The distributions are shown in the 2j1t control category before (left) and after (right) applying b tagging weights, for the 2016 (top), 2017 (center), and 2018 (bottom) data. Before and after b tagging efficiency correction, all other event weights considered in this thesis are applied. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

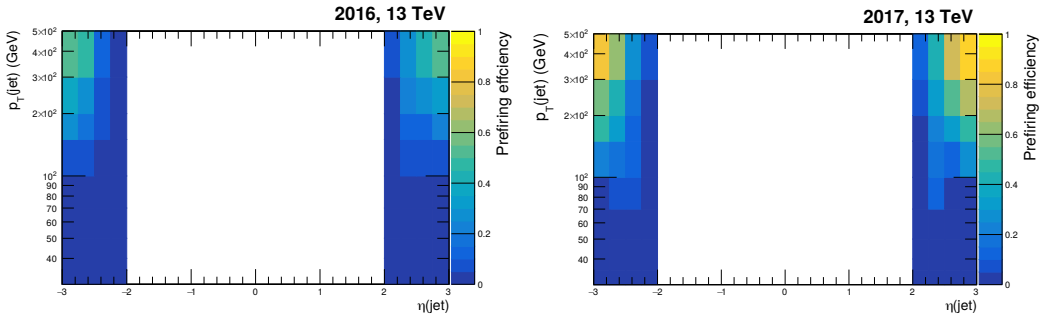

Figure 5.8: L1 ECAL prefiring efficiencies for jets. The prefiring efficiencies in dependence of jet p_T and η are centrally provided [194]. Modified version taken from Ref. [195].

5.6 QCD multijet background estimation

The modeling of the QCD multijet background contribution is challenging because of the high production cross section and the tiny event selection efficiency of this process. Extremely large simulation samples are required to accurately model QCD multijet production, which is a computationally intensive task.

Thus, an alternative approach is employed for the QCD multijet modeling using data from sideband regions enriched in QCD multijet events. The QCD-enriched sideband regions are obtained by inverting the isolation criterion of the selected lepton and removing the requirement $p_T^{miss} > 30\,GeV$, with the same requirements for jets and b-tagged jets as described in Section 5.3.2. As the modeling of other kinematic variables for QCD multijet production is independent of the applied isolation criterion, the QCD multijet modeling in the sideband regions can be used for the signal and control event categories. For muons, a relative isolation of $I_{rel} > 0.2$ is required, and electrons must fail the veto ID working point, described in Table 3.2, as the relative electron isolation is part of the ID requirements.

By performing a fit on a discriminating variable, the QCD contribution is estimated. With the shape of this variable, QCD multijet events can be distinguished from all other processes considered in this thesis. In the latest t-channel single top quark cross section measurement of the CMS Collaboration [158], different variables are used depending on the lepton flavor. For events with muons, the transverse W boson mass m_T^W is employed, whereas the missing transverse momentum p_T^{miss} is used in case of electrons. This difference is caused by a mismodeling of the m_T^W distribution for electrons. In this thesis, only p_T^{miss} is considered as a discriminating variable for the estimation of QCD multijet events, as no distinction is made between muons and electrons. The QCD template $Q(p_T^{miss})$ used for the estimation is obtained from the data distribution in each sideband region. As the sideband regions contain a small fraction of non-QCD multijet events, these are estimated from simulation and subtracted from the QCD template.

A maximum likelihood fit is performed on the p_T^{miss} distribution of the data in each event category. This fit employs the aforementioned QCD template and a non-QCD template

Table 5.8: Estimation of the QCD multijet background contribution in the 2j2t signal category. The event yields are obtained from the fit of the missing transverse momentum distribution and are displayed together with their statistical uncertainties for the full range and for $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$.

Year	QCD yield (n_{QCD}) full range	QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$	non-QCD yield ($n_{\text{non-QCD}}$) full range	non-QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$
2016	18 810 ± 210	9650 ± 150	197 360 ± 540	159 890 ± 450
2017	25 990 ± 300	16 470 ± 230	295 500 ± 710	243 020 ± 640
2018	36 350 ± 300	22 420 ± 240	491 700 ± 1200	402 100 ± 1100

$N(p_{\mathrm{T}}^{\mathrm{miss}})$, which is determined by the sum of all processes with isolated leptons, including the signal process, and has two fit parameters n_{QCD} and $n_{\text{non-QCD}}$, which represent the number of QCD and non-QCD multijet events, respectively. The $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in data is then parameterized as

$$D(p_{\mathrm{T}}^{\mathrm{miss}}) = n_{\mathrm{QCD}} \cdot Q(p_{\mathrm{T}}^{\mathrm{miss}}) + n_{\text{non-QCD}} \cdot N(p_{\mathrm{T}}^{\mathrm{miss}}). \tag{5.13}$$

The fit is performed separately for each event category and for each year of data taking. The fit result for the 2j2t signal category is shown in Fig. 5.9, and the obtained number of QCD and non-QCD multijet events in Table 5.8. In Appendix D, the outcome of the QCD estimation is presented for the 2j1t, 3j1t, and 3j2t control categories.

5.7 Event classification

As the signal-to-background ratio is still tiny after applying the event selection, advanced multivariate methods are employed to discriminate *s*-channel single top quark production from its main background contribution, top quark pair production. For this purpose, a deep neural network (DNN) implemented in Keras [145] is used, which is configured as shown in Table 5.9. A separate training of the DNN is performed for each year to take the different data-taking conditions into account. In all three years, the same signal and background processes, the same DNN configuration, and the same input variables are used. The available simulated signal and background samples are split into two subsets, one of which is only employed for the DNN training and the other one only used for the further analysis to avoid a biased event classification result. The signal process, *s*-channel single top quark production, is trained against semileptonically and dileptonically decaying top quark pair events in the 2j2t signal category. The signal and background samples are normalized according to their predicted cross section multiplied by the event selection efficiency. For technical reasons, the signal training sample comprises a third of all available signal events in 2016 and 2018, and half of the events in 2017. In all three years, 12% of the available $\mathrm{t\bar{t}}$ events are used as background training sample to have approximately twice as many background as signal events for the training. This makes sure that both semileptonically and dileptonically decaying $\mathrm{t\bar{t}}$ events are trained against the signal process. No

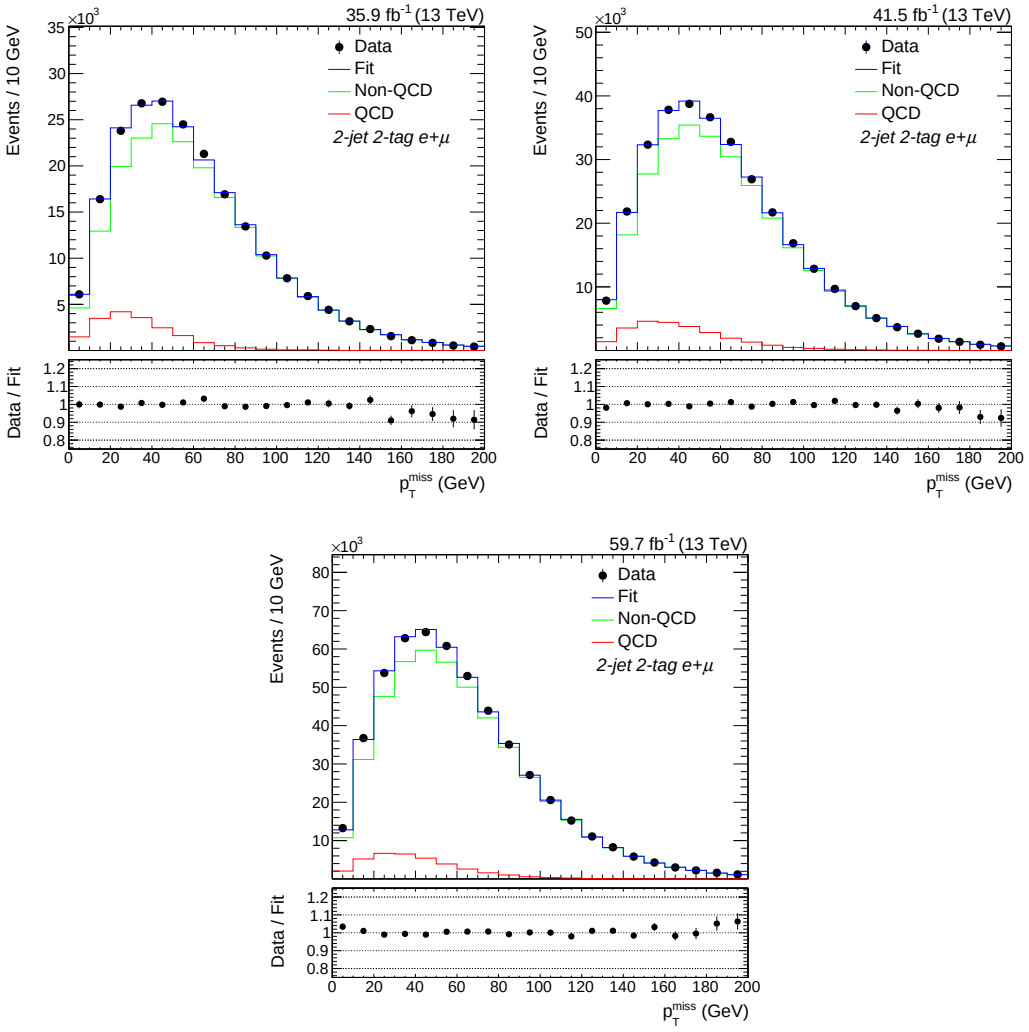

Figure 5.9: Estimation of the QCD multijet background contribution in the 2j2t signal category. The fit is performed to the $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) data. The error bars in the ratio panel correspond to statistical uncertainties.

Table 5.9: Parameter settings of the DNN training for event classification. The DNN is implemented with the Keras software package [145]. Details on the DNN parameters are explained in Section 4.1. The early-stopping method prevents overtraining of the DNN by stopping the training when the loss of a validation subset of the training data (*val_loss*) did not approved within 50 epochs. During one epoch, the entire training dataset is passed forward and backward through the DNN once.

Parameter	value
Topology	2 hidden layers (activation function f_{ReLU}, 100 neurons per layer, dropout for 25% of neurons) output layer (activation function f_{sig}, 1 neuron)
Epochs	100
Batch size	512
Metrics	accuracy
Loss	binary crossentropy
Optimizer	Adam (default settings, see Ref. [144])
Early stopping	monitor: val_loss, patience: 50

dedicated training is performed in the control categories as they serve to constrain the background contributions in the signal extraction procedure.

Different kinematic input variables are considered for the DNN training, which describe the characteristics of the *s*-channel process and are required to provide separation power between signal and background. These need to be well-modeled when compared to data. In addition, the correlation between the variables needs to be studied. The full list of the kinematic variables is given in Table 5.10. No unexpected correlation between the chosen input variables was found. Depending on the year of data taking, the importance of the input variables to the event classification differs. The variable importance is estimated by calculating the total sum of all weights of the first hidden layer for each input variable. The ranking of the input variables employed in the DNN training are shown for each year in Table 5.11. In addition, the sum of all weights of the first hidden layer is listed for each variable for all years in Appendix E.

For the 2016 analysis, the three most important variables are the invariant mass of the lepton and subleading b jet, the reconstructed lepton p_{T}, and the absolute difference between η of the lepton and η of the leading b jet. The leading and subleading b jet is the reconstructed b-tagged jet with highest and second highest p_{T}, respectively. The modeling and the shapes of these three variables are shown for the 2j2t signal category in Fig. 5.10.

For the 2017 analysis, the azimuthal difference between reconstructed top quark and leading b jet, $p_{\text{T}}^{\text{miss}}$, and the absolute η difference between the top quark and the b jet stemming from the virtual W boson, contribute the most. In Fig. 5.11, the modeling and the shape distributions of these variables are shown.

For the 2018 analysis, the invariant mass of the lepton and subleading b jet, $p_{\mathrm{T}}^{\mathrm{miss}}$, and the azimuthal difference between reconstructed top quark and leading b jet are the three most important variables and are shown in Fig. 5.12. For all presented input variables, a significant shape difference between the signal process and the background processes, in particular $\mathrm{t\bar{t}}$ production, is observed. The remaining input variables are shown in Appendix E.

The DNN training is checked for overtraining by comparing the DNN output distribution of the signal and background training samples, which are discarded from further analysis to avoid a bias on the final result, with separate test samples for signal and background. In Fig. 5.13, the result of the comparison is shown for each year and a clear separation between signal and background is observed. As a smaller number of signal events is available in 2016 compared to 2017 and 2018 and as the geometry of the CMS pixel detector was different in 2016, the shape of the DNN output distribution differs. In all three years, the training and test samples show no sign of overtraining. This is also verified by checking the loss function values of the training and test datasets with regard to the epoch of the DNN training. As the loss functions of the test samples do not increase towards the end of the DNN training, no potential overtraining is observed. For each year, the loss function values are shown in Appendix E. The area under the ROC curve is determined to be 82% for 2016 and 78% for 2017 and 2018, i. e., a slightly better separation between signal and background is achieved by the DNN training compared to 2017 and 2018. This effect is caused by the higher pileup contribution in 2017 and 2018 compared to 2016. The DNN is applied to the 2j2t signal category and all control categories. For control categories with only one b-tagged jet, the subleading b jet always corresponds to the untagged jet. In case of the 2j1t category, only one untagged jet exists and is thus chosen as subleading b jet, whereas two untagged jets are selected for the 3j1t control category. Here, the untagged jet with highest p_{T} is assigned as subleading b jet.

Table 5.10: Description of the input variables employed in the training of the DNN. Leading and subleading b jet refer to the reconstructed b jet with highest and second highest transverse momentum, respectively.

Variable	description
$p_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse momentum
$p_{\mathrm{T}}(\ell)$	transverse momentum of the charged lepton
$p_{\mathrm{T}}(\mathrm{b_l})$	transverse momentum of the leading b jet
$p_{\mathrm{T}}(\mathrm{b_s})$	transverse momentum of the subleading b jet
$p_{\mathrm{T}}^{\mathrm{bb}}$	vectorial sum of transverse momenta of the two b jets in the event
H_{T}	scalar p_{T} sum of all b jets, the charged lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event
$\eta(\mathrm{b_l})$	pseudorapidity of the leading b jet
$\eta(\mathrm{b_s})$	pseudorapidity of the subleading b jet
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	absolute pseudorapidity difference between the charged lepton and the leading b jet
$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	absolute pseudorapidity difference between the top quark reconstructed from the subleading b jet and the leading b jet
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	absolute pseudorapidity difference between the reconstructed top quark and the b jet from the virtual W boson
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	$\Delta\phi$ between the reconstructed top quark and the leading b jet
$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	$\Delta\phi$ between the two b jets of the event
$\cos\theta^*$	cosine of the angle between the charged lepton from top quark decay and the b jet from virtual W boson in the rest frame of the top quark
$m(\mathrm{t})$	invariant mass of the reconstructed top quark
$m_{\ell \mathrm{b2}}$	invariant mass of the charged lepton and the subleading b jet
Fox Wolfram #3	third-order Fox-Wolfram moment [196, 197] of the event
$q(\ell)$	lepton charge

Table 5.11: Ranking of input variables used in the training of the DNN. The ranking is an estimate of the importance of each input variable, determined by the sum of all weights of the first layer of the DNN for the variable under study. For each year, a different ranking is obtained, reflecting the different conditions during data taking per year and showing that the input variables are mainly of similar importance.

Variable	2016	2017	2018
$m_{\ell \mathrm{b}2}$	1	4	1
$p_{\mathrm{T}}(\ell)$	2	11	7
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	3	7	4
$p_{\mathrm{T}}(\mathrm{b_s})$	4	10	8
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	5	1	3
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	6	3	6
$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	7	15	5
$\eta(\mathrm{b_s})$	8	9	9
$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	9	6	10
Fox Wolfram #3	10	14	16
$q(\ell)$	11	8	11
$\eta(\mathrm{b_l})$	12	5	14
$\cos\theta^*$	13	12	15
$p_{\mathrm{T}}^{\mathrm{miss}}$	14	2	2
$m(\mathrm{t})$	15	13	13
$p_{\mathrm{T}}^{\mathrm{bb}}$	16	17	17
H_{t}	17	16	12
$p_{\mathrm{T}}(\mathrm{b_l})$	18	18	18

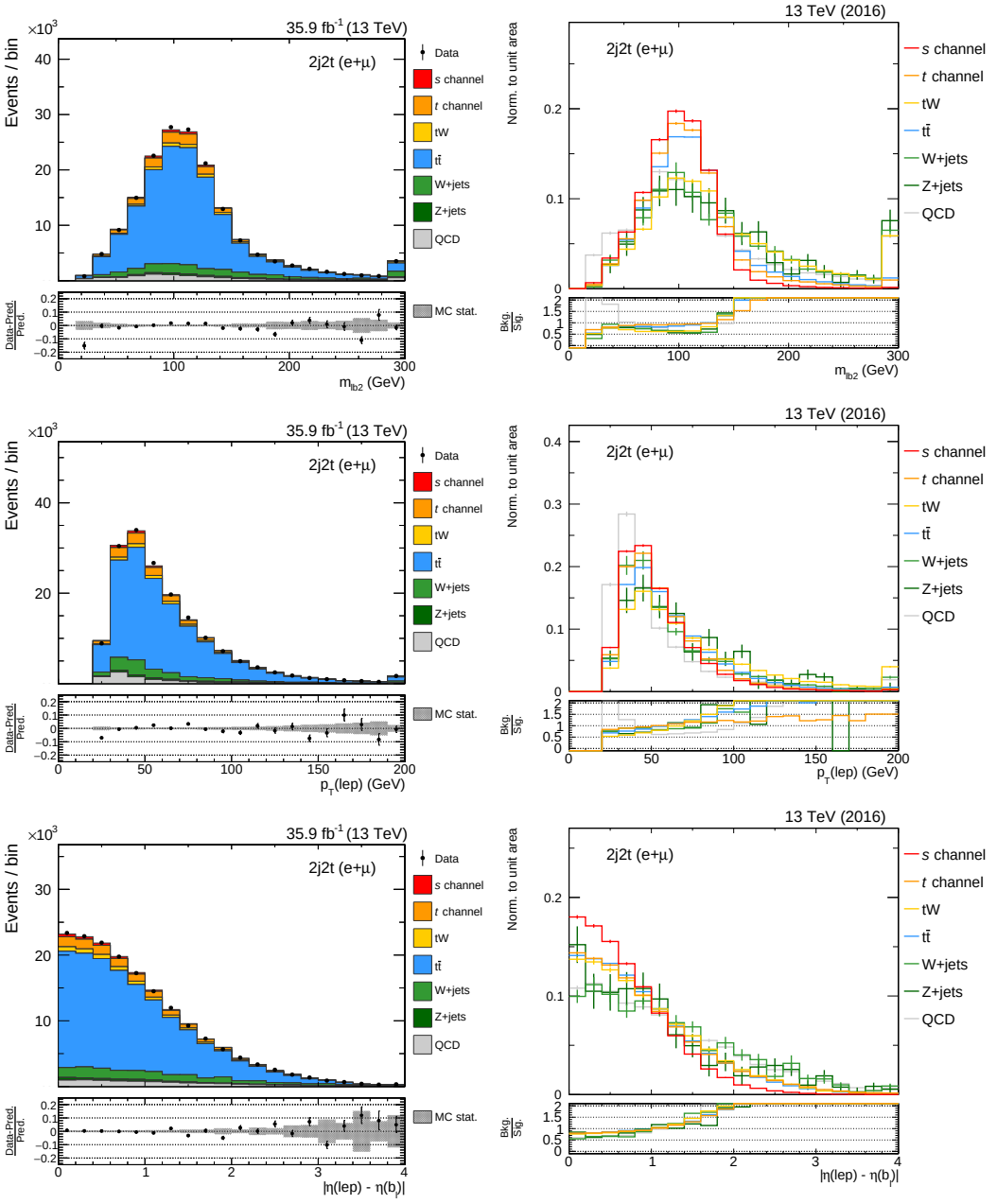

Figure 5.10: The three most important input variables for the DNN training in 2016. The distributions are shown in the 2j2t signal category for simulation and data (left) and for simulation only (right). In the distributions on the left-hand side, the prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation. The distributions on the right-hand side show the shapes of the input variables for each process.

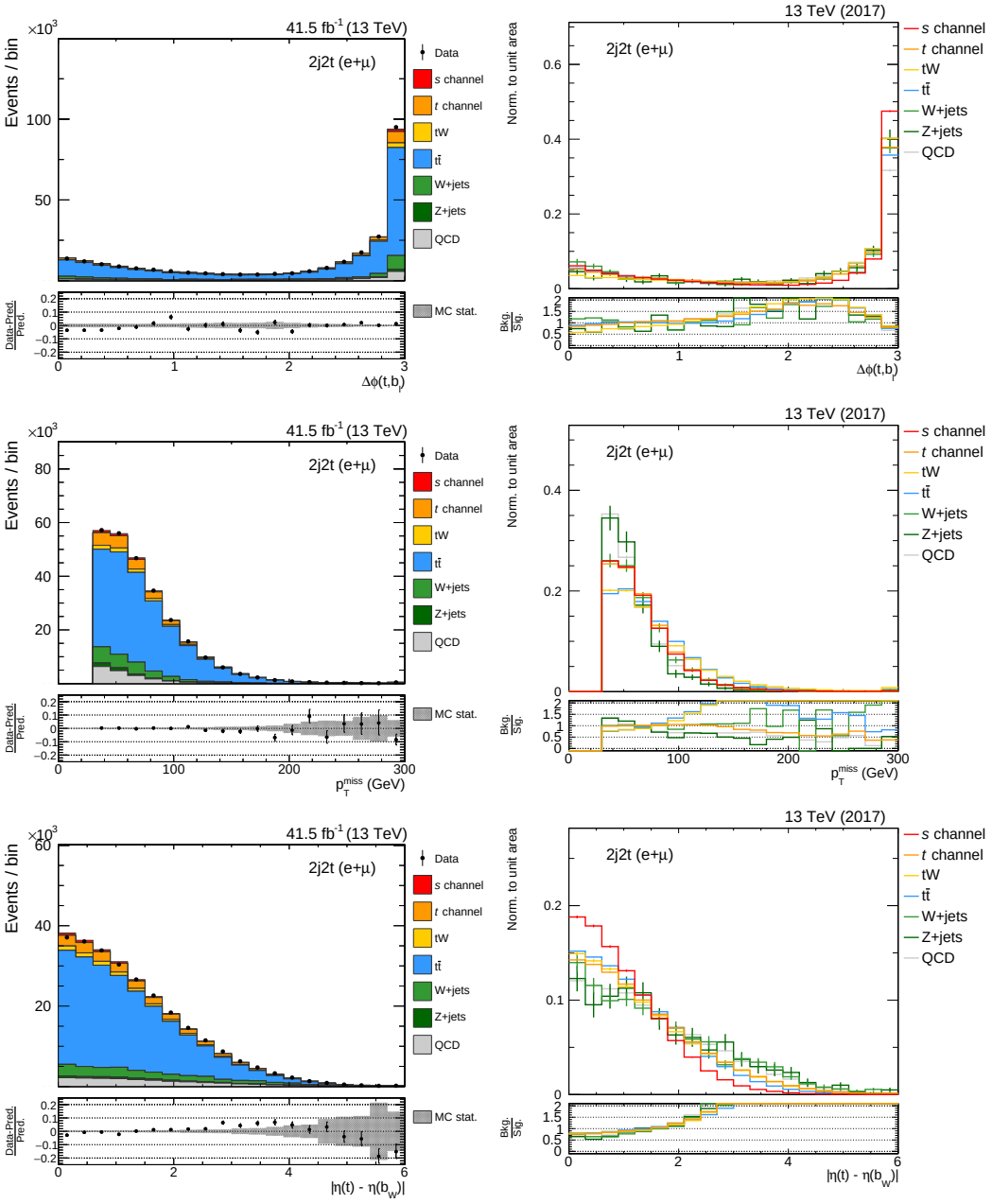

Figure 5.11: The three most important input variables for the DNN training in 2017. The distributions are shown in the 2j2t signal category for simulation and data (left) and for simulation only (right). In the distributions on the left-hand side, the prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation. The distributions on the right-hand side show the shapes of the input variables for each process.

