ETP-KA/2021-11

Measurements of the $t\bar{t}H(b\bar{b})$ Process with Neural Networks at the CMS Experiment - Current Status and Future Prospects

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)

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

DISSERTATION

von M.Sc. Philip Keicher

aus Karlsruhe

Tag der mündlichen Prüfung: 10.12.2021

Referent:	Prof. Dr. Ulrich Husemann	Institut für Experimentelle Teilchenphysik
Korreferent:	Prof. Dr. Thomas Müller	Institut für Experimentelle Teilchenphysik

Erklärung der selbstständigen Anfertigung der Dissertationsschrift
Hiermit erkläre ich, dass ich die Dissertation mit dem Titel

Measurements of the $t\bar{t}H(b\bar{b})$ Process with Neural Networks at the CMS Experiment - Current Status and Future Prospects.

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, 24.11.2021

.....
(**Philip Keicher**)

Contents

1 Introduction

High-energy physics aims at exploring the smallest components of nature, the elementary particles, as well as the underlying mechanisms of their interactions. These properties are described within the Standard Model of particle physics (SM). The last success of the SM was the prediction of the Higgs boson [1–5], the search for which being one of the main motivations for the construction of the Large Hadron Collider (LHC) at CERN in Switzerland. Using this accelerator and the detectors that are positioned along its circumference, it is possible to probe physics at unprecedented energy scales. This research resulted in the observation of a particle compatible with the expected properties of the Higgs boson with the ATLAS and CMS detectors in 2012 [6, 7].

While the predictions of the SM are compatible with observations at high-energy experiments all over the world, the SM does not include a description of gravity at the quantum level and cannot account for observations in other fields such as cosmology, where a strong evidence for non-radiative dark matter [8, 9] and dark energy [10] was observed. This necessitates a more fundamental theory that describes the mechanisms covered by the SM and physics beyond the SM (BSM). However, the absence of clear discrepancies of observed data and the SM predictions in controlled laboratory setups renders the formulation of such a new theory difficult. There are two general avenues scientists all over the world follow in the search for BSM physics: the direct search for New Physics (NP), e.g. new particles or interactions, and high-precision measurements of the SM predictions to search for deviations from the predicted properties.

Since the observation of the Higgs boson is relatively recent, there is still room for improvement in terms of accuracy in the measurements of its properties. The interactions of the Higgs boson with the other elementary particles are of special interest and are expected to be proportional to the particles' masses. Therefore, the research conducted in the scope of this thesis aims to contribute to the precise measurement of one of these interactions, the coupling to the top quark. The precise measurement of this interaction is crucial for understanding the validity of the SM [11, 12] and is expected to be sensitive to BSM physics [13–16].

While the decay of a Higgs boson into a top-quark-antiquark pair ($t\bar{t}$) is kinematically forbidden, the production of the Higgs boson in association with a $t\bar{t}$ pair ($t\bar{t}H$) provides the means to probe the top-Higgs coupling directly. Furthermore, the $t\bar{t}H$ process is expected to be sensitive to contributions of BSM physics as shown in Refs. [17, 18]. Consequently,

measurements involving the $t\bar{t}H$ production mode allow for the construction of powerful tests of the SM and beyond. The final state of the $t\bar{t}H$ process where the Higgs boson decays into a bottom-quark-antiquark pair ($t\bar{t}H(b\bar{b})$) is the most probable decay channel [19] and is thus very promising for statistical data analysis. Past measurements in this channel at the CMS experiment have seen evidence for this process with a statistical significance of 3.9 standard deviations [20, 21] and contributed crucial sensitivity to the observations of the $t\bar{t}H$ production [22] and the decay of the Higgs boson to a $b\bar{b}$-pair [23].