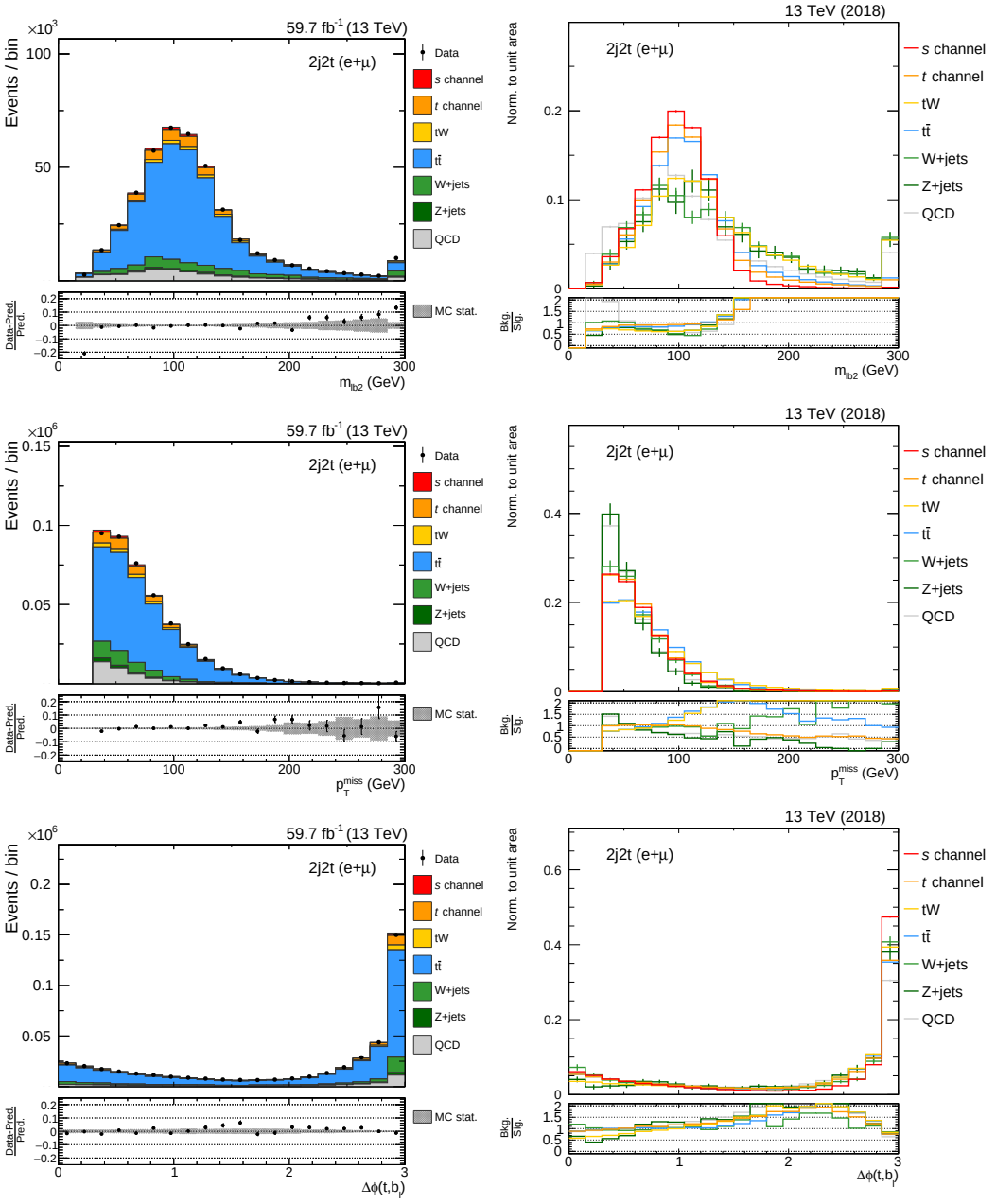

Figure 5.12: The three most important input variables for the DNN training in 2018. The distributions are shown in the 2j2t signal category for simulation and data (left) and for simulation only (right). In the distributions on the left-hand side, the prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation. The distributions on the right-hand side show the shapes of the input variables for each process.

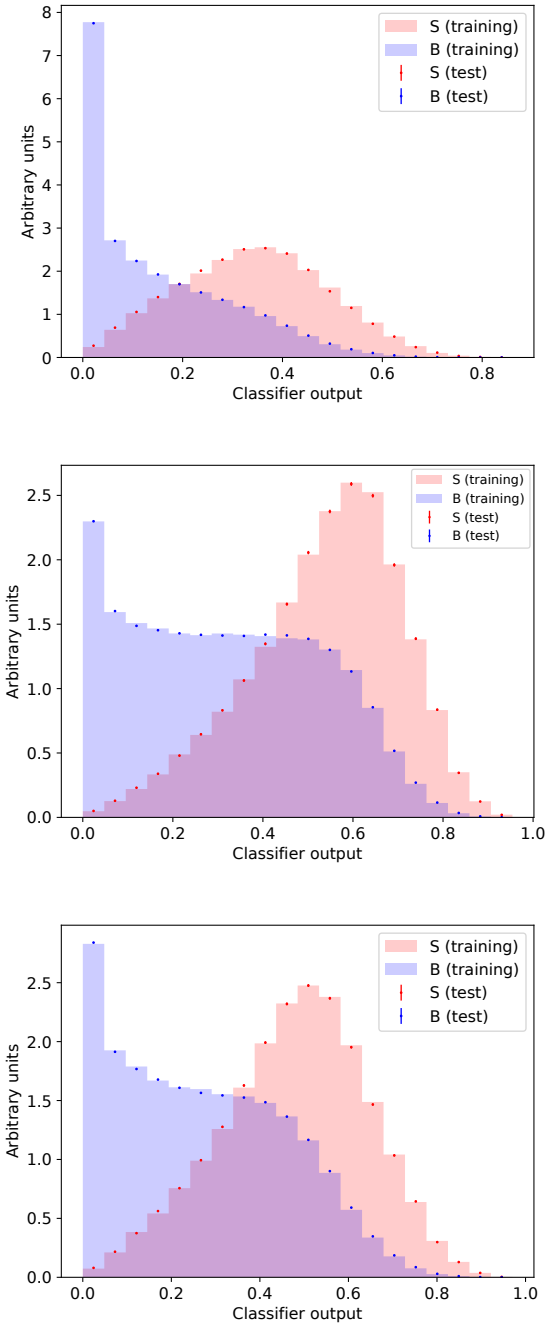

Figure 5.13: Result of the DNN training. The DNN output distributions of the signal and background training samples are compared to the corresponding test samples for the 2016 (top), 2017 (center), and 2018 (bottom) analysis. No overtraining is detected.

5.8 Signal extraction

The significance of s-channel single top quark production is extracted by performing a maximum likelihood fit, as described in Section 4.2.3, simultaneously on the DNN output distributions of all event categories. This means that four different DNN output distributions, corresponding to the signal category and the three control categories, are considered per year of data taking. In total, twelve DNN output distributions from three different years are included in the final fit for signal extraction. The DNN output distributions are shown in Fig. 5.14 for the 2j2t signal category and in Figs. E.11, E.12, and E.13 of Appendix E for the 2j1t, 3j1t, and 3j2t control categories. The binning of the DNN output distributions has been chosen such that no bins without any or very small background contribution exist, while still being sensitive to the difference in the shapes for signal and background events. This choice guarantees that the performed maximum likelihood fit is not too sensitive to bins with high statistical fluctuations of the background. The displayed uncertainty band in the DNN output distributions comprises not only the statistical uncertainty, but also all sources of systematic uncertainty that are considered in the fit. These distributions are referred to as *prefit* distributions. In the fit, the magnitude of each systematic uncertainty source is computed, which is later presented in Section 5.10. In general, two different significances are quoted: besides the observed significance computed by a fit to measured data, the expected significance is determined using an Asimov dataset [157]. The Asimov dataset is constructed from all simulated background processes and the signal process, i. e., it corresponds to the expectation of the signal-plus-background hypothesis. In addition, the measured cross section for s-channel single top quark production is determined from the fit.

5.9 Systematic uncertainties

Different uncertainty sources are taken into account in the search for s-channel single top quark production, which can be grouped into experimental and theoretical uncertainties. All uncertainty sources are incorporated as nuisance parameters (see Section 4.2.2) in the fitting procedure for signal extraction, either as a separate parameter per process or as a common parameter for several processes. As three years of data taking are analyzed and the same uncertainty sources are taken into account for each year, a dedicated model needs to be defined for the treatment of correlations among the systematic uncertainties.

5.9.1 Experimental uncertainties

Luminosity

As the integrated luminosities can only be determined within a certain uncertainty, a normalization uncertainty is applied on all simulated samples to consider this effect. The uncertainties in the integrated luminosities are measured to be 2.5% [198], 2.3% [199], and 2.5% [200] for the 2016, 2017, and 2018 datasets.

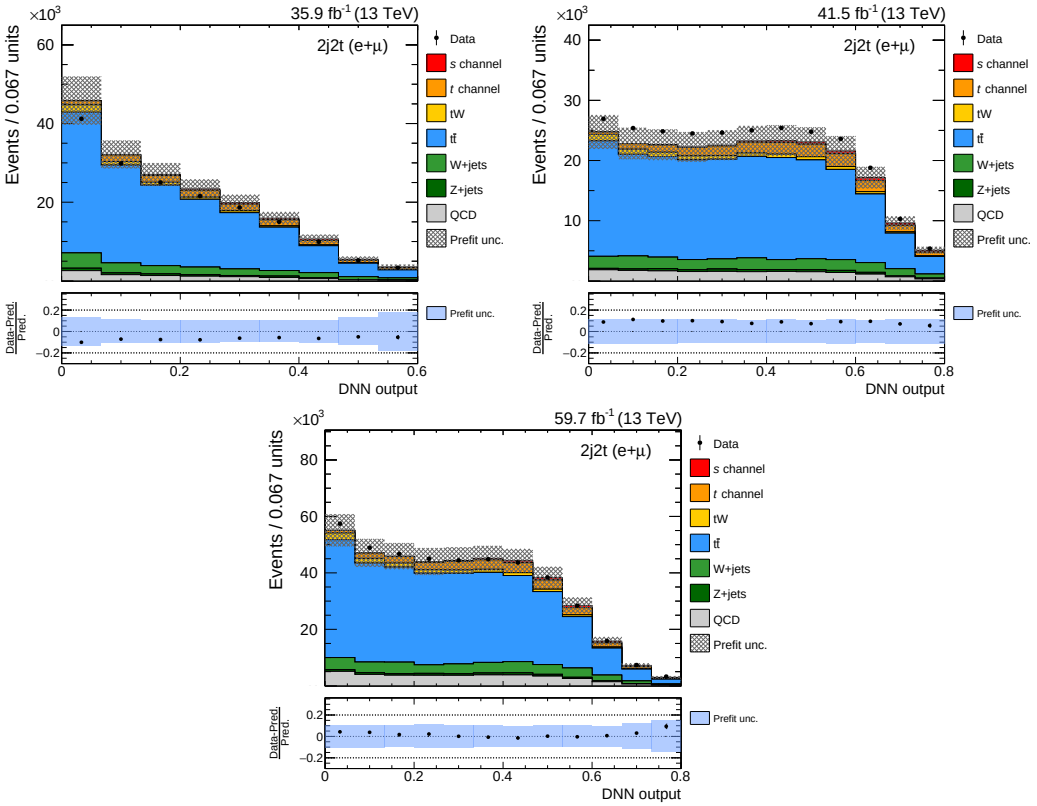

Figure 5.14: Prefit distributions of the DNN output values in the 2j2t signal category. The distributions are shown for the 2016 (top left), 2017 (top right), and 2018 analysis (bottom). The signal is depicted on top of the histogram. The hatched area in the main panel and the blue uncertainty band in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

Pileup

An uncertainty in the pileup reweighting procedure, which corrects the number of pileup interactions as described in Section 5.5.1, is applied by shifting the nominal inelastic proton-proton cross section by ±4.6% as recommended by the Luminosity POG [177]. This value covers both the uncertainty in the reweighting method [201] and in the measurement of the total inelastic proton-proton cross section [178].

Lepton efficiencies

For each computed identification, reconstruction, isolation, and trigger efficiency correction of selected muons and electrons, a dedicated uncertainty is applied (see Section 5.5.2). It is composed of the statistical uncertainty in the determination of the corresponding scale factors and the systematic uncertainty accounting for the different event topology used for the calculation of the efficiency corrections. The uncertainty is evaluated by applying event weights, which are varied up and down by the computed value and alter the shape of the nominal DNN output distribution.

b tagging efficiency

The correction of the b tagging efficiency applied to simulated events (see Section 5.5.3) leads to a systematic uncertainty that needs to be considered in this analysis. Each centrally available b tagging scale factor is computed within a certain uncertainty [191, 192, 193], which is employed to determine event weights that are shifted up and down by this value to evaluate this uncertainty source.

Jet energy scale

As described in Section 3.2.6, different jet energy corrections are applied on simulated events, which affect the p_{T} of all jets and therefore influence the jet energy scale (JES). Each uncertainty that arises from the computation of these corrections is taken into account using templates of the DNN output distribution for each process, which are shifted up and down by the determined uncertainty value. In total, 26 different JES uncertainty sources are derived, which represent the statistical and systematic component of the determined corrections for physics objects like light-flavored and heavy-flavored jets, and for effects caused by, e. g., pileup interactions and final-state radiation [202].

Jet energy resolution

Corrections are derived for jets in simulation, which smear the jet energy resolution (JER) to match the distribution observed in data [203]. The uncertainties associated to these corrections are taken into account by dedicated samples that provide templates of the DNN output distribution shifted up and down by the computed JER uncertainty for each process.

Unclustered energy

As different physics objects contribute to the reconstructed missing transverse momentum, uncertainties corresponding to the energy resolution of each of these objects have to be considered in this analysis [204]. Similar to the JES and JER uncertainties described above, the analysis is reiterated by employing templates of the DNN output distribution that are shifted up and down by the estimated uncertainty in the unclustered energy to take this uncertainty source into account.

L1 ECAL prefiring

As described in Section 5.5.4, a significant number of the high-η L1 ECAL trigger primitives was mistakenly associated to the previous bunch crossing during the 2016 and 2017 data-taking periods. The event weights accounting for the prefiring probability are shifted up and down according to their uncertainties. For each event weight, the maximum between the statistical uncertainty of the corresponding (η, p_{T}) bin and 20% of the prefiring probability is assigned as conservative uncertainty [194].

Limited size of simulated samples

Due to the finite number of simulated events of the samples used for the modeling of all background processes (see Section 5.3.1), an additional uncertainty has to be considered. The uncertainty is implemented as a bin-by-bin uncertainty and evaluated with the Barlow–Beeston method [155, 205]. The basic idea of this method is to assign one nuisance parameter to each bin of the DNN output distribution for each process in each event category, based on the number of events in the specific bin. This procedure introduces hundreds of additional nuisance parameters, which require high computing power for evaluation. Therefore, a modified approach is employed in this analysis, which assigns a single nuisance parameter per bin for all processes combined to decrease the total number of nuisance parameters. This approach is referred to as Barlow–Beeston light method. As a significantly higher number of signal events is simulated in general compared to the expected number in data, the statistical uncertainty caused by fluctuations between bins of the simulation sample is negligible with regard to the statistical uncertainty in data. Therefore, the Barlow–Beeston light method is only applied to the background processes.

5.9.2 Theoretical uncertainties

PDF and α_{s}

Uncertainties due to the choice of the PDF and the value for α_{s} are taken into account. For each simulated process, a normalization uncertainty according to the predicted PDF and α_{s} uncertainty is applied. A rate uncertainty of 3% is applied to t-channel single top quark production [30]. In case of the tW and $\mathrm{t\bar{t}}$ processes, an uncertainty of 5% [28, 30] is assigned. Rate uncertainties of 4% and 1% [206] are considered for W and Z boson production in association with jets, respectively. The applied normalization uncertainties are preliminary, they will be replaced by dedicated uncertainties that describe the effect

of the choice of the PDF and α_s value on the shape of the DNN output distribution. These uncertainties will be obtained by reweighting the DNN output distribution to account for 100 different shape variations of the nominal PDF and two variations of α_s, all derived from the eigenvector variations of the PDF set NNPDF3.0 NLO [83] for W/Z+jets production in 2016 and NNPDF3.1 NNLO [84] for all other simulation samples. The final uncertainty will be determined by calculating the full envelope of the up and down variations for each process, following the PDF4LHC recommendations [207, 208, 209].

Renormalization and factorization scales

The uncertainty due to the choice of the renormalization and factorization scales μ_R and μ_F is derived by employing the LHE reweighting procedure [210]. With this approach, shifted templates of the DNN output distribution are obtained for each process, corresponding to different combinations of double and half the nominal μ_R and μ_F values. The envelope of these templates is taken as the final template, omitting variations in which the two scales are shifted in opposite directions.

Matching of parton shower and matrix element

The matching of the parton shower generated by PYTHIA to the matrix element calculated by POWHEG is controlled by a damping function, which depends on the parameter h_{damp} and which in addition regulates the high-p_T radiation of partons [170]. As this can lead to additional jets that pass the selection requirements of this analysis, the effect of the damping function needs to be considered as uncertainty source. The uncertainty is evaluated for the $t\bar{t}$ process using dedicated simulation samples that have been generated with a variation of h_{damp} according to the uncertainties of the nominal value, $1.379^{+0.926}_{-0.5052} \cdot m_t$ [170, 211, 212], with m_t set to 172.5 GeV. For single top quark processes, this uncertainty source is found to be negligible.

Initial-state and final-state radiation

The impact on the modeling of the initial-state (ISR) and final-state radiation (FSR) caused by the choice of the α_s value for the parton shower is taken into account by shifted event weights. These event weights correspond to doubled and halved probabilities for additional gluon radiation in the initial and final state and are provided by PYTHIA for all simulated single top quark processes and $t\bar{t}$ production in 2016 and 2017. In 2018, these event weights are available for all simulated processes used in the analysis.

Underlying event

The underlying event and multiparton interactions, i. e., interactions between constituents stemming from different colliding partons, of the $t\bar{t}$ process are modeled by the generator tune CP5. Uncertainties connected to the tune settings are evaluated using dedicated $t\bar{t}$ simulation samples, which provide varied settings for multiparton interactions and for color string formations between final partons from independent hard scatterings, referred to as color reconnection [170].

Table 5.12: Rate uncertainties for the normalization of the 2016 W/Z+jets simulation samples with regard to different tunes. The uncertainties are calculated from the difference in the selection efficiency of the *s*-channel single top quark process when switching from tune CUETP8M1 to CP5. Rate uncertainties slightly higher than the calculated efficiency differences are applied to be conservative.

Event category	*s*-channel efficiency change (%)	W/Z+jets rate uncertainty (%)
2j1t	+1.4	±2.0
2j2t	+12.0	±12.0
3j1t	-8.5	±9.0
3j2t	+0.4	±1.0

W/Z+jets underlying event normalization

As described in Section 5.3.1, the simulation samples for W/Z+jets production used in 2016 are generated with tune CUETP8M1, whereas all other simulation samples are generated with the latest tune CP5. Differences in the modeling of the underlying event need to be considered when using a mixture of simulation samples with different tunes. Under the assumption that the yields of the W/Z+jets background processes change in the same way as for the signal process, conservative rate uncertainties are derived for the W/Z+jets normalization based on the change in the signal selection efficiencies when switching to the new tune. As the shape of the signal process is insensitive with regard to the tune switch, no shape uncertainties need to be derived. The relative change of the selection efficiency of *s*-channel single top quark production for each event category and the corresponding W/Z+jets rate uncertainties are listed in Table 5.12.

Top quark p_T

According to differential measurements of the top quark p_T distribution in $t\bar{t}$ events, the predicted p_T spectrum is shifted towards higher values compared to the observed spectrum [213]. As it is not recommended by Top Quark Physics Analysis Group of the CMS Collaboration [214] to correct this effect by reweighting the top quark p_T distribution for the $t\bar{t}$ process, the mismodeling is taken into account by assigning an additional uncertainty. The uncertainty is assessed with a template in the DNN output distribution that is shifted by the measured difference between simulation and data [214]. As this effect is one-sided, no interpolation between up- and down-varied templates is performed for this uncertainty source.

Background normalization

For each background process, a normalization uncertainty is applied. In case of the *t*-channel single top quark and $t\bar{t}$ processes, the uncertainty is taken from theory prediction, as the former shows no mismodeling in the selected phase space and as the cross section

of the latter was determined by CMS measurements [215, 216] with an uncertainty of the same size as predicted. The uncertainty in the cross section for t-channel single top quark and $\mathrm{t\bar{t}}$ production is predicted to be 4% [31, 32] and 6% [29], respectively. A conservative uncertainty of 10% is assigned to the W/Z+jets processes, to cover possible effects caused by the selection of heavy-flavored jets. For the tW process, an uncertainty of 11%, taken from the most recent CMS measurement [159], is applied to account for systematic uncertainties arising from a possible overlap with the $\mathrm{t\bar{t}}$ process at NLO accuracy. As the QCD multijet contribution is estimated with a data-driven method, no dedicated uncertainty is applied. Instead, the normalization of this background process is included in the fit as a free-floating rate parameter.

5.9.3 Correlation model of systematic uncertainties

Because of the different physics processes and the different years of data taking considered in this thesis, it is inevitable to define a dedicated correlation model for the treatment of systematic uncertainty sources. The decision to either fully correlate, partially correlate or uncorrelate a specific source of systematic uncertainty is driven by physics motivation and by technical implementation. In the following, the chosen correlation model between different processes and different years are presented.

Correlation between processes

As experimental uncertainty sources affect each considered process in the same way, they are all treated as fully correlated, meaning that all processes share one common nuisance parameter per experimental uncertainty source. In case of theoretical uncertainties, different normalization uncertainties are assigned for each background process (see Section 5.9.2), leading to an uncorrelated treatment of this uncertainty source. For the uncertainty in the renormalization and factorization scales, a distinction is made between processes that are induced by electroweak (EWK) interactions, i. e., all three single top quark production modes and W/Z+jets production, and the processes induced by strong (QCD) interactions, i. e., $\mathrm{t\bar{t}}$ production. This results into two nuisance parameters for this uncertainty source, one taking care of EWK-induced and the other for QCD-induced processes. The same approach is employed for the correlation of the uncertainties in the initial-state radiation, as the parton shower approximates specific higher-order matrix elements. For instance, the matrix element describing additional gluon emission is different in single top quark processes compared to the $\mathrm{t\bar{t}}$ process [217]. As these effects do not occur in the final-state radiation, this uncertainty source is treated as fully correlated between all processes. Furthermore, the uncertainty caused by the choice of the PDF set and the α_s value is treated as correlated between all processes that have been generated with the same PDF set. No dedicated correlation model needs to be defined for the uncertainties in the matching of matrix element with parton shower, in the underlying event, and in the top quark p_T distribution, as these uncertainty sources are considered only for top quark pair production.

Correlation between years

When calculating the significance and extracting the cross section for s-channel single top quark production using the combined 2016, 2017, and 2018 datasets, a potential correlation between certain systematic uncertainty sources needs to be considered. In particular, experimental uncertainties comprise a systematic component, as the same method is used in all years of data taking, and a statistical component, because statistically independent samples per year are employed to estimate each uncertainty source.

In case of the experimental uncertainties, the following correlation model is chosen: The uncertainty in the luminosity is partially correlated by splitting the corresponding nuisance parameter into one uncorrelated component per year and components that are either correlated between all years or only between two of the three years [218]. The uncertainty in the number of pileup interactions is considered to be fully correlated between all data-taking periods [177]. The uncertainty in the muon and electron efficiencies is treated separately for each source of efficiency correction. In case of the muon identification efficiency, separate scale factors are provided by the MUO POG [181, 182, 183] to account for the systematic and statistical component of this uncertainty source, therefore allowing to treat the former as correlated and the latter as uncorrelated. As privately produced scale factors, which are derived with a different dataset per year, are employed to estimate the uncertainty in the muon isolation and trigger paths, these two sources are considered as uncorrelated. The uncertainties in the electron efficiency for identification and reconstruction are fully correlated, as the systematic component dominates the estimation of these efficiencies [187]. The electron trigger efficiency is determined using privately computed scale factors, based on different HLT paths per year, and is therefore treated as an uncorrelated uncertainty source. The uncertainty in the b tagging efficiency is considered as uncorrelated, as different datasets and working points of the b tagging algorithm per year are employed for efficiency calculation. For the 26 different JES uncertainty sources taken into account in this thesis, a dedicated correlation model is employed (see Appendix F), which treats sources with dominant systematic component as correlated and sources taking statistical uncertainties into account as uncorrelated [202]. In addition, the JER uncertainties and the uncertainties in the unclustered energy between the different years of data taking are uncorrelated. The uncertainty in the prefiring issue of the L1 ECAL trigger primitives, which only occurred in 2016 and 2017, is fully correlated between these two years. As different samples per year are employed in this thesis, the resulting uncertainty due to the limited size of these samples is considered as uncorrelated between all data-taking periods.

In principle, all theoretical uncertainty sources are treated as fully correlated among the different years, as the theory predictions are independent of the year of data taking. One exception is the uncertainty in the choice of the PDF and the α_s value for the associated production of W and Z bosons with jets, as they use a different PDF set in 2016 compared to the samples employed in 2017 and 2018 (see Section 5.3.1). Therefore, this uncertainty source is treated as uncorrelated between 2016 and the other two data-taking periods for these two processes.

Table 5.13: Correlation of the systematic uncertainties between different years of data taking. The systematic uncertainties are included as nuisance parameters in the fit, and are either considered as fully correlated, partially correlated, or uncorrelated between the years.

Uncertainty source	corr. 2016/2017 (%)	corr. 2016/2018 (%)	corr. 2017/2018 (%)
Experimental uncertainties			
Luminosity	≈20	≈30	≈30
Pileup	100	100	100
Muon eff. (ID, syst.)	100	100	100
Muon eff. (ID, stat.)	0	0	0
Muon eff. (isolation)	0	0	0
Muon eff. (trigger)	0	0	0
Electron eff. (ID)	100	100	100
Electron eff. (reconstruction)	100	100	100
Electron eff. (trigger)	0	0	0
b tagging efficiency	0	0	0
JES	see Appendix F	see Appendix F	see Appendix F
JER	0	0	0
Unclustered energy	0	0	0
L1 ECAL prefiring	100	–	–
Limited sample size	0	0	0
Theoretical uncertainties			
PDF+α_s (W/Z+jets)	0	0	100
PDF+α_s ($t\bar{t}$ and single top)	100	100	100
μ_R/μ_F scale (EWK-induced)	100	100	100
μ_R/μ_F scale (QCD-induced)	100	100	100
h_{damp}	100	100	100
ISR (EWK-induced)	100	100	100
ISR (QCD-induced)	100	100	100
FSR	100	100	100
Underlying event	100	100	100
Top quark p_T	100	100	100
Background normalization	100	100	100

A summary of the correlation model of all experimental and theoretical uncertainty sources is presented in Table 5.13.

5.9.4 Impact of systematic uncertainty sources

In the fit employed to extract the signal, the systematic uncertainties affect the calculated significance. The impact of each uncertainty source is computed as the relative change of the expected significance of *s*-channel single top quark production when not considering the systematic uncertainty source under study. These contributions are determined by calculating the significance for each different uncertainty source, with the nuisance parameters corresponding to the respective uncertainty source being fixed at its prefit value. The final impact is then determined as the relative change of the significance to the nom-

Table 5.14: Expected impact of systematic uncertainties for each year and for the combined 2016–2018 analysis. The experimental and theoretical uncertainty sources are sorted by descending impact on the combined 2016–2018 analysis. For simplification, the different uncertainties in the muon and electron efficiencies, as well as the different PDF+α_s uncertainties, are each summarized into one single group.

Uncertainty source	change (%), 2016	change (%), 2017	change (%), 2018	change (%), 2016–2018
Experimental uncertainties				
Limited sample size	49.5	36.9	74.7	62.7
Pileup	9.0	6.6	18.7	7.8
JES	2.1	8.5	5.8	5.6
b tagging efficiency	12.9	0.9	6.0	3.2
Lepton efficiencies	1.4	2.1	0.2	1.7
Unclustered energy	0.8	<0.1	0.7	0.6
Luminosity	0.6	0.2	<0.1	0.6
JER	0.7	<0.1	1.3	0.3
L1 ECAL prefiring	0.5	0.4	–	0.2
Theoretical uncertainties				
Background normalization	4.3	6.1	1.5	5.1
FSR	1.2	4.6	0.9	2.9
h_{damp}	<0.1	1.1	2.2	1.9
μ_R/μ_F scale (QCD-induced)	1.2	2.7	0.4	1.6
ISR (QCD-induced)	5.1	0.2	0.2	1.5
Underlying event	<0.1	<0.1	2.1	0.9
ISR (EWK-induced)	<0.1	0.4	1.3	0.6
PDF+α_s	0.5	<0.1	0.2	0.3
μ_R/μ_F scale (EWK-induced)	0.4	4.3	0.7	0.3
Top quark p_T	0.5	<0.1	<0.1	0.1

inal value. The impacts are determined for different uncertainty groups and are listed in Table 5.14 for each year of data taking and for the combined 2016–2018 analysis. In case of the uncertainties in the JES and in the limited sample size, all associated nuisance parameters are grouped into a single uncertainty source during the impact computation. The experimental uncertainty source with by far the largest impact on the significance of the combined 2016–2018 analysis is the uncertainty in the limited sample size. The normalization of the background processes is the theoretical uncertainty that affects the significance the most. This means that the analysis is sensitive to the number of events of the available simulation samples due to the small selection efficiency. Variations of the impact of several systematic uncertainties between the years of data-taking periods occur. These are caused by a different shape of the DNN output distribution per year, by potential differences in the estimation of systematic uncertainty sources between the years, and by potential cancellations of systematic impacts due to correlation effects.

5.10 Results

As described in Section 5.8, the observed and expected significances are determined by a maximum likelihood fit of the DNN output distributions of the 2j2t signal category and the 2j1t, 3j1t, and 3j2t control categories to data from all three years of data taking. The significance is computed separately for each year and for the combined 2016–2018 analysis, the values are listed in Table 5.15. For the combined 2016–2018 analysis, single top quark production in the s-channel is observed with a significance of 6.0 standard deviations, with 4.7 standard deviations expected. While the observed significance in 2017 data alone is not sufficient to claim evidence for s-channel single top quark production, evidence is found in the analyses of the 2016 and 2018 datasets independently. Starting from the expected significance of 2.3 standard deviations for the 2016 analysis, one naively expects an increase of the significance per year by a factor of $\sqrt{N}$, with N being the ratio between the luminosity of the dataset under study and the one of the 2016 dataset. With this formula, significances of 2.5, 3.0, and 4.5 standard deviations are expected for the 2017, 2018, and the combined 2016–2018 analysis. These values agree well with the ones shown in Table 5.15. Small differences can for instance arise from the CMS pixel detector, which has been upgraded before the start of the data taking in 2017 and leads to an improved b tagging of jets. In Fig 5.15, all DNN output distributions are combined by sorting the corresponding bins by their signal-to-background ratio. A significant excess of the measured data compared to the background-only hypothesis is observed for the bins with high signal-to-background ratio values, which can be described by the presence of s-channel single top quark production.

In addition, the measured cross section of the combined 2016–2018 analysis is determined from the fit used for the significance extraction. During the fit, the actual impact of each systematic uncertainty source listed in Section 5.9 is computed and the overall fit uncertainty is determined, which is depicted in the *postfit* DNN output distributions shown in Figs. 5.16, 5.17, and 5.18. The result of the maximum likelihood estimation is the signal strength μ, which is multiplied with the predicted value of $\sigma^{\text{theo.}}_{s\text{-ch.}} = 10.32^{+0.29}_{-0.24}\,(\text{scale}) \pm 0.27\,(\text{PDF}+\alpha_\text{s})\,\text{pb}$ to obtain the measured cross section. With a measured signal strength of $\mu = 1.42^{+0.24}_{-0.23}$, where the uncertainty comprises the statistical and systematic uncertainties, the measured cross section is

$$\sigma_{s\text{-ch.}} = 14.65^{+2.48}_{-2.37}\,(\text{stat.} + \text{syst.})\,\text{pb}, \tag{5.14}$$

which is in agreement with the SM prediction within two standard deviations.

The measured cross section is used to measure the absolute value of the CKM matrix element V_tb, which is given by

$$|f_\text{LV} V_\text{tb}| = \sqrt{\frac{\sigma_{s\text{-ch.}}}{\sigma^{\text{theo.}}_{s\text{-ch.}}}}, \tag{5.15}$$

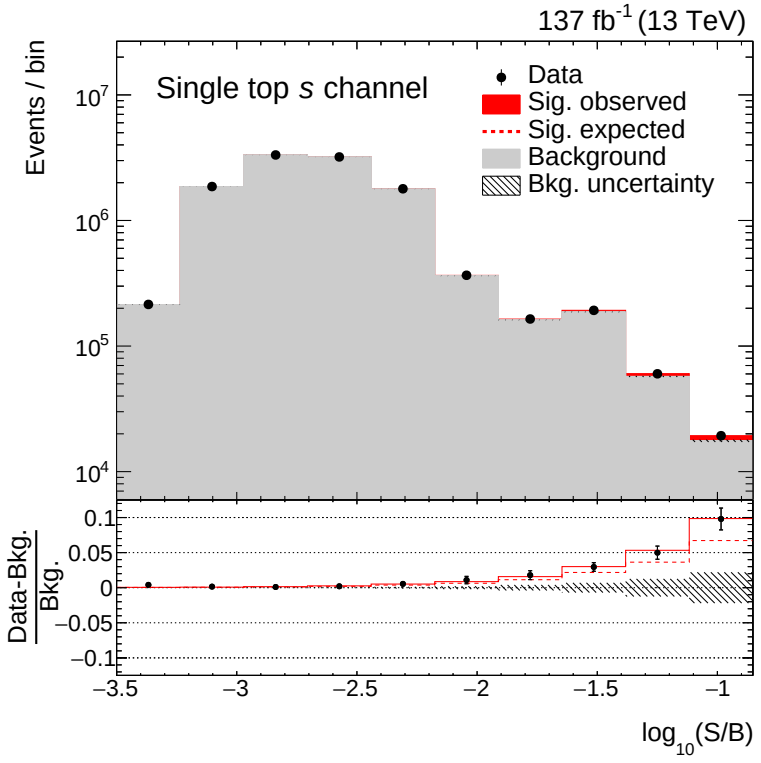

Figure 5.15: Comparison of prediction and data for the combined 2016–2018 analysis. The bins of all DNN output distributions included in the fit are sorted in ascending order of the logarithm of their predicted signal-to-background ratio. These values are evaluated for the background prediction (gray), for the expected and observed signal (red), and for the measured data. A significant excess is observed for high values, which can be described within two standard deviations by SM s-channel single top quark production and therefore excludes the background-only hypothesis in favor of the signal+background hypothesis.

Table 5.15: Observed and expected significances for s-channel single top quark production. The significance values are calculated separately per year of data taking and for the combined 2016–2018 analysis.

Year	significance (std. dev.)	
	obs.	exp.
2016	3.5	2.3
2017	1.9	2.5
2018	4.7	3.1
Combined	6.0	4.7

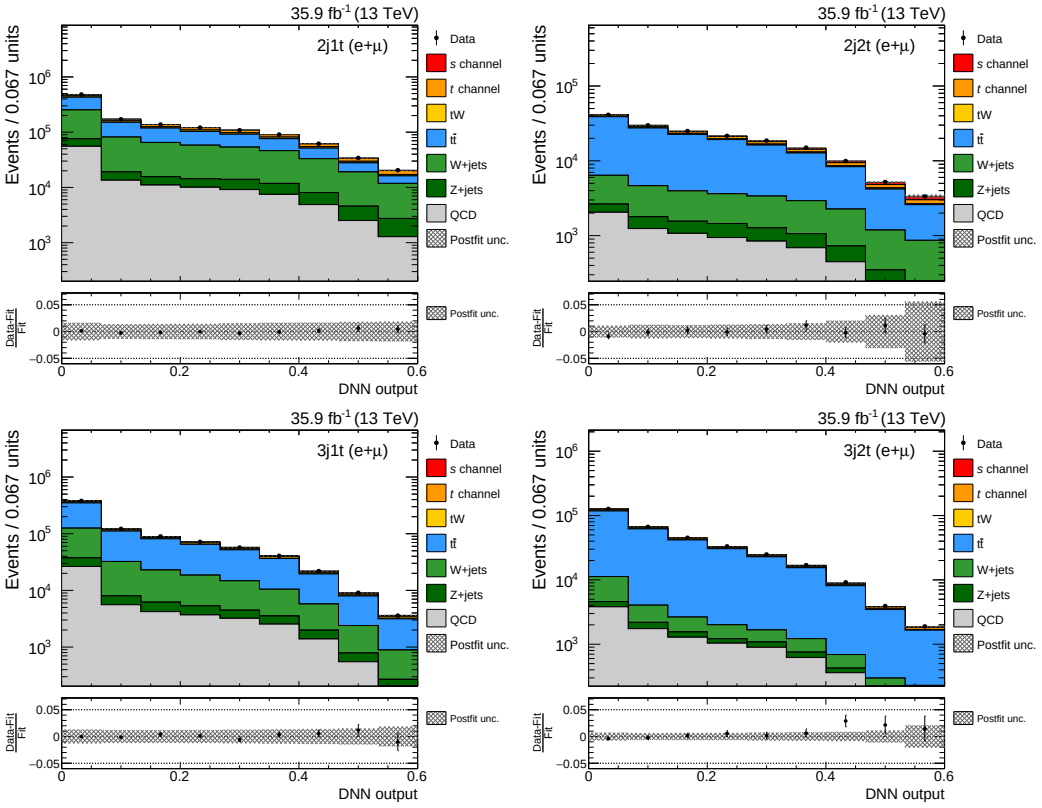

Figure 5.16: Postfit distributions of the DNN output values in the 2016 analysis. The distributions are shown for all event categories. The signal is depicted on top of the histogram. The hatched areas in the main and in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

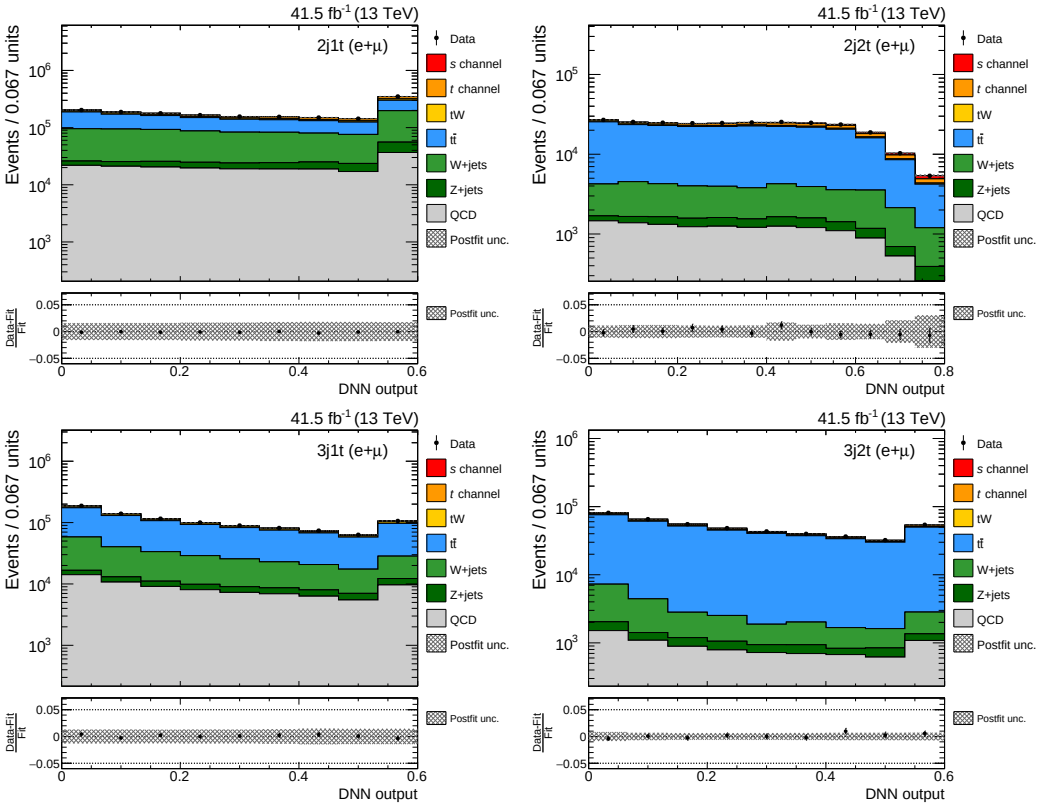

Figure 5.17: Postfit distributions of the DNN output values in the 2017 analysis. The distributions are shown for all event categories. The signal is depicted on top of the histogram. The hatched areas in the main and in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

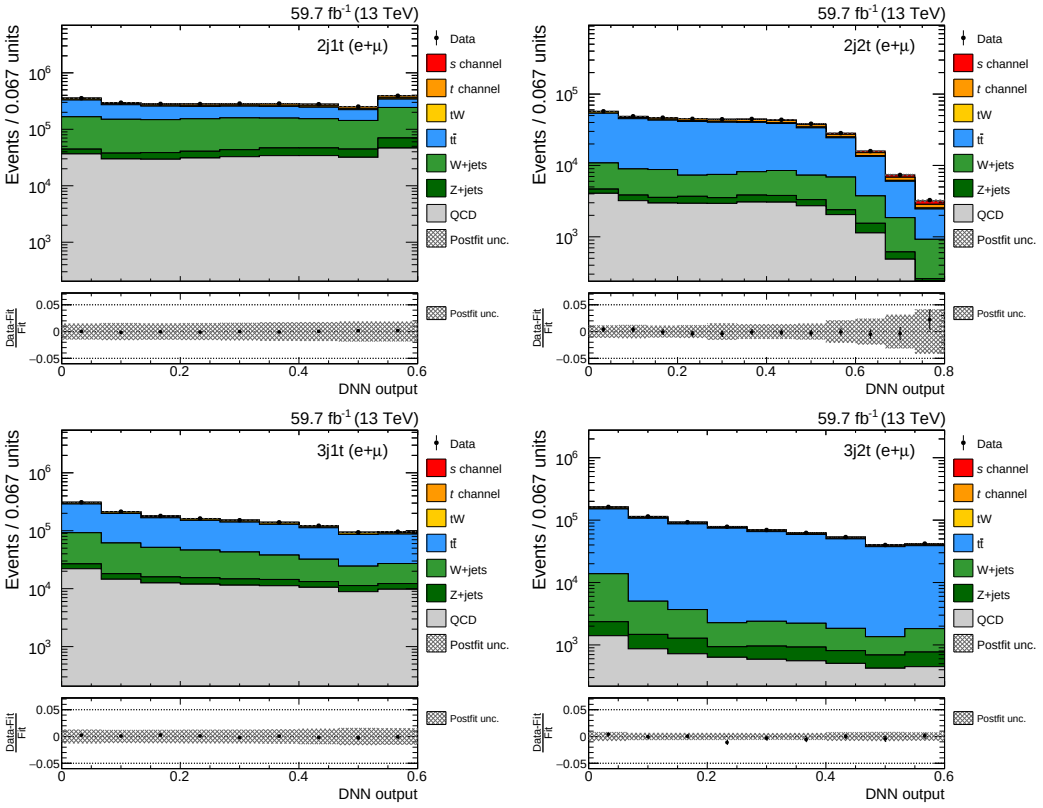

Figure 5.18: Postfit distributions of the DNN output values in the 2018 analysis. The distributions are shown for all event categories. The signal is depicted on top of the histogram. The hatched areas in the main and in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

under the assumption that the CKM matrix elements $|V_{td}|$ and $|V_{ts}|$ are negligible compared to $|V_{tb}|$. Here, $\sigma^{\text{theo.}}_{s\text{-ch.}}$ is the predicted cross section, calculated with $|V_{tb}| = 1$, and f_{LV} is a form factor, which considers the possible presence of anomalous Wtb couplings [5]. In case that the Wtb interaction is a left-handed weak coupling as predicted by the SM, f_{LV} is set to 1, whereas for couplings caused by physics beyond the SM, it is set to $f_{LV} \neq 1$. The measured cross section leads to an absolute CKM matrix element V_{tb} value of

$$|f_{LV}V_{tb}| = 1.19 \pm 0.09\,(\text{exp.}) \pm 0.02\,(\text{theo.}), \tag{5.16}$$

with the experimental uncertainty derived from the total uncertainty in the measured cross section, and the theoretical uncertainty computed from the uncertainty in the SM prediction. In addition, lower and upper exclusion limits on $|V_{tb}|$ of 0.99 and 1.36, respectively, are determined at 95% confidence level with the Feldman–Cousins unified approach [219], assuming the unitarity of the CKM matrix.

6. Conclusion

In this thesis, the search for s-channel single top quark production using data recorded by the CMS detector in 2016, 2017, and 2018 at a center-of-mass energy of 13 TeV has been presented. This process is the only single top quark production mode that has not been previously observed at the LHC. The presented analysis is the first s-channel single top quark analysis performed with data taken during Run 2 and reports an observed significance for the signal process of 6.0 standard deviations, with 4.7 standard deviations expected. The cross section and the CKM matrix element V_{tb} have been measured to be $\sigma_{s\text{-ch.}} = 14.65^{+2.48}_{-2.37}\,(\text{stat.} + \text{syst.})\,\text{pb}$ and $|f_{\mathrm{LV}}V_{\mathrm{tb}}| = 1.19 \pm 0.09\,(\text{exp.}) \pm 0.02\,(\text{theo.})$, which are within two standard deviations of the SM predictions.

The analysis of the CMS Collaboration performed with data taken during Run 1 of the LHC measured s-channel single top quark production with an observed significance of 2.5 standard deviations [162]. Compared to this previous analysis, a different analysis strategy has been employed by choosing a DNN for the event classification instead of BDTs. With this more elaborate analysis tool and with more data being analyzed, an increased sensitivity compared to the previous analysis has been achieved. In addition, the result of this analysis is improved compared to the Run 1 result of the ATLAS Collaboration, which reported an observed significance of 3.2 standard deviations [39].

The sensitivity of the presented analysis can be improved by employing simulated samples with a higher number of events, as the analysis is dominated by the uncertainty in the limited size of the employed background samples. In particular, the analysis would profit from a larger simulated sample for the W+jets background process, as this process has a relatively large cross section, but low event selection efficiency. By employing a multiclassification DNN, which is in addition to the signal and background separation capable of distinguishing different background contributions, the sensitivity could be further increased. With this approach, not only the dominant background process, $\mathrm{t\bar{t}}$, but also other single top quark production modes and W+jets production could be separated from s-channel single top quark production.

As the LHC is expected to collect an amount of data of around 350 fb^{-1} during Run 3 of the LHC [220], planned to start in May 2021, the precision of the measured s-channel single top quark production cross section will be further improved, especially if Run 3 is operated at a center-of-mass energy of 13 TeV. In this case, a combined analysis of Run 2 and Run 3 data would be possible, leading to a higher precision because of a larger number of measured events. Furthermore, the high-luminosity LHC (HL-LHC) [220, 221], which is designed to provide an instantaneous luminosity of 5 times the design value of the LHC and requires a major upgrade of the current LHC, aims to record around 3000 fb^{-1} of collision data at a center-of-mass energy of 14 TeV. With this amount of measured data, it will be possible to perform precision measurements of, e. g., differential cross sections, allowing to detect possible deviations from the SM predictions. Such small deviations for s-channel single top quark production can be caused by a non-SM mediator, for instance charged Higgs bosons [40] or heavy W′ bosons [41], or by contributions of anomalous flavor-changing top quark couplings [42].

In future circular and linear electron-positron colliders like the FCC-ee or the ILC, single top quarks can be produced, either via SM interactions or via potential flavor-changing neutral currents [222, 223]. As neither pileup interactions nor underlying events occur in electron-positron collisions, they provide much cleaner final states compared to the production via proton-proton collisions and therefore enable high-precision measurements, where small deviations could be more easily detected.