This thesis presents the $t\bar{t}H(b\bar{b})$ analysis with the full Run-II data set recorded at the CMS experiment corresponding to $137.5\,\text{fb}^{-1}$ in the semileptonic decay mode of the $t\bar{t}$ system. Apart from the challenges arising from the small production cross section of $\sigma_{t\bar{t}H(b\bar{b})} = 0.29\,\text{pb}$ [19] of this process, one of the crucial aspects of the analysis is the control of the production process of a $t\bar{t}$ pair in association with a bottom-quark-antiquark pair ($t\bar{t} + b\bar{b}$), which is an irreducible background in the analysis that is $\mathcal{O}(10)$ times more probable [24, 25]. This necessitates the utilization of sophisticated multivariate analysis techniques, such as neural networks which enhance the separation between the signal and background processes and yield a high sensitivity to the $t\bar{t}H(b\bar{b})$ process.

In the absence of significant deviations from the SM prediction above five standard deviations, novel techniques are needed in the search for NP. In this thesis, two approaches are of special interest. In the first approach, a generic extension to the formalism used in the SM is added to account for BSM contributions well above the currently accessible energy scale. This results in the Standard Model Effective Field Theory (SMEFT) [26–28], which involves many higher-dimension contributions to a given phase space and therefore requires a large amount of sensitivity to constrain them. The second approach exploits the high statistical precision of the Run-II data set and divides the available phase space according to different observables of the physical system. The definitions of these regions are synchronized between experimental searches and the theoretical calculations and are referred to as the Simplified Template Cross Section (STXS) technique [29, 30]. This provides a powerful framework for cross section measurements that can probe individual parts of the total phase space due to the differential information that is employed. In the scope of this thesis, the experimental sensitivity to NLO SMEFT contributions in the context of a measurement in the STXS framework is gauged in the $t\bar{t}H(b\bar{b})$ channel for the first time at the CMS experiment.

Furthermore, the imminent Run 3 and the planned upgrade of the LHC to the High Luminosity Large Hadron Collider (HL-LHC) [31, 32] is expected to yield an even larger amount of data. The studies conducted during this thesis contributed to the formulation of the European strategy for particle physics [33], and are reviewed in more detail here.

This thesis is divided into four parts. The principles this thesis relies upon are summarized in Part I. Specifically, this includes a summary of the theoretical concepts in Chap. 2, the experimental setup of the CMS detector in Chap. 3, the definition of the relevant physics processes in Chap. 4 and their simulation in Chap. 5 as well as the relevant methods of statistical data analysis in Chap. 6 and the methods of machine learning in Chap. 7 used for the statistical inference. Part II presents the results of the inclusive $t\bar{t}H(b\bar{b})$ analysis in the semileptonic channel (Chap. 8) and its contribution to the combined $t\bar{t}H(b\bar{b})$ analysis (Chap. 9). Novel analysis methods in the context of the $t\bar{t}H(b\bar{b})$ measurement are presented in Part III of the thesis, which consists of a cross section measurement in the STXS framework and first feasibility studies of a measurement in the scope of the SMEFT framework (Chap. 10) as well as projected sensitivities for future data sets (Chap. 11). Finally, the conclusions drawn from these studies are presented in Part IV.

Part I

Foundations

2 Theoretical background

The goal of this thesis is the inference of physical properties by the comparison of theoretical predictions and experimentally observed data. The theoretical foundations are briefly discussed in this chapter. First, the Standard Model of particle physics (SM) and its effective field theory extension (SMEFT) are introduced. Afterward, the aspects relevant for the description of proton-proton collisions are reviewed. This chapter is intended as a short overview of these topics. More thorough discussions and reviews can be found in Refs. [34–36], which are the basis for the summary presented in this chapter. This thesis is based entirely on the natural unit system, i.e. the speed of light and Plank's constant are equal to one.

2.1 The Standard Model of particle physics

This section introduces the relevant aspects of the SM. First, the SM is briefly summarized. Since this thesis focuses on an analysis in the Higgs sector, the Higgs mechanism [1–5] is reviewed in the last part of this section.

2.1.1 Introduction to the SM

The SM is a quantum field theory that describes and predicts the behavior and interactions of elementary particles. It expresses these particles in four-dimensional Minkowski space using the Euler-Lagrange formalism founded on the principle of least action. The SM is a powerful theory that is able to make predictions based on first principles like gauge symmetry, gauge invariance and renormalizibility.