A. List of MC simulation samples and datasets

A.1 2016 analysis

Table A.1: Nominal simulation samples for the 2016 analysis. The samples are listed together with the generated number of events and the cross section times branching ratio (BR). If not stated otherwise, the cross section obtained from the generator is used. The fragment "RunIISummer16NanoAODv5-PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1…" and the postfix "/NANOAODSIM" are omitted everywhere.

Sample name	events ($\times 10^6$)	cross section $\times$ BR (pb)
ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	9.8	3.36 (NLO [14, 30])
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	31.8	136.02 (NLO [30])
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	17.8	80.95 (NLO [30])
ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	5.0	35.85 (NNLO [30])
ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	5.0	35.85 (NNLO [30])
TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8	65.9	88.2 (NNLO [14, 28])
TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	107.3	365.3 (NNLO [14, 28])
WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	259.8	61526.7 (NNLO [206])
WToLNu_0J_13TeV-amcatnloFXFX-pythia8	98.9	50131.98 (NNLO)
WToLNu_1J_13TeV-amcatnloFXFX-pythia8	41.4	8426.09 (NNLO)
WToLNu_2J_13TeV-amcatnloFXFX-pythia8	253.2	3172.96 (NNLO)
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	120.8	5765.4 (NNLO [206])

Table A.2: Systematically varied $t\bar{t}$ samples for the 2016 analysis. The samples are privately produced using official miniAODv3 samples and account for systematic uncertainties in the matching of parton shower to matrix element, controlled by the parameter h_{damp}, and in the underlying event. They are scaled with the same cross section times branching ratio as for the nominal samples. The fragment "mullerd-privNanoAODv5_2016-bf4f02b1b38ffdc426e6dbc4bfc0a03c/USER" is omitted everywhere.

Sample name	events ($\times 10^6$)
h_{damp}	
TTTo2L2Nu_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	14.9
TTTo2L2Nu_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	14.5
TTToSemiLeptonic_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	29.8
TTToSemiLeptonic_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	29.7
Underlying event	
TTTo2L2Nu_TuneCP5down_PSweights_13TeV-powheg-pythia8	11.1
TTTo2L2Nu_TuneCP5up_PSweights_13TeV-powheg-pythia8	13.9
TTToSemiLeptonic_TuneCP5down_PSweights_13TeV-powheg-pythia8	23.4
TTToSemiLeptonic_TuneCP5up_PSweights_13TeV-powheg-pythia8	21.7

Table A.3: Datasets for the 2016 analysis. The different 2016 data taking periods with the corresponding run ranges and integrated luminosities are listed. The fragments "SingleMuon" and "SingleElectron" are indicated with "…".

Period	run range	int. luminosity (fb^{-1})
/…/Run2016B_ver2-Nano1June2019_ver2-v1/NANOAOD	272007–275376	5.8
/…/Run2016C-Nano1June2019-v1/NANOAOD	275657–276283	2.6
/…/Run2016D-Nano1June2019-v1/NANOAOD	276315–276811	4.2
/…/Run2016E-Nano1June2019-v1/NANOAOD	276831–277420	4.0
/…/Run2016F-Nano1June2019-v1/NANOAOD	277772–278808	3.1
/…/Run2016G-Nano1June2019-v1/NANOAOD	278820–280385	7.5
/…/Run2016H-Nano1June2019-v1/NANOAOD	280919–284044	8.6
Total	272007–284044	35.9

A.2 2017 analysis

Table A.4: Nominal simulation samples for the 2017 analysis. The samples are listed together with the generated number of events and the cross section times branching ratio (BR). If not stated otherwise, the cross section obtained from the generator is used. The fragment "RunIIFall17NanoAODv5-PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1…" and the postfix "/NANOAODSIM" are omitted everywhere.

Sample name	events ($\times 10^6$)	cross section $\times$ BR (pb)
ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	9.8	3.36 (NLO[30])
ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	9.6	3.36 (NLO[30])
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	122.7	136.02 (NLO[30])
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	64.8	80.95 (NLO[30])
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	5.1	19.6 (NNLO[30])
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	5.6	19.6 (NNLO[30])
TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8	69.7	88.2 (NNLO[28])
TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	111.3	365.3 (NNLO[28])
WJetsToLNu_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	180.7	50131.98 (NNLO)
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	169.9	8426.09 (NNLO)
WJetsToLNu_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	98.3	3172.96 (NNLO)
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	186.0	5765.4 (NNLO [206])

Table A.5: Systematically varied $t\bar{t}$ samples for the 2017 analysis. The samples account for systematic uncertainties in the matching of parton shower to matrix element, controlled by the parameter h_{damp}, and in the underlying event. They are scaled with the same cross section times branching ratio as for the nominal samples. The fragment "RunIIFall17NanoAODv5-PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1…" and the postfix "/NANOAODSIM" are omitted everywhere.

Sample name	events ($\times 10^6$)
h_{damp}	
TTTo2L2Nu_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	5.5
TTTo2L2Nu_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	3.3
TTToSemiLeptonic_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	26.4
TTToSemiLeptonic_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	24.0
Underlying event	
TTTo2L2Nu_TuneCP5down_PSweights_13TeV-powheg-pythia8	5.5
TTTo2L2Nu_TuneCP5up_PSweights_13TeV-powheg-pythia8	5.5
TTToSemiLeptonic_TuneCP5down_PSweights_13TeV-powheg-pythia8	27.1
TTToSemiLeptonic_TuneCP5up_PSweights_13TeV-powheg-pythia8	20.1

Table A.6: Datasets for the 2017 analysis. The different 2017 data taking periods with the corresponding run ranges and integrated luminosities are listed. The fragments "SingleMuon" and "SingleElectron" are indicated with "…".

Period	run range	int. luminosity (fb^{-1})
/…/Run2017B-Nano1June2019-v1/NANOAOD	297046–299329	4.8
/…/Run2017C-Nano1June2019-v1/NANOAOD	299368–302029	9.7
/…/Run2017D-Nano1June2019-v1/NANOAOD	302030–303434	4.3
/…/Run2017E-Nano1June2019-v1/NANOAOD	303824–304797	9.3
/…/Run2017F-Nano1June2019-v1/NANOAOD	305040–306462	13.5
Total	297046–306462	41.5

A.3 2018 analysis

Table A.7: Nominal simulation samples for the 2018 analysis. The samples are listed together with the generated number of events and the cross section times branching ratio (BR). If not stated otherwise, the cross section obtained from the generator is used. The fragment "RunIIAutumn18NanoAODv5-Nano1June2019_102X_upgrade2018_realistic_v19-v1…" and the postfix "/NANOAODSIM" are omitted everywhere.

Sample name	events ($\times 10^6$)	cross section $\times$ BR (pb)
ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	20.0	3.36 (NLO[30])
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	154.3	136.02 (NLO[30])
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	79.1	80.95 (NLO[30])
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	7.6	19.6 (NNLO[30])
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	5.8	19.6 (NNLO[30])
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	64.3	88.2 (NNLO[28])
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	101.6	365.3 (NNLO[28])
WJetsToLNu_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	192.2	50131.98 (NNLO)
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	171.7	8426.09 (NNLO)
WJetsToLNu_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	98.3	3172.96 (NNLO)
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	182.2	5765.4 (NNLO [206])

Table A.8: Systematically varied $t\bar{t}$ samples for the 2018 analysis. The samples account for systematic uncertainties in the matching of parton shower to matrix element, controlled by the parameter h_{damp}, and in the underlying event. They are scaled with the same cross section times branching ratio as for the nominal samples. The fragment "RunIIAutumn18NanoAODv5-Nano1June2019_102X_upgrade2018_realistic_v19-v1…" and the postfix "/NANOAODSIM" are omitted everywhere.

Sample name	events ($\times 10^6$)
h_{damp}	
TTTo2L2Nu_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	5.5
TTTo2L2Nu_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	5.3
TTToSemiLeptonic_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	25.9
TTToSemiLeptonic_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	26.9
Underlying event	
TTTo2L2Nu_TuneCP5down_PSweights_13TeV-powheg-pythia8	5.0
TTTo2L2Nu_TuneCP5up_PSweights_13TeV-powheg-pythia8	5.4
TTToSemiLeptonic_TuneCP5down_PSweights_13TeV-powheg-pythia8	20.5
TTToSemiLeptonic_TuneCP5up_PSweights_13TeV-powheg-pythia8	26.9

Table A.9: Datasets for the 2018 analysis. The different 2018 data taking periods with the corresponding run ranges and integrated luminosities are listed. The fragments "SingleMuon" and "EGamma" are indicated with "…".

Period	run range	int. luminosity (fb^{-1})
/…/Run2018A-Nano1June2019-v1/NANOAOD	315252–316995	14.0
/…/Run2018B-Nano1June2019-v1/NANOAOD	317080–319310	7.1
/…/Run2018C-Nano1June2019-v1/NANOAOD	319337–320065	6.9
/…/Run2018D-Nano1June2019-v1/NANOAOD	320673–325175	31.9
Total	315252–325175	59.7

B. Distributions of reconstructed top quark mass

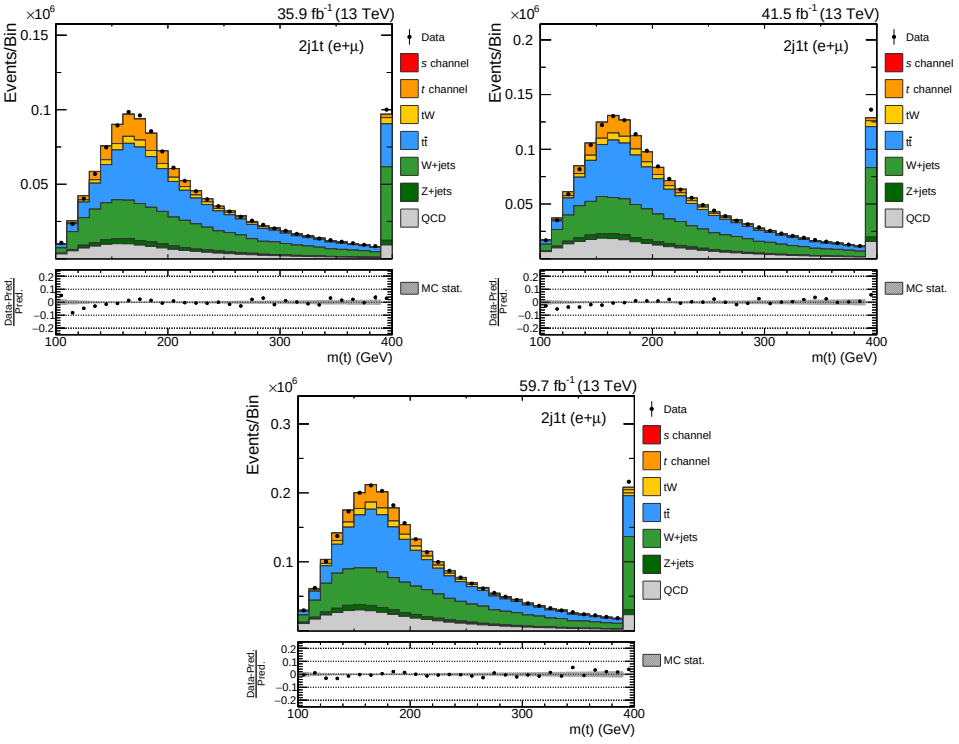

Figure B.1: Reconstructed top quark mass in the 2j1t control category. The distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) are shown. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

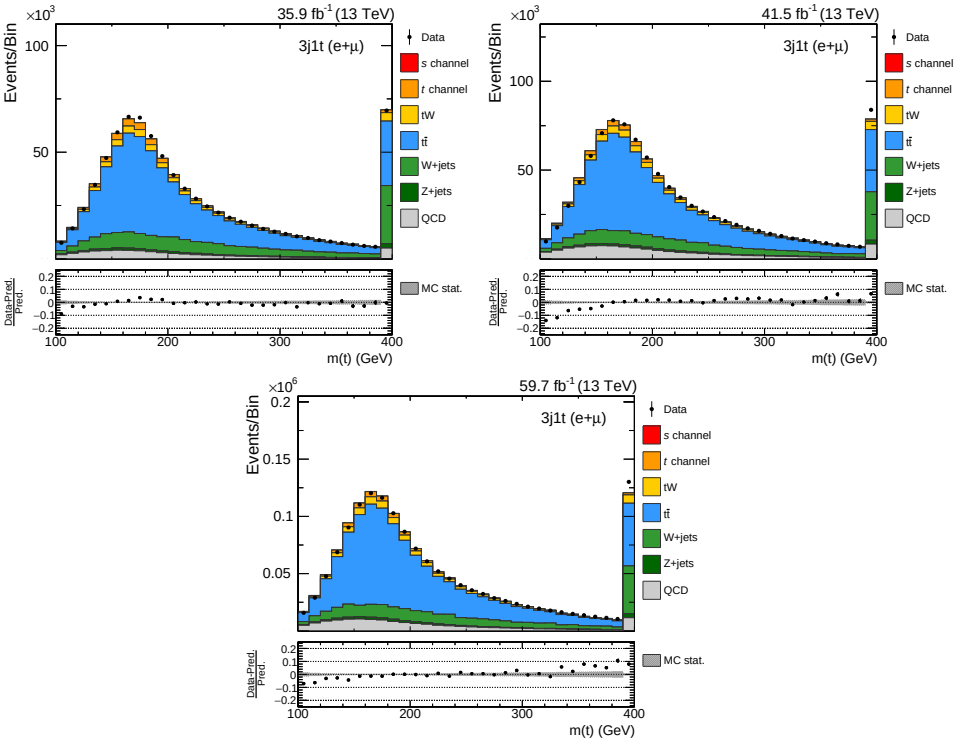

Figure B.2: Reconstructed top quark mass in the 3j1t control category. The distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) are shown. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

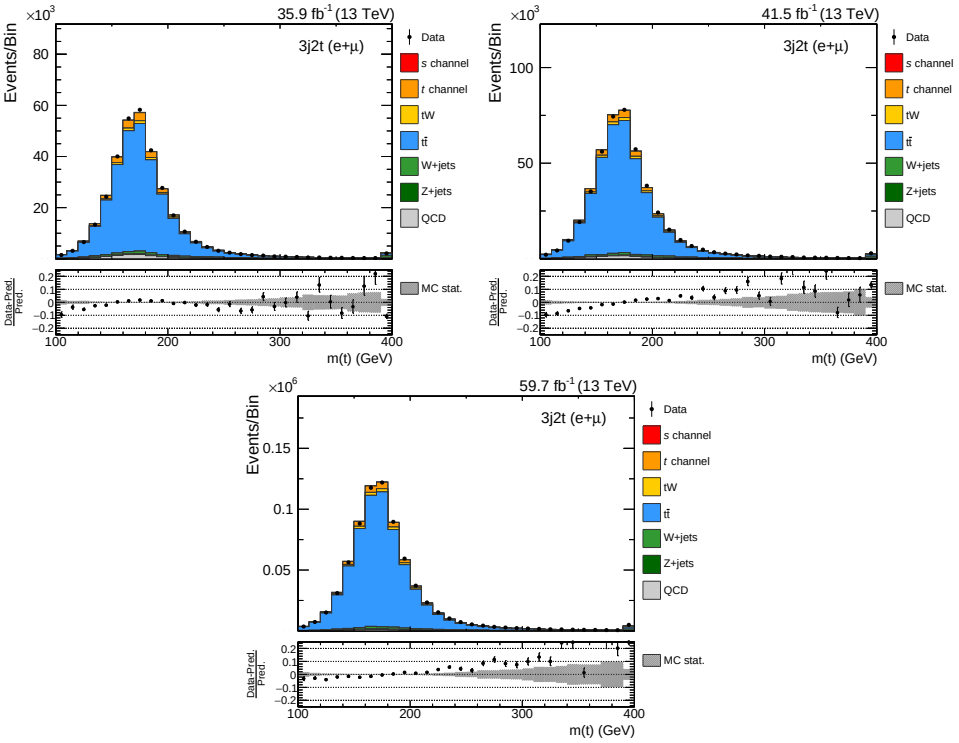

Figure B.3: Reconstructed top quark mass in the 3j2t control category. The distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) are shown. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

C. Calculation of electron trigger efficiencies

C.1 2017 analysis

Table C.1: Electron trigger scale factors in dependence of electron p_T and $|\eta|$ for the 2017 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($\lvert\eta_e\rvert$, $p_{T,e}$)	32–37 GeV	37–45 GeV	45–55 GeV	55–200 GeV
0–0.8	0.876 ± 0.022	0.933 ± 0.020	0.936 ± 0.019	0.943 ± 0.013
0.8–1.5	0.790 ± 0.025	0.912 ± 0.024	0.936 ± 0.024	0.946 ± 0.015
1.5–2.1	0.652 ± 0.027	0.847 ± 0.029	0.904 ± 0.030	0.920 ± 0.019

Table C.2: Electron trigger scale factors in dependence of electron and leading jet p_T for the 2017 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($p_{T,jet1}$, $p_{T,e}$)		electron			
		32–37 GeV	37–45 GeV	45–55 GeV	55–200 GeV
jet	40–50 GeV	0.771 ± 0.063	0.896 ± 0.058	0.931 ± 0.058	0.939 ± 0.035
	50–100 GeV	0.808 ± 0.020	0.908 ± 0.018	0.930 ± 0.018	0.940 ± 0.013
	100–200 GeV	0.803 ± 0.025	0.918 ± 0.023	0.930 ± 0.023	0.942 ± 0.015

Table C.3: Electron trigger scale factors in dependence of leading jet p_T and $|\eta|$ for the 2017 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($\lvert\eta_{\text{jet1}}\rvert$, $p_{T,\text{jet1}}$)	40–50 GeV	50–100 GeV	100–200 GeV
0–0.8	0.922 ± 0.037	0.918 ± 0.014	0.929 ± 0.020
0.8–1.5	0.900 ± 0.042	0.922 ± 0.016	0.922 ± 0.024
1.5–2.0	0.911 ± 0.064	0.913 ± 0.022	0.908 ± 0.034
2.0–2.4	0.888 ± 0.077	0.914 ± 0.031	0.923 ± 0.050

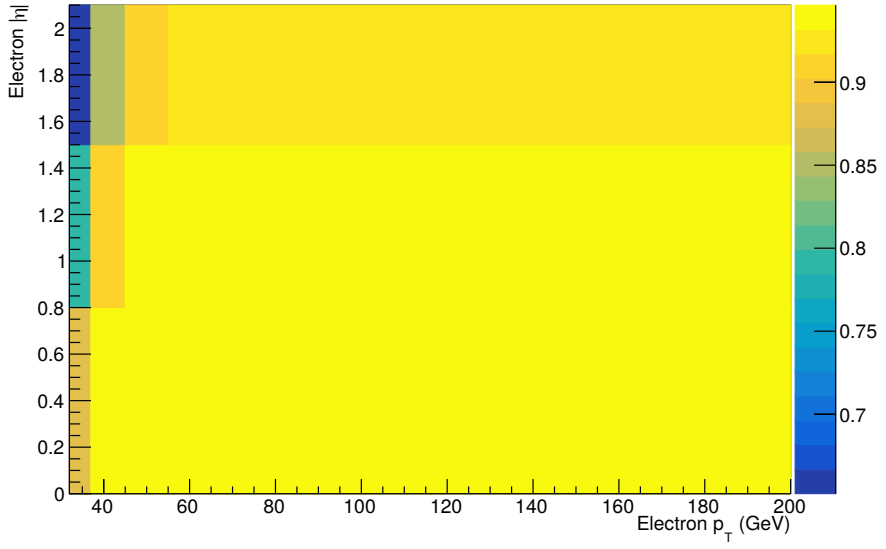

Figure C.1: Electron trigger scale factors in dependence of electron p_T and $|\eta|$. The scale factors are calculated using 2017 single muon data samples and simulated dileptonic $t\bar{t}$ events.

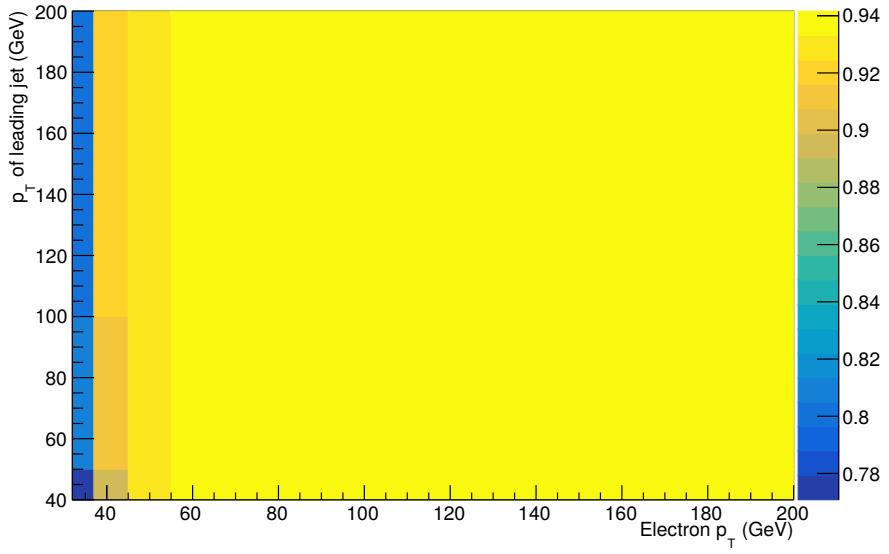

Figure C.2: Electron trigger scale factors in dependence of electron and leading jet p_T. The scale factors are calculated using 2017 single muon data samples and simulated dileptonic $t\bar{t}$ events.

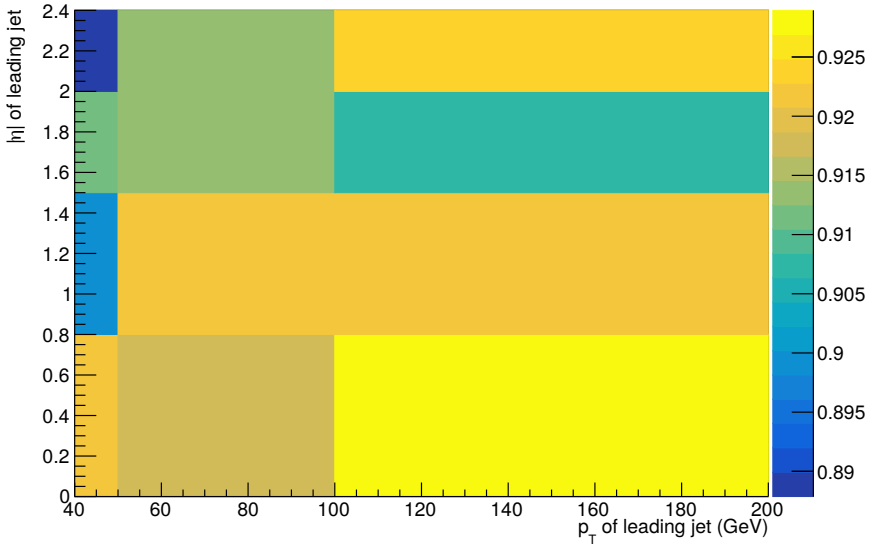

Figure C.3: Electron trigger scale factors in dependence of leading jet p_T and $|\eta|$. The scale factors are calculated using 2017 single muon data samples and simulated dileptonic $t\bar{t}$ events.

C.2 2018 analysis

Table C.4: Electron trigger scale factors in dependence of electron p_T and $|\eta|$ for the 2018 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($\lvert\eta_e\rvert$, $p_{T,e}$)	32–37 GeV	37–45 GeV	45–55 GeV	55–200 GeV
0–0.8	0.932 ± 0.020	0.953 ± 0.017	0.959 ± 0.017	0.967 ± 0.012
0.8–1.5	0.868 ± 0.023	0.952 ± 0.021	0.952 ± 0.021	0.968 ± 0.014
1.5–2.1	0.820 ± 0.027	0.912 ± 0.025	0.936 ± 0.026	0.948 ± 0.017

Table C.5: Electron trigger scale factors in dependence of electron and leading jet p_T for the 2018 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($p_{T,jet1}$, $p_{T,e}$)		electron			
		32–37 GeV	37–45 GeV	45–55 GeV	55–200 GeV
jet	40–50 GeV	0.899 ± 0.058	0.958 ± 0.050	0.972 ± 0.050	0.958 ± 0.029
	50–100 GeV	0.888 ± 0.018	0.945 ± 0.016	0.958 ± 0.016	0.965 ± 0.012
	100–200 GeV	0.892 ± 0.023	0.947 ± 0.020	0.954 ± 0.020	0.964 ± 0.013

Table C.6: Electron trigger scale factors in dependence of leading jet p_T and $|\eta|$ for the 2018 analysis. In addition to statistical uncertainties, a systematic uncertainty of 1% is applied.

SF($\lvert\eta_{jet1}\rvert$, $p_{T,jet1}$)	40–50 GeV	50–100 GeV	100–200 GeV
0–0.8	0.922 ± 0.037	0.918 ± 0.014	0.929 ± 0.020
0.8–1.5	0.900 ± 0.042	0.922 ± 0.016	0.922 ± 0.024
1.5–2.0	0.911 ± 0.064	0.913 ± 0.022	0.908 ± 0.034
2.0–2.4	0.888 ± 0.077	0.914 ± 0.031	0.923 ± 0.050

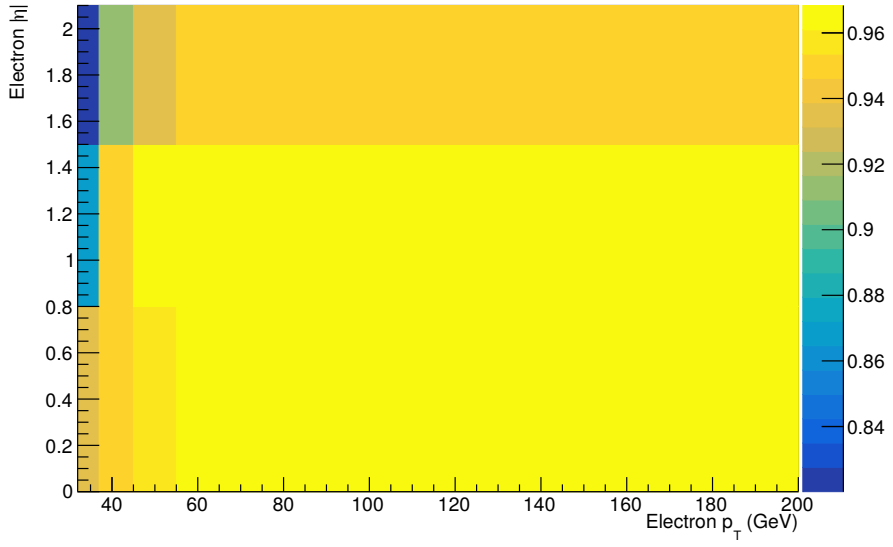

Figure C.4: Electron trigger scale factors in dependence of electron p_T and $|\eta|$. The scale factors are calculated using 2018 single muon data samples and simulated dileptonic $t\bar{t}$ events.

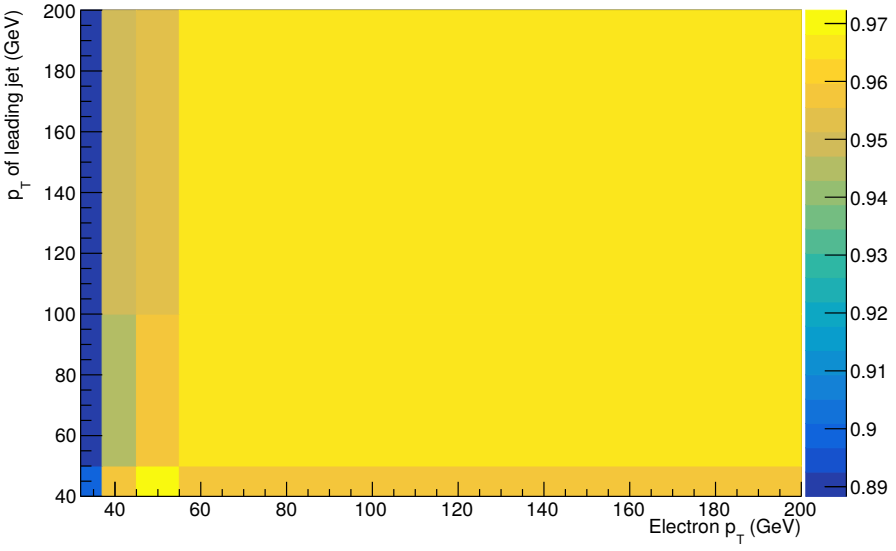

Figure C.5: Electron trigger scale factors in dependence of electron and leading jet p_T. The scale factors are calculated using 2018 single muon data samples and simulated dileptonic $t\bar{t}$ events.

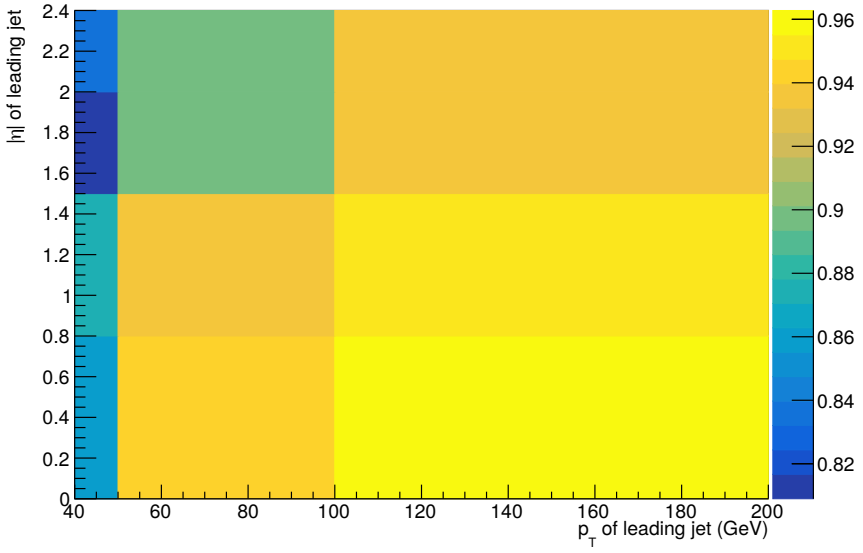

Figure C.6: Electron trigger scale factors in dependence of leading jet p_T and $|\eta|$. The scale factors are calculated using 2018 single muon data samples and simulated dileptonic $t\bar{t}$ events.

D. QCD multijet background estimation in the control categories

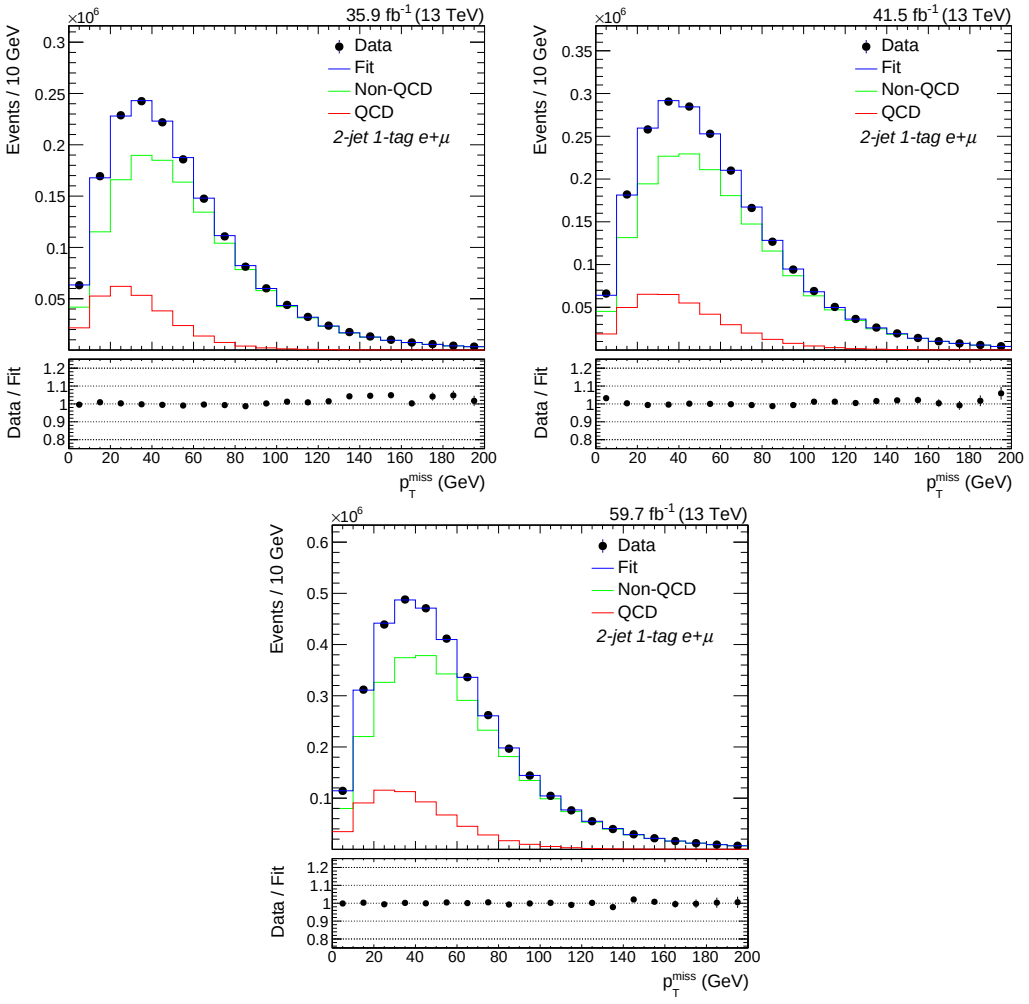

Figure D.1: Estimation of the QCD multijet background contribution in the 2j1t control category. The fit is performed to the $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) data. The error bars in the ratio panel correspond to statistical uncertainties.

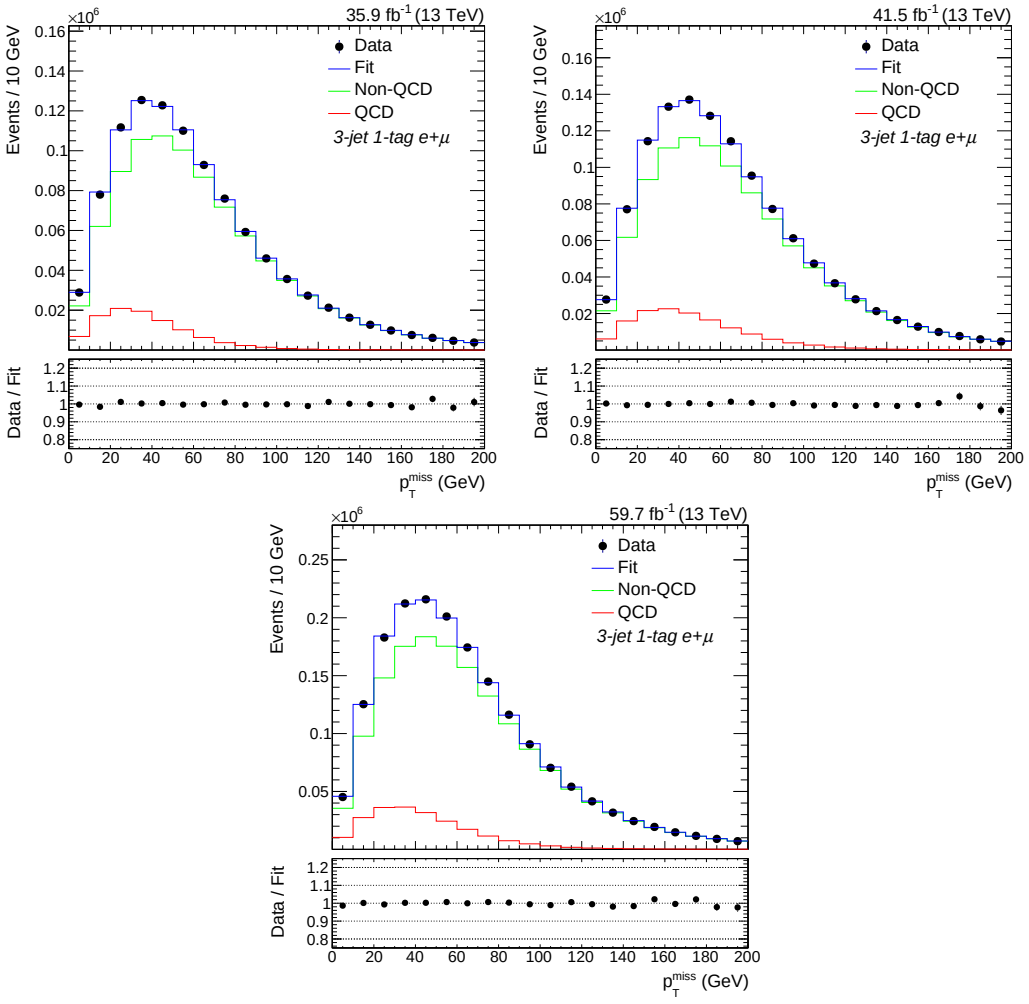

Figure D.2: Estimation of the QCD multijet background contribution in the 3j1t control category. The fit is performed to the $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) data. The error bars in the ratio panel correspond to statistical uncertainties.

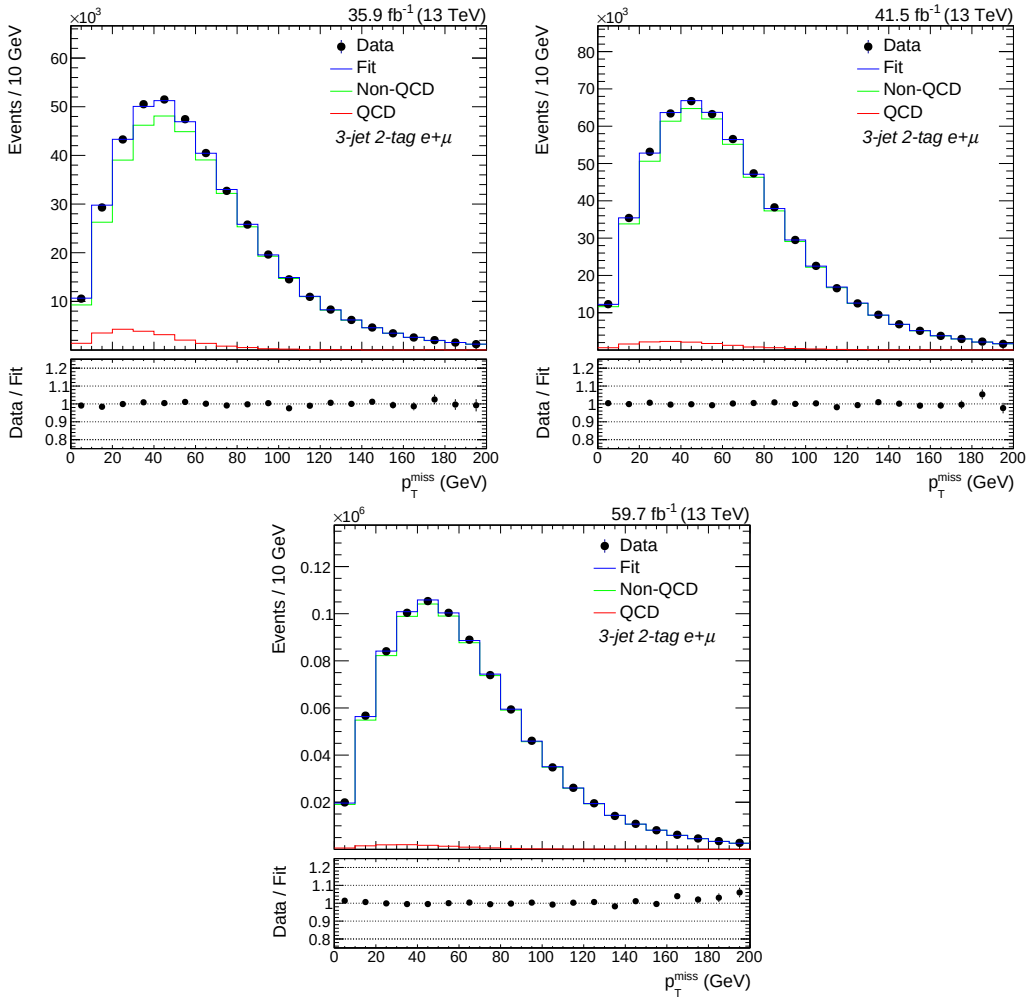

Figure D.3: Estimation of the QCD multijet background contribution in the 3j2t control category. The fit is performed to the $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions for 2016 (top left), 2017 (top right), and 2018 (bottom) data. The error bars in the ratio panel correspond to statistical uncertainties.

Table D.1: Estimation of the QCD multijet background contribution in the 2j1t control category. The event yields are obtained from the fit to the missing transverse momentum distribution and are displayed together with their statistical uncertainties for the full range and for $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.

Year	QCD yield (n_{QCD}) full range	QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	non-QCD yield ($n_{\text{non-QCD}}$) full range	non-QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV
2016	281 790 ± 540	145 460 ± 390	1 404 500 ± 2300	1 081 500 ± 2000
2017	377 720 ± 840	243 890 ± 670	1 814 600 ± 3400	1 443 500 ± 2900
2018	627 300 ± 1000	386 400 ± 810	2 946 100 ± 5400	2 319 800 ± 4700

Table D.2: Estimation of the QCD multijet background contribution in the 3j1t control category. The event yields are obtained from the fit to the missing transverse momentum distribution and are displayed together with their statistical uncertainties for the full range and for $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.

Year	QCD yield (n_{QCD}) full range	QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	non-QCD yield ($n_{\text{non-QCD}}$) full range	non-QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV
2016	105 260 ± 410	60 310 ± 310	907 500 ± 1200	733 800 ± 1100
2017	141 450 ± 660	97 840 ± 550	1 034 500 ± 1800	858 000 ± 1600
2018	216 330 ± 750	142 380 ± 610	1 607 700 ± 2600	1 326 400 ± 2300

Table D.3: Estimation of the QCD multijet background contribution in the 3j2t control category. The event yields are obtained from the fit to the missing transverse momentum distribution and are displayed together with their statistical uncertainties for the full range and for $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.

Year	QCD yield (n_{QCD}) full range	QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	non-QCD yield ($n_{\text{non-QCD}}$) full range	non-QCD yield $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV
2016	21 760 ± 520	12 590 ± 390	389 830 ± 480	315 260 ± 420
2017	14 610 ± 430	10 180 ± 350	542 040 ± 750	445 990 ± 680
2018	11 750 ± 250	7750 ± 200	864 900 ± 1000	708 760 ± 930

E. DNN for event classification

E.1 DNN training

Table E.1: Sum of all weights of first hidden layer per variable for the 2016 DNN training. The variables are sorted by the highest value for the sum of all weights obtained in the first hidden layer of the 2016 DNN.

Variable	∑weights	Variable	∑weights
$m_{\ell \mathrm{b}2}$	5.6	Fox Wolfram #3	−0.6
$p_{\mathrm{T}}(\ell)$	2.5	$q(\ell)$	−0.8
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	2.2	$\eta(\mathrm{b_l})$	−1.0
$p_{\mathrm{T}}(\mathrm{b_s})$	1.8	$\cos\theta^*$	−1.1
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	1.8	$p_{\mathrm{T}}^{\mathrm{miss}}$	−1.2
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	0.3	$m(\mathrm{t})$	−1.4
$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	0.3	$p_{\mathrm{T}}^{\mathrm{bb}}$	−2.2
$\eta(\mathrm{b_s})$	−0.3	H_{t}	−2.3
$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	−0.4	$p_{\mathrm{T}}(\mathrm{b_l})$	−4.5

Table E.2: Sum of all weights of first hidden layer per variable for the 2017 DNN training. The variables are sorted by the highest value for the sum of all weights obtained in the first hidden layer of the 2017 DNN.

Variable	∑weights	Variable	∑weights
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	1.9	$p_\mathrm{T}(\mathrm{b_s})$	−0.4
$p_\mathrm{T}^\mathrm{miss}$	0.8	$p_\mathrm{T}(\ell)$	−0.5
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	0.7	$\cos\theta^*$	−0.6
$m_{\ell\mathrm{b}2}$	0.6	$m(\mathrm{t})$	−0.8
$\eta(\mathrm{b_l})$	0.3	Fox Wolfram #3	−1.0
$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	0.1	$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	−1.5
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	0.1	H_t	−1.9
$q(\ell)$	0.0	p_T^bb	−3.2
$\eta(\mathrm{b_s})$	−0.2	$p_\mathrm{T}(\mathrm{b_l})$	−5.2

Table E.3: Sum of all weights of first hidden layer per variable for the 2018 DNN training. The variables are sorted by the highest value for the sum of all weights obtained in the first hidden layer of the 2018 DNN.

Variable	∑weights	Variable	∑weights
$m_{\ell\mathrm{b}2}$	2.8	$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	0.1
$p_\mathrm{T}^\mathrm{miss}$	2.6	$q(\ell)$	0.0
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	1.7	H_t	−0.2
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	1.7	$m(\mathrm{t})$	−0.7
$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	1.6	$\eta(\mathrm{b_l})$	−1.3
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	1.0	$\cos\theta^*$	−2.2
$p_\mathrm{T}(\ell)$	0.4	Fox Wolfram #3	−2.3
$p_\mathrm{T}(\mathrm{b_s})$	0.2	p_T^bb	−2.9
$\eta(\mathrm{b_s})$	0.2	$p_\mathrm{T}(\mathrm{b_l})$	−4.6

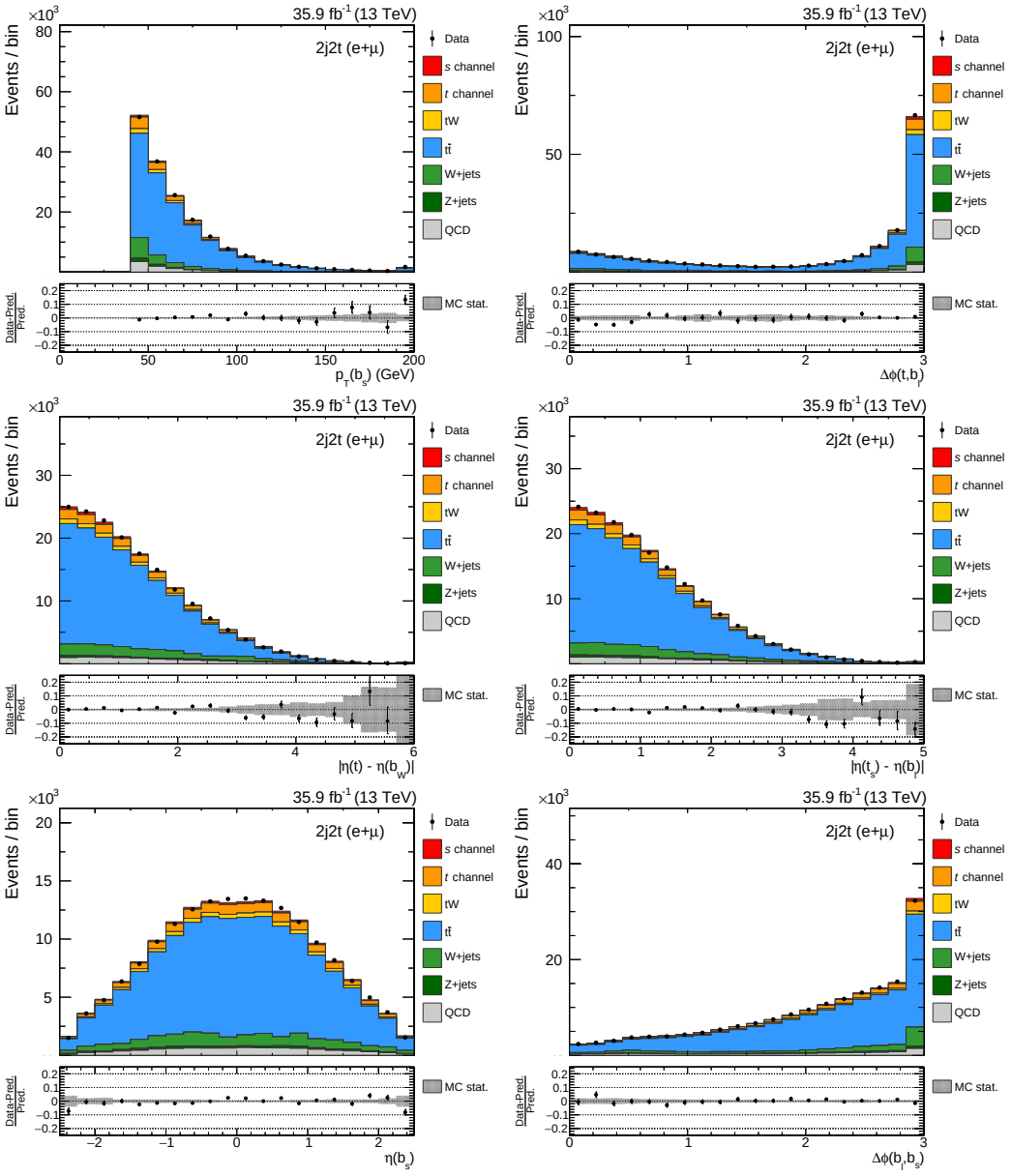

Figure E.1: Input variables ranked 4th to 9th for the DNN training in 2016. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

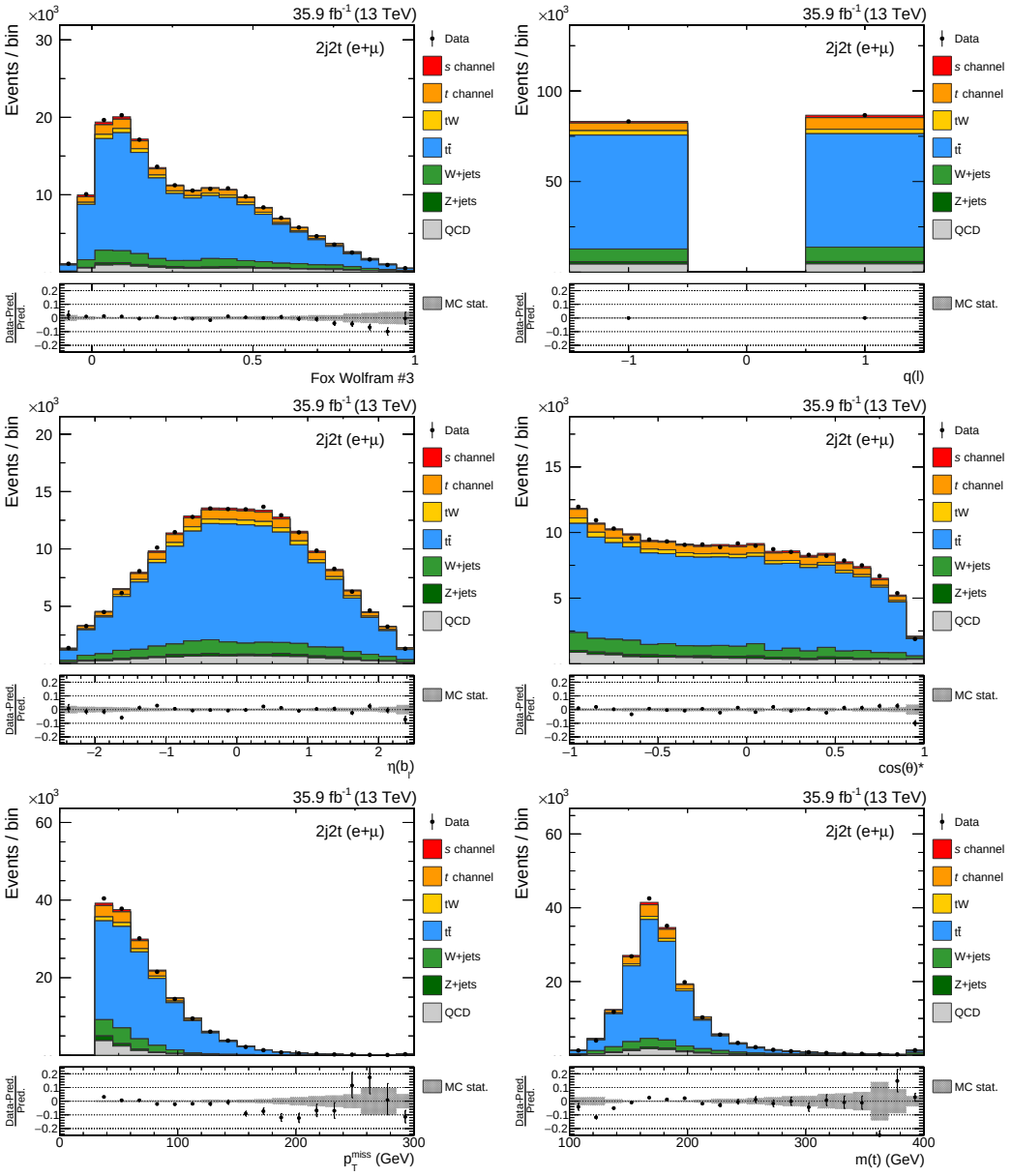

Figure E.2: Input variables ranked 10th to 15th for the DNN training in 2016. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

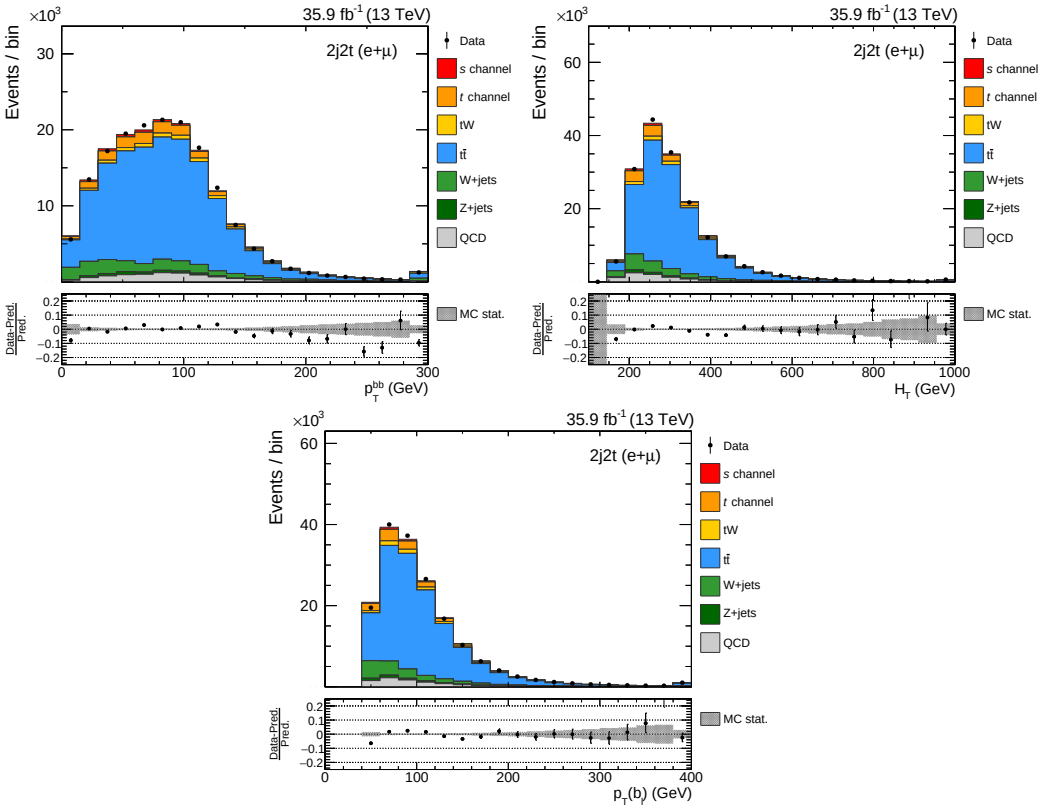

Figure E.3: Input variables ranked 16th to 18th for the DNN training in 2016. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

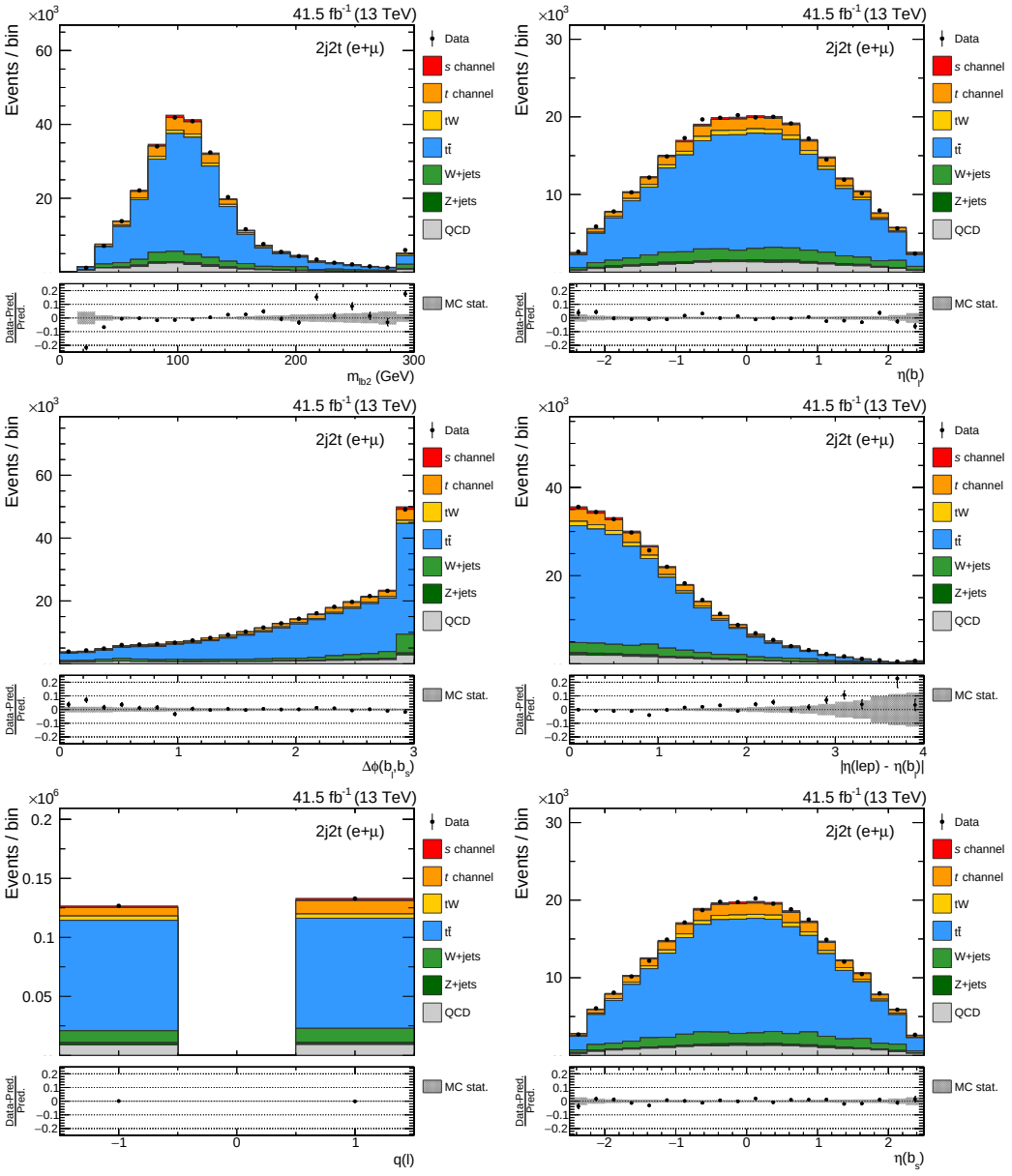

Figure E.4: Input variables ranked 4th to 9th for the DNN training in 2017. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

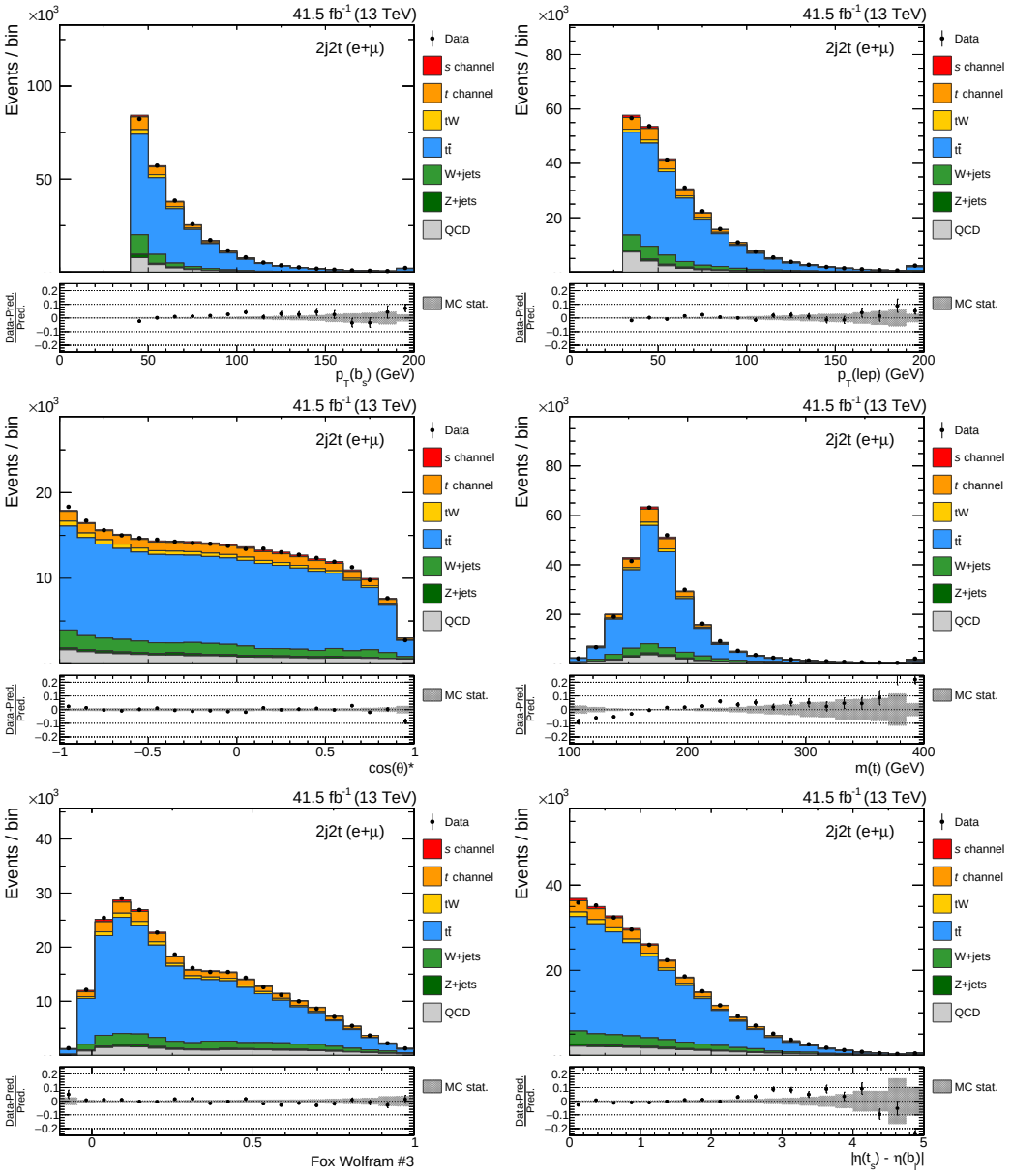

Figure E.5: Input variables ranked 10th to 15th for the DNN training in 2017. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

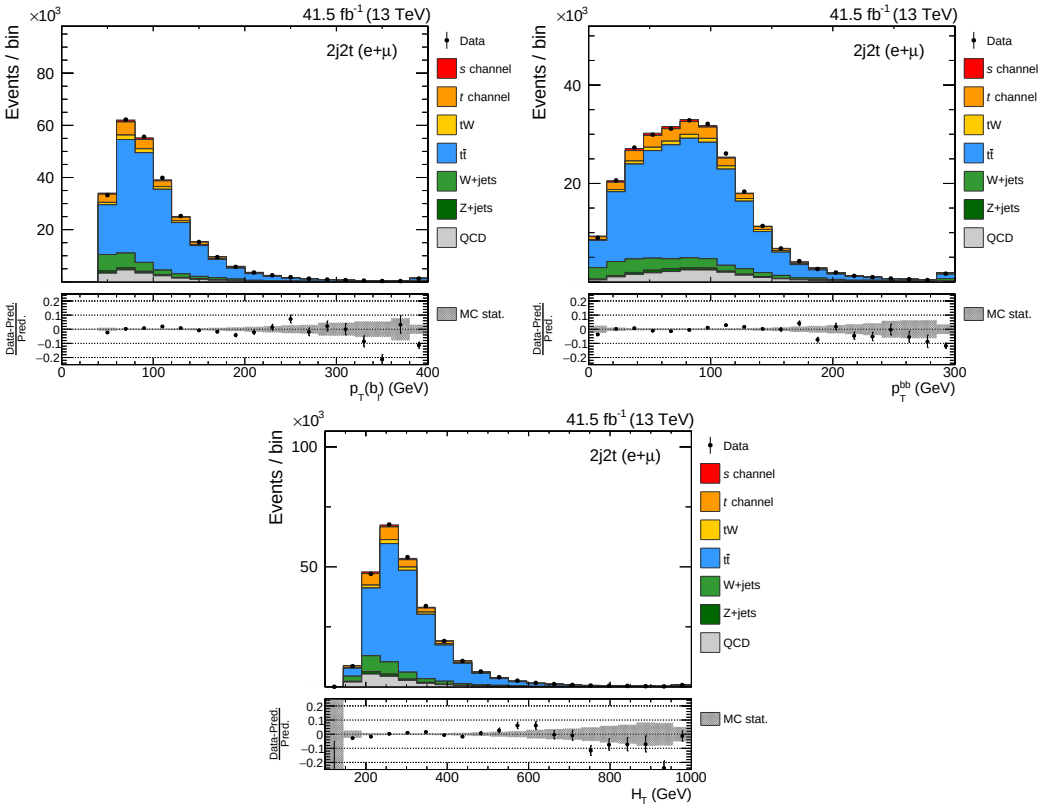

Figure E.6: Input variables ranked 16th to 18th for the DNN training in 2017. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

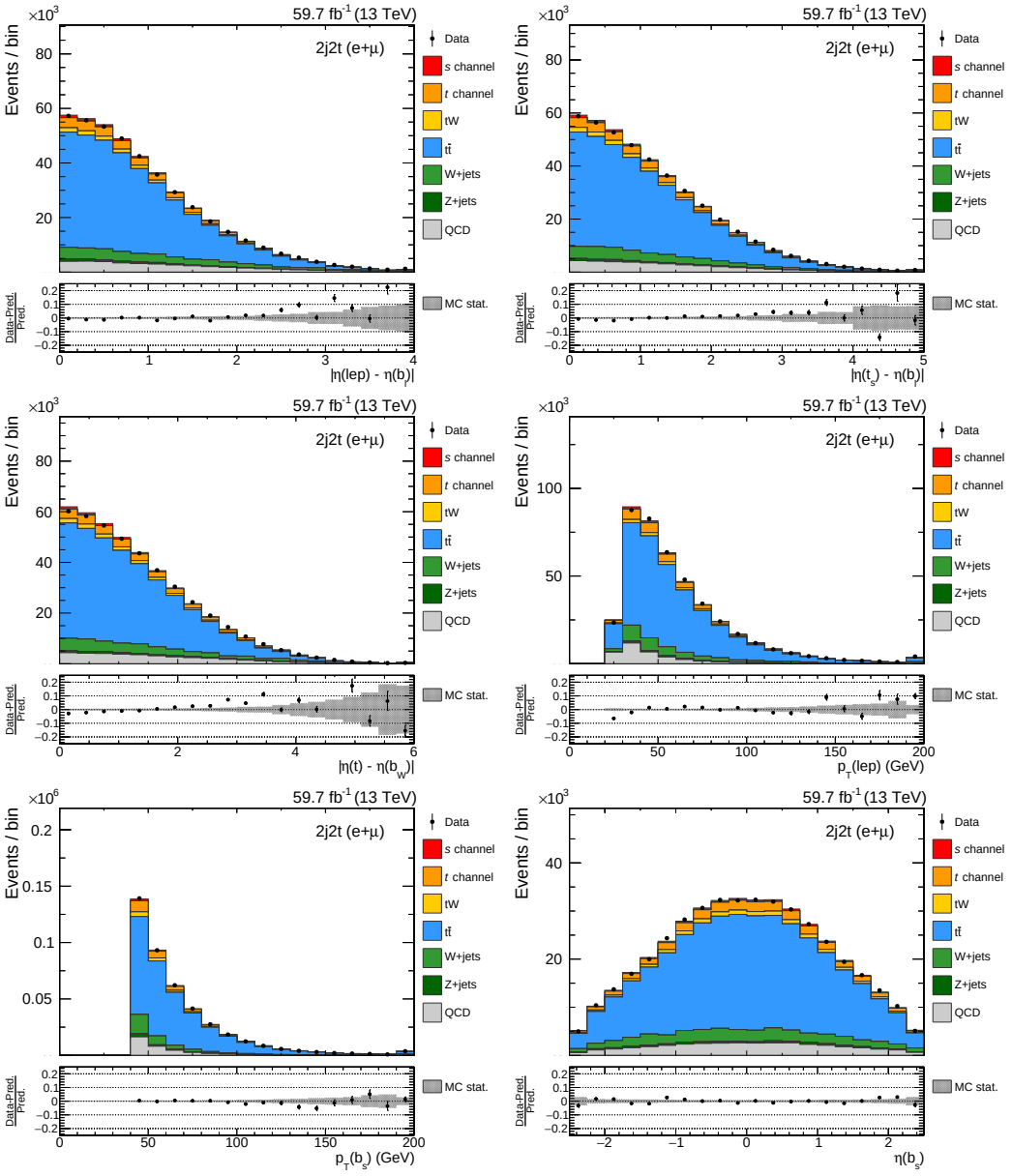

Figure E.7: Input variables ranked 4th to 9th for the DNN training in 2018. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

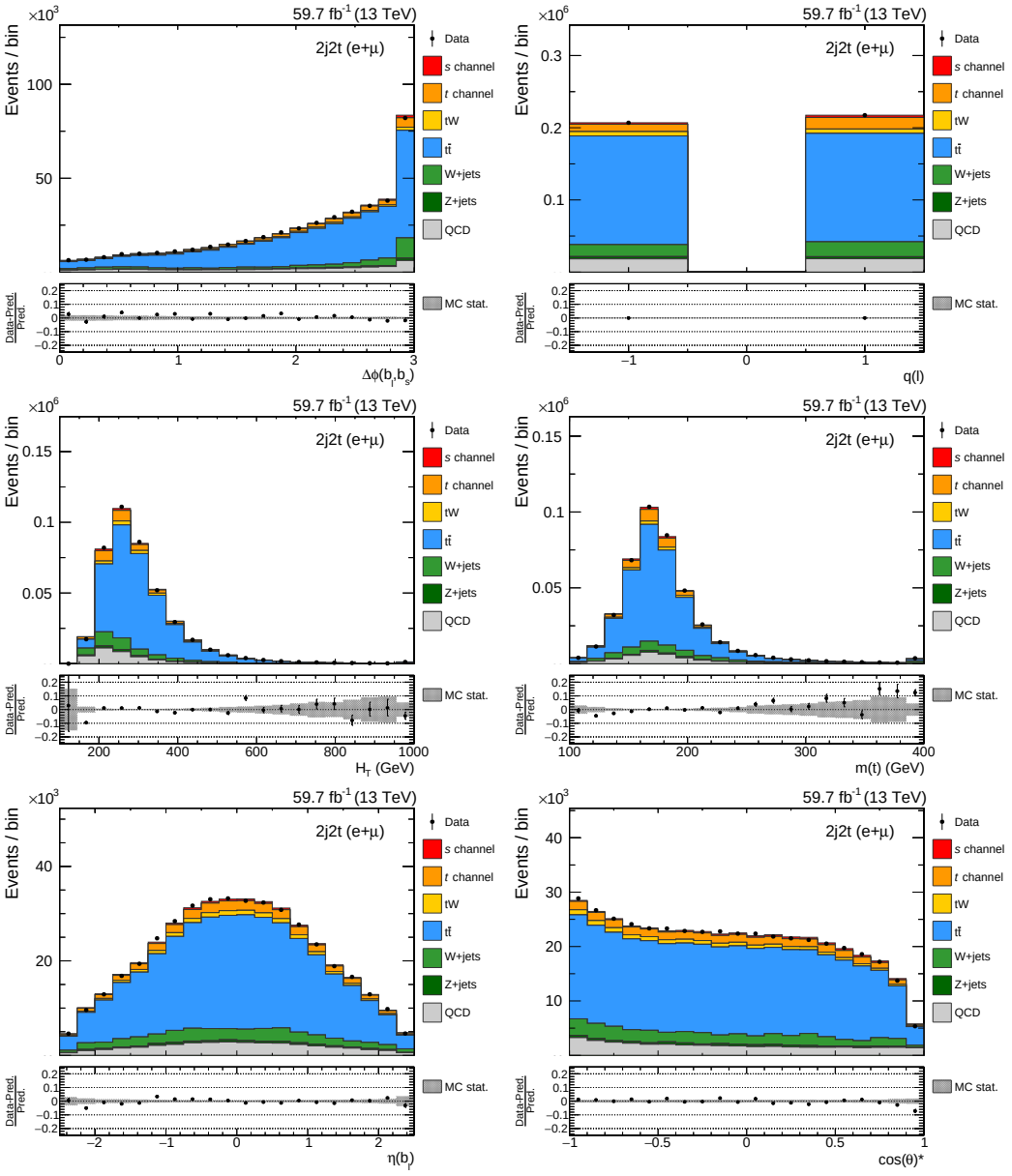

Figure E.8: Input variables ranked 10th to 15th for the DNN training in 2018. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

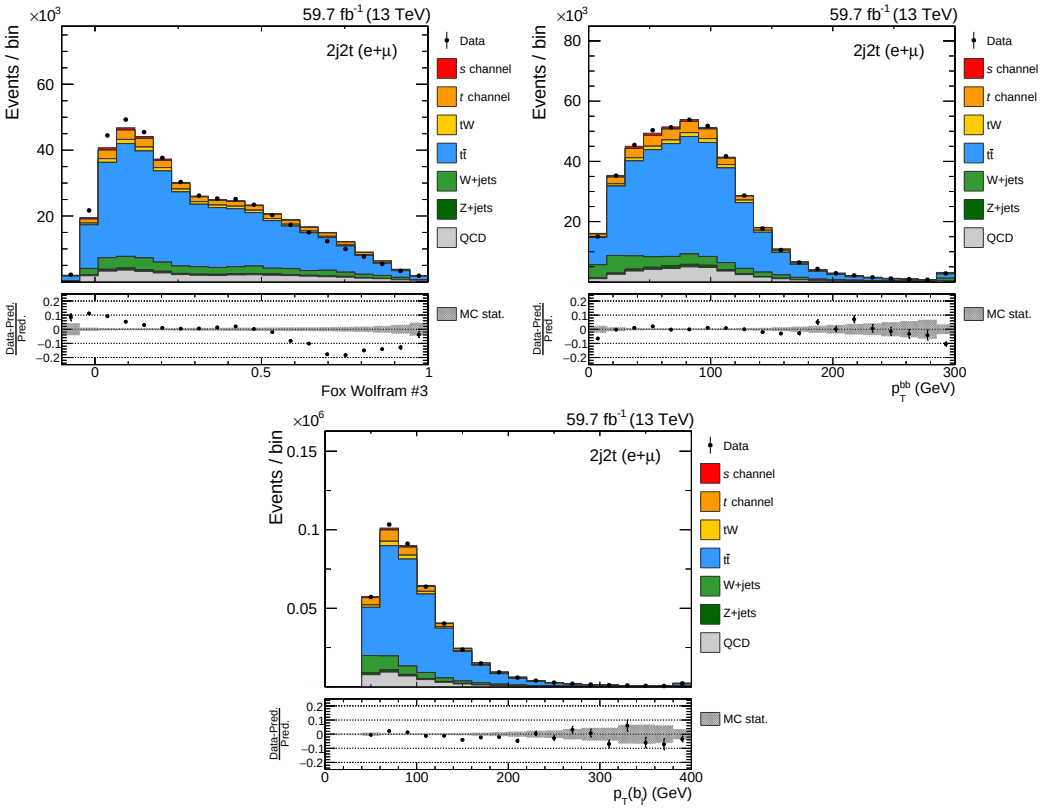

Figure E.9: Input variables ranked 16th to 18th for the DNN training in 2018. The distributions are shown in the 2j2t signal category for simulation and data. The prediction is scaled to the number of observed events. The gray band in the ratio panel corresponds to statistical uncertainties of the simulation.

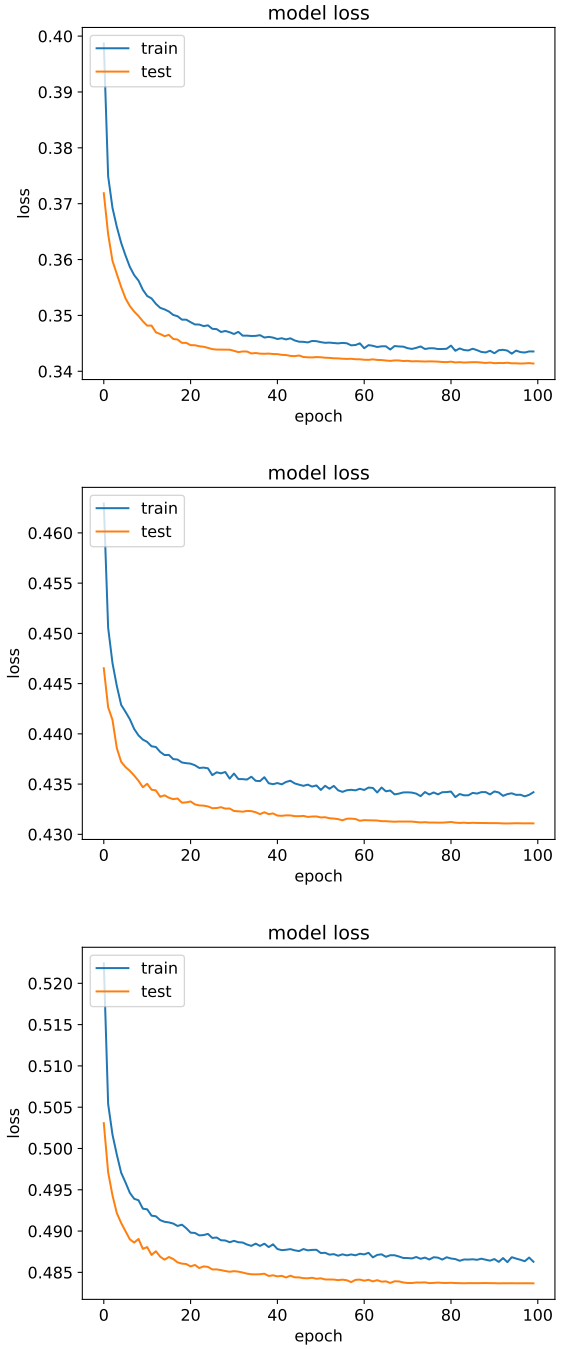

Figure E.10: Loss function of the DNN training. The loss function values in dependence of the epoch are shown for the training and test samples for the 2016 (top), 2017 (center), and 2018 (bottom) analysis. No overtraining is detected.

E.2 Prefit DNN output distributions

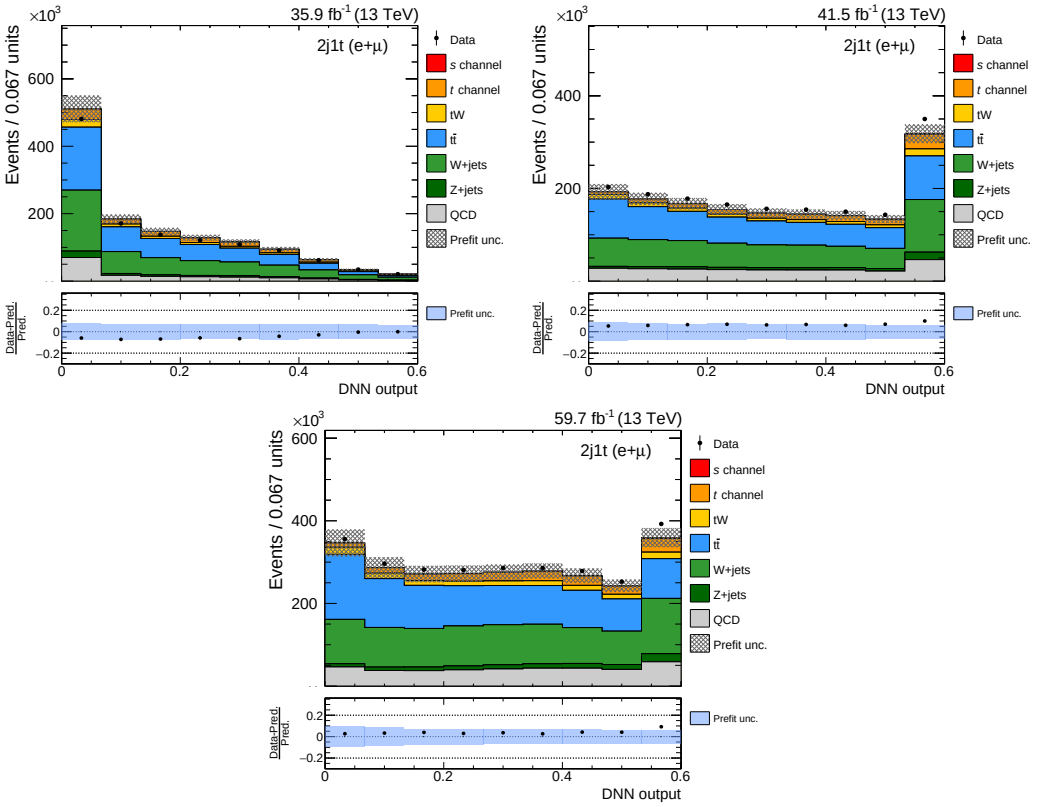

Figure E.11: Prefit distributions of the DNN output values in the 2j1t control category. The distributions are shown for the 2016 (top left), 2017 (top right), and 2018 analysis (bottom). The hatched area in the main panel and the blue uncertainty band in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

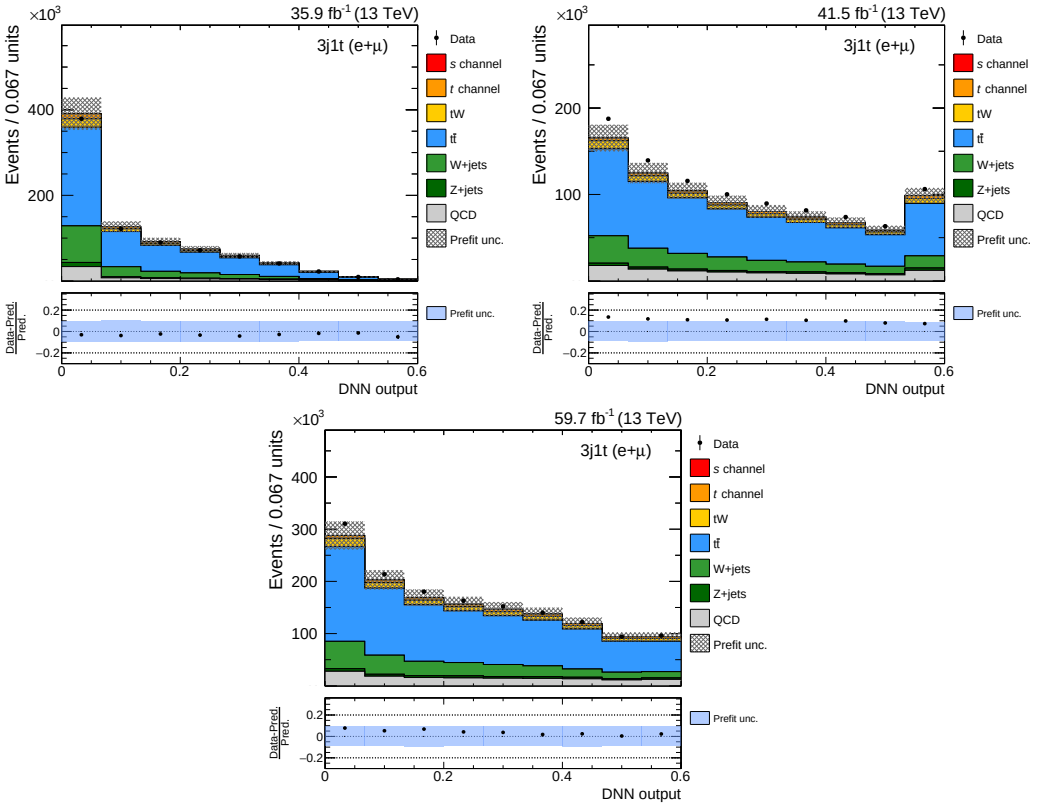

Figure E.12: Prefit distributions of the DNN output values in the 3j1t control category. The distributions are shown for the 2016 (top left), 2017 (top right), and 2018 analysis (bottom). The hatched area in the main panel and the blue uncertainty band in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

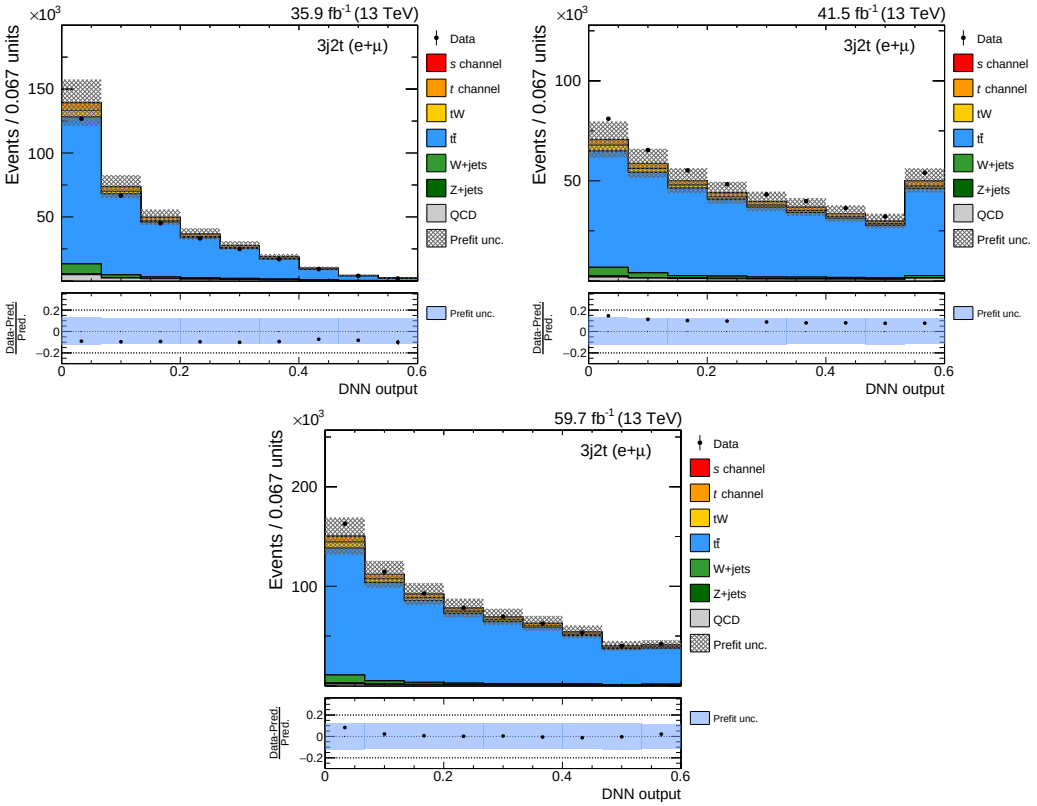

Figure E.13: Prefit distributions of the DNN output values in the 3j2t control category. The distributions are shown for the 2016 (top left), 2017 (top right), and 2018 analysis (bottom). The hatched area in the main panel and the blue uncertainty band in the ratio panel comprise the statistical uncertainty and all systematic uncertainty sources.

F. Correlation model for JES uncertainty sources

Table F.1: Correlation of the JES uncertainties between different years of data taking. The systematic uncertainties are included as nuisance parameters in the fit, and are either considered as fully correlated or uncorrelated.

JES uncertainty source	Corr. 2016/2017 (%)	Corr. 2016/2018 (%)	Corr. 2017/2018 (%)
AbsoluteMPFBias	100	100	100
AbsoluteScale	100	100	100
AbsoluteStat	0	0	0
FlavorQCD	100	100	100
Fragmentation	100	100	100
PileUpDataMC	100	100	100
PileUpPtBB	100	100	100
PileUpPtEC1	100	100	100
PileUpPtEC2	100	100	100
PileUpPtHF	100	100	100
PileUpPtRef	100	100	100
RelativeFSR	100	100	100
RelativeJEREC1	0	0	0
RelativeJEREC2	0	0	0
RelativeJERHF	100	100	100
RelativePtBB	100	100	100
RelativePtEC1	0	0	0
RelativePtEC2	0	0	0
RelativePtHF	100	100	100
RelativeBal	100	100	100
RelativeStatEC	0	0	0
RelativeStatFSR	0	0	0
RelativeStatHF	0	0	0
SinglePionECAL	100	100	100
SinglePionHCAL	100	100	100
TimePtEta	0	0	0

List of Figures

List of Tables

Bibliography

[1] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, pp. 1–29, 2012.

[2] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B*, vol. 716, pp. 30–61, 2012.

[3] F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[4] S. Abachi *et al.*, "Observation of the top quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995.

[5] J. A. Aguilar-Saavedra, "A minimal set of top anomalous couplings," *Nucl. Phys. B*, vol. 812, pp. 181–204, 2009.

[6] T. Aaltonen *et al.*, "Observation of electroweak single top-quark production," *Phys. Rev. Lett.*, vol. 103, p. 092002, 2009.

[7] V. M. Abazov *et al.*, "Observation of single top-quark production," *Phys. Rev. Lett.*, vol. 103, p. 092001, 2009.

[8] M. Gell-Mann and Y. Ne'eman, "The eightfold way," Frontiers in physics, 1964.

[9] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the color octet gluon picture," *Phys. Lett. B*, vol. 47, no. 4, pp. 365 – 368, 1973.

[10] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.

[11] E. Noether, "Invariante Variationsprobleme," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918.

[12] Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998.

[13] Q. R. Ahmad *et al.*, "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory," *Phys. Rev. Lett.*, vol. 87, p. 071301, 2001.

[14] M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev. D*, vol. 98, p. 030001, 2018.

[15] J. Angrik *et al.*, "KATRIN design report 2004," 2005. FZKA-7090.

[16] M. Aker *et al.*, "Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN," *Phys. Rev. Lett.*, vol. 123, p. 221802, 2019.

[17] E. Fermi, "Versuch einer Theorie der β-Strahlen. I," *Zeitschrift für Physik*, vol. 88, no. 3, pp. 161–177, 1934.

[18] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Phys. Rev.*, vol. 105, pp. 1413–1415, 1957.

[19] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[20] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[21] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[22] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[23] L. Álvarez-Gaumé and J. Ellis, "Eyes on a prize particle," *Nature Physics*, vol. 7, p. 2, 2010.

[24] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, 1963.

[25] 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.

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

[27] S. Mandelstam, "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity. General Theory," *Phys. Rev.*, vol. 112, pp. 1344–1360, 1958.

[28] LHC Top Physics Working Group, "NNLO+NNLL top-quark-pair cross sections," https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO. Accessed: 2019-12-01.

[29] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders," *Comput. Phys. Commun.*, vol. 185, p. 2930, 2014.

[30] LHC Top Physics Working Group, "NLO single-top channel cross sections," https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec. Accessed: 2019-12-01.

[31] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR," *Comput. Phys. Commun.*, vol. 182, pp. 1034–1046, 2011.

[32] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions," *Comput. Phys. Commun.*, vol. 191, pp. 74–89, 2015.

[33] M. Brucherseifer, F. Caola, and K. Melnikov, "On the NNLO QCD corrections to single-top production at the LHC," *Phys. Lett. B*, vol. 736, pp. 58–63, 2014.

[34] E. L. Berger, J. Gao, C. P. Yuan, and H. X. Zhu, "NNLO QCD Corrections to t-channel Single Top-Quark Production and Decay," *Phys. Rev. D*, vol. 94, no. 7, p. 071501, 2016.

[35] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-," *Phys. Rev. D*, vol. 82, p. 054018, 2010.

[36] N. Kidonakis, "Top Quark Production," in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, July 15-28, 2013*, pp. 139–168, 2014.

[37] Z. L. Liu and J. Gao, "s-channel single top quark production and decay at next-to-next-to-leading-order in QCD," *Phys. Rev. D*, vol. 98, no. 7, p. 071501, 2018.

[38] T. A. Aaltonen *et al.*, "Observation of *s*-channel production of single top quarks at the Tevatron," *Phys. Rev. Lett.*, vol. 112, p. 231803, 2014.

[39] G. Aad *et al.*, "Evidence for single top-quark production in the *s*-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method," *Phys. Lett. B*, vol. 756, p. 228, 2016.

[40] M. Hashemi, "Observability of Heavy Charged Higgs through s-channel Single Top Events at LHC," *JHEP*, vol. 11, p. 005, 2013.

[41] F. Huang, H.-L. Li, S.-Y. Li, Z.-G. Si, W. Su, and Z.-J. Yang, "Search for W′ signal in single top quark production at the LHC," *Chinese Physics C*, vol. 42, no. 3, p. 033103, 2018.

[42] T. M. P. Tait and C. P. Yuan, "Single top quark production as a window to physics beyond the standard model," *Phys. Rev. D*, vol. 63, p. 014018, 2000.

[43] L. Evans and P. Bryant, "LHC machine," *JINST*, vol. 3, no. 08, pp. S08001–S08001, 2008.

[44] CERN, "LHC Guide," CERN-Brochure-2017-002-Eng, 2017.

[45] CERN, "The Large Hadron Collider," https://home.cern/science/accelerators/large-hadron-collider. Accessed: 2019-12-09.

[46] E. Mobs, "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019," 2019. General Photo, CERN-GRAPHICS-2019-002.

[47] CERN, "CERN releases analysis of LHC incident," https://home.cern/news/press-release/cern/cern-releases-analysis-lhc-incident. Accessed: 2019-12-29.

[48] CERN, "A new schedule for the LHC and its successor," https://home.cern/news/news/accelerators/new-schedule-lhc-and-its-successor. Accessed: 2019-12-29.

[49] CERN, "LHC performance reaches new highs," https://home.cern/news/news/accelerators/lhc-performance-reaches-new-highs. Accessed: 2019-12-29.

[50] CERN, "Public CMS Luminosity Information," https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. Accessed: 2019-12-29.

[51] S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[52] CMS Collaboration, "CMS Detector," https://cms.cern/detector. Accessed: 2019-12-29.

[53] UZH CMS Group, "CMS Wiki Pages," https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates. Accessed: 2019-12-29.

[54] V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen, and A. Cattai, "The CMS tracker system project: Technical Design Report," CERN-LHCC-98-006, 1997.

[55] CMS Collaboration, "The CMS tracker: addendum to the Technical Design Report," CERN-LHCC-2000-016, 2000.

[56] A. Dominguez, "The CMS pixel detector," *Nucl. Instrum. Methods Phys. Res. A*, vol. 581, pp. 343 – 346, 2007.

[57] A. Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung, *et al.*, "CMS Technical Design Report for the Pixel Detector Upgrade," CERN-LHCC-2012-016. CMS-TDR-11, 2012.

[58] J.-L. Agram, "CMS Silicon Strip Tracker Performance," *Physics Procedia*, vol. 37, pp. 844 – 850, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).

[59] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report," CERN-LHCC-97-033, 1997.

[60] P. Bloch, R. Brown, P. Lecoq, and H. Rykaczewski, "Changes to CMS ECAL electronics: addendum to the Technical Design Report," CERN-LHCC-2002-027, 2002.

[61] Q. Ingram, "Energy resolution of the barrel of the CMS electromagnetic calorimeter," *JINST*, vol. 2, no. 04, pp. P04004–P04004, 2007.

[62] A. Benaglia, "The CMS ECAL performance with examples," *JINST*, vol. 9, no. 02, p. C02008, 2014.

[63] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report," CERN-LHCC-97-031, 1997.

[64] V. Kryshkin and A. Ronzhin, "An optical fiber readout for scintillator calorimeters," *Nucl. Instrum. Methods Phys. Res. A*, vol. 247, no. 3, pp. 583 – 585, 1986.

[65] M. Albrow, G. Arnison, J. Bunn, D. Clarke, C. Cochet, P. Colas, D. Dallman, J. de Brion, B. Denby, E. Eisenhandler, *et al.*, "A uranium scintillator calorimeter with plastic-fibre readout," *Nucl. Instrum. Methods Phys. Res. A*, vol. 256, no. 1, pp. 23 – 37, 1987.

[66] CMS Collaboration, "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays," *JINST*, vol. 5, no. 03, pp. T03021–T03021, 2010.

[67] J. G. Layter, "The CMS muon project: Technical Design Report," CERN-LHCC-97-032, 1997.

[68] A. M. Sirunyan *et al.*, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV," *JINST*, vol. 13, no. 06, p. P06015, 2018.

[69] A. Colaleo, A. Safonov, A. Sharma, and M. Tytgat, "CMS Technical Design Report for the Muon Endcap GEM Upgrade," CERN-LHCC-2015-012. CMS-TDR-013, 2015.

[70] CMS Collaboration, "The Phase-2 Upgrade of the CMS Muon Detectors," CERN-LHCC-2017-012. CMS-TDR-016, CERN, 2017.

[71] S. Dasu *et al.*, "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems," CERN-LHCC-2000-038, 2000.

[72] S. Cittolin, A. Rácz, and P. Sphicas, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project," CERN-LHCC-2002-026, 2002.

[73] A. Tapper and D. Acosta, "CMS Technical Design Report for the Level-1 Trigger Upgrade," CERN-LHCC-2013-011. CMS-TDR-12, 2013.

[74] V. Khachatryan *et al.*, "The CMS trigger system," *JINST*, vol. 12, no. 01, p. P01020, 2017.

[75] M. Jeitler, "Upgrade of the trigger system of CMS," *Nucl. Instrum. Methods Phys. Res. A*, vol. 718, p. 11, 2013.

[76] I. Bird, K. Bos, N. Brook, D. Duellmann, C. Eck, I. Fisk, D. Foster, B. Gibbard, C. Grandi, F. Grey, *et al.*, "LHC computing Grid. Technical design report," CERN-LHCC-2005-024, 2005.

[77] I. Bird *et al.*, "Update of the Computing Models of the WLCG and the LHC Experiments," CERN-LHCC-2014-014, LCG-TDR-002, 2014.

[78] Worldwide LHC Computing Grid, "WLCG Document Repository," http://wlcg-docs.web.cern.ch/wlcg-docs/?dir=outreach/images. Accessed: 2019-05-10.

[79] S. Höche, "Introduction to parton-shower event generators," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 235–295, 2015.

[80] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[81] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972. [Yad. Fiz.15,781(1972)].

[82] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language," *Nucl.Phys. B*, vol. 126, no. 2, pp. 298 – 318, 1977.

[83] R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015.

[84] R. D. Ball *et al.*, "Parton distributions from high-precision collider data," *Eur. Phys. J. C*, vol. 77, p. 663, 2017.

[85] G. T. Bodwin, "Factorization of the Drell-Yan Cross-Section in Perturbation Theory," *Phys. Rev. D*, vol. 31, p. 2616, 1985. [Erratum: Phys. Rev.D34,3932(1986)].

[86] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization for Short Distance Hadron - Hadron Scattering," *Nucl. Phys. B*, vol. 261, pp. 104–142, 1985.

[87] J. C. Collins, D. E. Soper, and G. F. Sterman, "Soft Gluons and Factorization," *Nucl. Phys. B*, vol. 308, pp. 833–856, 1988.

[88] V. V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics," *Sov. Phys. JETP*, vol. 3, pp. 65–71, 1956. [Zh. Eksp. Teor. Fiz.30,87(1956)].

[89] A. Sen, "Asymptotic Behavior of the Sudakov Form-Factor in QCD," *Phys. Rev. D*, vol. 24, p. 3281, 1981.

[90] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions," *JHEP*, vol. 01, p. 013, 2007.

[91] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO," *JHEP*, vol. 12, p. 061, 2012.

[92] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983.

[93] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949.

[94] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[95] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP*, vol. 06, p. 128, 2011.

[96] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 06, p. 029, 2002.

[97] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP*, vol. 11, p. 040, 2004.

[98] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[99] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010.

[100] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[101] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015.

[102] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Methods Phys. Res. A*, vol. 506, no. 3, pp. 250 – 303, 2003.

[103] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.*, vol. 53, p. 270, 2006.

[104] J. Allison *et al.*, "Recent developments in Geant4," *Nucl. Instrum. Methods Phys. Res. A*, vol. 835, pp. 186 – 225, 2016.

[105] A. M. Sirunyan *et al.*, "Particle-flow reconstruction and global event description with the CMS detector," *JINST*, vol. 12, no. 10, p. P10003, 2017.

[106] G. Petrucciani, A. Rizzi, and C. Vuosalo, "Mini-AOD: A New Analysis Data Format for CMS," in *Proceedings, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP 2015): Okinawa, Japan, April 13-17, 2015*, p. 072052, 2015.

[107] S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. P10009, 2014.

[108] P. Billoir, "Progressive track recognition with a Kalman-like fitting procedure," *Comput. Phys. Commun.*, vol. 57, no. 1, pp. 390 – 394, 1989.

[109] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method," *Nucl. Instrum. Methods Phys. Res. A*, vol. 294, no. 1, pp. 219 – 228, 1990.

[110] R. Mankel, "A concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system," *Nucl. Instrum. Methods Phys. Res. A*, vol. 395, no. 2, pp. 169 – 184, 1997.

[111] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2210–2239, 1998.

[112] R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting," CMS-NOTE-2007-008, 2007.

[113] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, "Electron reconstruction in CMS," *Eur. Phys. J. C*, vol. 49, pp. 1099–1116, 2007.

[114] V. Khachatryan *et al.*, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV," *JINST*, vol. 10, p. P06005, 2015.

[115] R. Frühwirth, "Track fitting with non-gaussian noise," *Comput. Phys. Commun.*, vol. 100, no. 1, pp. 1 – 16, 1997.

[116] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC," CMS-NOTE-2005-001, 2005.

[117] V. Khachatryan *et al.*, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV," *JINST*, vol. 10, no. 08, p. P08010, 2015.

[118] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[119] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms," *JHEP*, vol. 08, p. 001, 1997.

[120] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering," in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 270–279, 1998.

[121] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys. B*, vol. 406, pp. 187–224, 1993.

[122] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys. Rev. D*, vol. 48, pp. 3160–3166, 1993.

[123] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," *Phys. Lett. B*, vol. 641, pp. 57–61, 2006.

[124] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur. Phys. J. C*, vol. 72, p. 1896, 2012.

[125] "Introduction to Jet Energy Corrections at CMS.," https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC. Restricted webpage. Accessed: 2020-01-18.

[126] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS," *JINST*, vol. 6, no. 11, pp. P11002–P11002, 2011.

[127] V. Khachatryan *et al.*, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV," *JINST*, vol. 12, no. 02, pp. P02014–P02014, 2017.

[128] S. Chatrchyan *et al.*, "Identification of b-Quark Jets with the CMS Experiment," *JINST*, vol. 8, p. P04013, 2013.

[129] A. M. Sirunyan *et al.*, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV," *JINST*, vol. 13, p. P05011, 2018.

[130] M. Stoye, "Deep learning in jet reconstruction at CMS," *Journal of Physics: Conference Series*, vol. 1085, p. 042029, 2018.

[131] CMS Collaboration, "Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector," CMS-DP-2018-033, 2018.

[132] W. Zhang, K. Itoh, J. Tanida, and Y. Ichioka, "Parallel distributed processing model with local space-invariant interconnections and its optical architecture," *Appl. Opt.*, vol. 29, no. 32, pp. 4790–4797, 1990.

[133] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[134] M. Verzetti, "Machine learning techniques for jet flavour identification at CMS," *EPJ Web of Conferences*, vol. 214, p. 06010, 2019.

[135] A. M. Sirunyan *et al.*, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}$ = 13 TeV using the CMS detector," *JINST*, vol. 14, no. 07, p. P07004, 2019.

[136] CMS Collaboration, "Baseline muon selections for Run-II," https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2. Restricted webpage. Accessed: 2020-01-24.

[137] S. Chatrchyan *et al.*, "Search for neutral Higgs bosons decaying to tau pairs in *pp* collisions at $\sqrt{s}$ = 7 TeV," *Phys. Lett. B*, vol. 713, pp. 68–90, 2012.

[138] CMS Collaboration, "Cut Based Electron ID for Run 2," https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2. Restricted webpage. Accessed: 2020-01-24.

[139] CMS Collaboration, "Jet Identification," https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID. Restricted webpage. Accessed: 2020-01-24.

[140] CMS Collaboration, "Recommendation for Using b-tag Objects in Physics Analyses," https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation. Restricted webpage. Accessed: 2020-01-24.

[141] CMS Collaboration, "MET Filter Recommendations for Run II," https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2. Restricted webpage. Accessed: 2020-01-24.

[142] A. Lenail, "NN-SVG: Publication-ready NN-architecture schematics," http://alexlenail.me/NN-SVG/index.html. Accessed: 2019-05-20.

[143] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[144] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2014. arXiv:1412.6980 [cs.LG].

[145] F. Chollet *et al.*, "Keras," https://keras.io, 2015. Accessed: 2020-01-27.

[146] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016. arXiv:1603.04467 [cs.DC].

[147] A. L. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," *G. Ist. Ital. Attuari*, vol. 4, pp. 83–91, 1933.

[148] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, no. 2, pp. 279–281, 1948.

[149] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012. arXiv:1207.0580 [cs.NE].

[150] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.

[151] ATLAS and CMS Collaborations, LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011," CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, 2011.

[152] CMS Higgs Working Group, "Combine," https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/. Accessed: 2020-01-27.

[153] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf*, vol. C0303241, p. MOLT007, 2003. [,186(2003)].

[154] A. L. Read, "Linear interpolation of histograms," *Nucl. Instrum. Methods Phys. Res. A*, vol. 425, pp. 357–360, 1999.

[155] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra," in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011*, pp. 115–120, 2011.

[156] J. Neyman and E. S. Pearson, "IX. On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Phil. Trans. R. Soc. Lond. A*, vol. 231, pp. 289–337, 1933.

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

[158] A. M. Sirunyan *et al.*, "Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s}$ = 13 TeV," *Phys. Lett. B*, vol. 800, p. 135042, 2020.

[159] A. M. Sirunyan *et al.*, "Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s}$ = 13 TeV," *JHEP*, vol. 10, p. 117, 2018.

[160] V. M. Abazov *et al.*, "Evidence for s-Channel Single Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV," *Phys. Lett. B*, vol. 726, pp. 656–664, 2013.

[161] S. Sumowidagdo, "A pedagogical derivation of the matrix element method in particle physics data analysis," *Journal of Physics: Conference Series*, vol. 988, p. 012003, 2018.

[162] V. Khachatryan *et al.*, "Search for s channel single top quark production in pp collisions at $\sqrt{s}$ = 7 and 8 TeV," *JHEP*, vol. 09, p. 027, 2016.

[163] A. Giammanco, "Single top quark production at the LHC," *Rev. Phys.*, vol. 1, pp. 1–12, 2016.

[164] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," *JHEP*, vol. 09, p. 111, 2009. [Erratum: JHEP02,011(2010)].

[165] R. Frederix, E. Re, and P. Torrielli, "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO," *JHEP*, vol. 09, p. 130, 2012.

[166] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur. Phys. J. C*, vol. 71, p. 1547, 2011.

[167] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction," *JHEP*, vol. 09, p. 126, 2007.

[168] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J. C*, vol. 74, no. 8, p. 3024, 2014.

[169] V. Khachatryan *et al.*, "Event generator tunes obtained from underlying event and multiparton scattering measurements," *Eur. Phys. J. C*, vol. 76, no. 3, p. 155, 2016.

[170] A. M. Sirunyan *et al.*, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements," *Eur. Phys. J. C*, vol. 80, no. 1, p. 4, 2020.

[171] CMS Collaboration, "Discussion of noisy EE jets in 2017 data," https://docs.google.com/presentation/d/1PMWr9QEaHYAbyi6yAFLHEE6KT-u7Cjwm3wniJZCixcU/edit#slide=id.p. Accessed: 2020-02-08.

[172] CMS Collaboration, "Treatment of the HEM15/16 region in 2018 data," https://hypernews.cern.ch/HyperNews/CMS/get/JetMET/2000.html. Restricted webpage. Accessed: 2020-02-08.

[173] T. Chwalek, "Messung der W-Boson-Helizitätsanteile in Top-Quark-Zerfällen mit dem CDF II Experiment und Studien zu einer frühen Messung des $t\bar{t}$-Wirkungsquerschnitts mit dem CMS Experiment," 2010. CERN-THESIS-2010-255.

[174] CMS Collaboration, "CMS Luminosity – Public Results," https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. Accessed: 2020-02-10.

[175] CMS Collaboration, "Pileup in Monte Carlo," https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. Accessed: 2020-02-10.

[176] CMS Collaboration, "Pileup Reweighting Utilities," https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupMCReweightingUtilities. Restricted webpage. Accessed: 2020-02-10.

[177] CMS Collaboration, "Utilities for Accessing Pileup Information for Data," https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData. Restricted webpage. Accessed: 2020-02-10.

[178] A. M. Sirunyan *et al.*, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV," *JHEP*, vol. 07, p. 161, 2018.

[179] CMS Collaboration, "Tag and Probe," https://twiki.cern.ch/twiki/bin/view/CMS/MuonTagAndProbe. Restricted webpage. Accessed: 2020-02-11.

[180] CMS Collaboration, "Muon T&P Instructions for Run-II," https://twiki.cern.ch/twiki/bin/view/CMS/MuonTagAndProbeTreesRun2. Restricted webpage. Accessed: 2020-02-11.

[181] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2016 legacy re-reco data.," https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco. Restricted webpage. Accessed: 2020-02-11.

[182] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2017 data.," https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017. Restricted webpage. Accessed: 2020-02-11.

[183] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2018 data.," https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018. Restricted webpage. Accessed: 2020-02-11.

[184] CMS Collaboration, "Event selection for t-channel single-top-quark analyses with the 2016 dataset at 13 TeV," CMS Analysis Note AN-2017/056, CERN, 2017.

[185] CMS Collaboration, "Tag and Probe," https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe. Accessed: 2020-02-11.

[186] CMS Collaboration, "Electron Tag-and-Probe," https://twiki.cern.ch/twiki/bin/view/CMSPublic/ElectronTagAndProbe. Accessed: 2020-02-11.

[187] CMS Collaboration, "E/gamma RunII Recommendations," https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations. Restricted webpage. Accessed: 2020-02-11.

[188] P. Ott, "Inclusive cross section measurement for the production of single top quarks in the t channel at 13 TeV with the CMS experiment," 2017. Master thesis, Karlsruhe Institute of Technology (KIT), ETP-KA/2017-32.

[189] CMS Collaboration, "Measurement of the 2016 Trigger Efficiencies for a dilepton selection for a ttbar analysis," CMS Analysis Note AN-2016/392, CERN, 2016.

[190] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors," https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods. Restricted webpage. Accessed: 2020-02-11.

[191] CMS Collaboration, "Heavy flavour tagging for 13 TeV 2016 Legacy data," https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation2016Legacy. Restricted webpage. Accessed: 2020-02-11.

[192] CMS Collaboration, "Heavy flavour tagging for 13 TeV data in 2017 and 94X MC," https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation94X. Restricted webpage. Accessed: 2020-02-11.

[193] CMS Collaboration, "Heavy flavour tagging for 13 TeV data in 2018 and 10_2_X MC," https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation102X. Restricted webpage. Accessed: 2020-02-11.

[194] CMS Collaboration, "Reweighting recipe to emulate Level 1 ECAL prefiring.," https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1ECALPrefiringWeightRecipe. Restricted webpage. Accessed: 2020-02-11.

[195] CMS Collaboration, "Preliminary 2016/2017 L1 Prefiring maps," https://lathomas.web.cern.ch/lathomas/TSGStuff/L1Prefiring/PrefiringMaps_2016and2017/. Restricted webpage. Accessed: 2020-03-11.

[196] G. C. Fox and S. Wolfram, "Observables for the analysis of event shapes in e^+e^- annihilation and other processes," *Phys. Rev. Lett.*, vol. 41, pp. 1581–1585, 1978.

[197] R. D. Field, Y. Kanev, and M. Tayebnejad, "Topological analysis of the top quark signal and background at hadron colliders," *Phys. Rev. D*, vol. 55, pp. 5685–5697, 1997.

[198] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period," CMS-PAS-LUM-17-001, 2017.

[199] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV," CMS-PAS-LUM-17-004, 2018.

[200] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s}$ = 13 TeV," CMS-PAS-LUM-18-002, 2019.

[201] CMS Collaboration, "Estimating Systematic Errors Due to Pileup Modeling," https://twiki.cern.ch/twiki/bin/view/CMS/PileupSystematicErrors. Restricted webpage. Accessed: 2020-02-14.

[202] CMS Collaboration, "Jet energy scale uncertainty sources," https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources. Restricted webpage. Accessed: 2020-02-15.

[203] CMS Collaboration, "Jet Energy Resolution," https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution. Restricted webpage. Accessed: 2020-02-15.

[204] CMS Collaboration, "MET Corrections and Uncertainties for Run-II," https://twiki.cern.ch/twiki/bin/view/CMS/MissingETRun2Corrections. Restricted webpage. Accessed: 2020-02-15.

[205] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples," *Comput. Phys. Commun.*, vol. 77, no. 2, pp. 219 – 228, 1993.

[206] CMS Collaboration, "Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing," https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns. Restricted webpage. Accessed: 2020-02-06.

[207] S. Alekhin *et al.*, "The PDF4LHC Working Group Interim Report," 2011. arXiv:1101.0536 [hep-ph].

[208] M. Botje *et al.*, "The PDF4LHC Working Group Interim Recommendations," 2011. arXiv:1101.0538 [hep-ph].

[209] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys. G*, vol. 43, p. 023001, 2016.

[210] A. Kalogeropoulos and J. Alwall, "The SysCalc code: A tool to derive theoretical systematic uncertainties," 2018. arXiv:1801.08401 [hep-ph].

[211] CMS Collaboration, "hdamp for TuneCP5," https://gitlab.cern.ch/cms-gen/Tuning/merge_requests/11. Restricted webpage. Accessed: 2020-02-15.

[212] CMS Collaboration, "Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $\mathrm{t\bar{t}}$ at $\sqrt{s}$ = 8 and 13 TeV," CMS-PAS-TOP-16-021, 2016.

[213] V. Khachatryan *et al.*, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s}$ = 8 TeV," *Eur. Phys. J. C*, vol. 75, no. 11, p. 542, 2015.

[214] CMS Collaboration, "pt(top-quark) based reweighting of ttbar MC," https://twiki.cern.ch/twiki/bin/view/CMS/TopPtReweighting. Restricted webpage. Accessed: 2020-02-15.

[215] A. M. Sirunyan *et al.*, "Measurement of the $\mathrm{t\bar{t}}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s}$ = 13 TeV," *JHEP*, vol. 09, p. 051, 2017.

[216] A. M. Sirunyan *et al.*, "Measurement of the $\mathrm{t\bar{t}}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C*, vol. 79, no. 5, p. 368, 2019.

[217] CMS Collaboration, "TOP Systematic Uncertainties (Run2)," https://twiki.cern.ch/twiki/bin/view/CMS/TopSystematics. Restricted webpage. Accessed: 2020-02-18.

[218] CMS Collaboration, "Luminosity Physics Object Group (Lumi POG)," https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiLUM. Restricted webpage. Accessed: 2020-02-19.

[219] G. J. Feldman and R. D. Cousins, "A Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D*, vol. 57, pp. 3873–3889, 1998.

[220] CERN, "The HL-LHC project," https://hilumilhc.web.cern.ch/content/hl-lhc-project. Accessed: 2020-02-26.

[221] G. Apollinari, I. Béjar Alonso, O. Brüning, P. Fessia, M. Lamont, L. Rossi, and L. Tavian, "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1," CERN Yellow Reports: Monographs, 2017.

[222] A. Escamilla, A. O. Bouzas, and F. Larios, "Single-top production at linear e^-e^+ colliders," *Phys. Rev. D*, vol. 97, p. 033004, 2018.

[223] H. Khanpour, S. Khatibi, M. Khatiri Yanehsari, and M. Mohammadi Najafabadi, "Single top quark production as a probe of anomalous $tq\gamma$ and tqZ couplings at the FCC-ee," *Phys. Lett. B*, vol. 775, pp. 25–31, 2017.

Danksagung

Zuallererst möchte ich mich bei Prof. Dr. Thomas Müller für die Ermöglichung meiner Promotion im Anschluss meiner Masterarbeit innerhalb seiner Arbeitsgruppe, sowie meiner Teilnahme an zahlreichen internationalen Konferenzen, Workshops und Schulen bedanken. Insbesondere bin ich ihm sehr dankbar für meinen halbjährigen Aufenthalt am CERN, der mir einen wissenschaftlichen Austausch mit Physikern aus aller Welt ermöglicht hat.

Ebenso geht mein Dank an Prof. Dr. Ulrich Husemann für die Übernahme des Korreferats während meiner Promotion und auch während meiner Masterarbeit, sowie für den Einstieg in die Teilchenphysik durch die Ermöglichung meiner Bachelorarbeit in seiner Arbeitsgruppe.

Weiterhin möchte ich mich bei Dr. Thorsten Chwalek und Dr. Nils Faltermann für die hervorragende Betreuung meiner Doktorarbeit für das geduldige Beantworten aller offenen Fragen bedanken.

Mein Dank geht an Darius Bühler, Dr. Nils Faltermann, Dr. Kevin Flöh, Max Neukum, Dr. Daniela Schäfer, David Seith und Priv. Doz. Dr. Roger Wolf für das Korrekturlesen meiner Doktorarbeit. Ich danke Jochen Gemmler für die gemeinsame und sehr hilfreiche Prüfungsvorbereitung (die aufgrund der Coronavirus-Pandemie hauptsächlich über Skype und Zoom stattfinden musste). Darüberhinaus möchte ich mich bei allen Mitgliedern meiner Arbeitsgruppe sowie den Mitgliedern des Instituts für die gute Atmosphäre bedanken. Ebenso danke ich allen derzeitigen und ehemaligen Mitgliedern des Admin-Teams für die Bereitstellung und Instandhaltung der lokalen Computingressourcen.

Dem Graduiertenkolleg *GRK 1694: Elementarteilchenphysik bei höchster Energie und höchster Präzision* und der Graduiertenschule *Karlsruher Schule für Elementarteilchen- und Astroteilchenphysik: Wissenschaft und Technologie (KSETA)* möchte ich für das zahlreiche Angebot an Kolloquien sowie für die vielfältigen Weiterbildungskurse danken. An dieser Stelle geht mein Dank an die aktuellen und ehemaligen Sprecher des Graduiertenkollegs für die gute Zusammenarbeit während meiner Zeit als Doktorandensprecherin.

Zuallerletzt möchte ich mich bei meinen Freunden und bei meiner Familie für ihre Unterstützung während der Promotion sowie in allen anderen Lebenssituationen bedanken.