The elementary particles are described as wave functions embedded into quantum fields. In general, they are divided into two groups based on their intrinsic spin. Fermions have spin 1/2 and follow the Fermi-Dirac statistics introduced in Ref. [38]. They can be further divided according to an additional degree of freedom called color, which is present for quarks and is absent for leptons. These particles are summarized in Tabs. 2.1 and 2.2, respectively.

The SM itself does not predict masses for neutrinos. However, the observation of the oscillation of neutrinos between the three lepton flavors presented in Refs. [39, 40] shows that the physical mass eigenstates of the neutrinos ν are not equivalent to the flavor eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$. This is clear evidence for the existence of neutrino masses. While

Table 2.1: **Charge and mass values of the quarks**. The blocks shown in the table correspond to the first, second and third generation of quarks. Up-type quarks are listed at the top of each block, while down-type quarks are at the bottom. The mass value of the top quark is extracted from direct measurements, while all other quark masses are calculated in the $\overline{\text{MS}}$ renormalization scheme. All values taken from Ref. [37].

Particle	Electric Charge in units of e	Mass
up quark u	$+2/3$	$2.16^{+0.49}_{-0.26}$ MeV
down quark d	$-1/3$	$4.67^{+0.50}_{-0.17}$ MeV
charm quark c	$+2/3$	$1.27^{+0.02}_{-0.02}$ GeV
strange quark s	$-1/3$	93^{+11}_{-5} MeV
top quark t	$+2/3$	$172.76^{+0.30}_{-0.30}$ GeV
bottom quark b	$-1/3$	$4.18^{+0.03}_{-0.02}$ GeV

Table 2.2: **Charge and mass values of the leptons**. The blocks shown in the table correspond to the first, second and third generation of leptons. The masses of the neutrinos are unknown at this time. Therefore, the current limits are shown in the table. The uncertainties on the masses of electrons (muons) are of order $\mathcal{O}(10^{-9})$ ($\mathcal{O}(10^{-6})$) MeV and are therefore not shown here. All values taken from Ref. [37].

Particle	Electric Charge in units of e	Mass in MeV
electron neutrino ν_e	0	$<1.10 \times 10^{-6}$
electron e	-1	0.51
muon neutrino ν_μ	0	<0.19
muon μ	-1	105.66
tau neutrino ν_τ	0	<18.20
tau τ	-1	1776.86 ± 0.12

Table 2.3: **Charge and mass values of the bosons**. Values for the Higgs boson are taken from Ref. [45]. All other values are taken from Ref. [37].

Particle	Electric Charge in units of e	Spin	Mass in GeV
gluon g	0	1	massless
photon γ	0	1	massless
W boson	± 1	1	80.379 ± 0.012
Z boson	0	1	91.188 ± 0.002
Higgs boson H	0	0	125.38 ± 0.14

the measurement of the neutrinos' properties is a very active field of research, it is not essential for this thesis and will therefore not be further discussed here. More information can be found in Refs. [37, 41–44].

Elementary particles with an integer spin are referred to as bosons and are described using Bose-Einstein statistics [46]. Their relevant properties are summarized in Tab. 2.3. The listed particles mediate the fundamental forces in the particle interactions, which are further reviewed in the following.

Currently, the interactions between particles are characterized by four fundamental interactions: the strong, the electromagnetic and the weak force as well as gravitation. The SM describes all of these forces except for gravitation, which is currently not described at the quantum level. However, since this force is much smaller than the other three at the scales relevant for high-energy physics, its effects are negligible in the scope of this thesis and will not be further discussed.

The strong force is described by quantum chromodynamics (QCD). This theory is represented by the non-Abelian SU(3) gauge group [47, 48] and is constructed using the Yang-Mills theorem [49]. Quarks transform in the fundamental representation of this group, which has three dimensions corresponding to the three color charges red, blue and green. Analogously, gluons transform in the adjoint representation of the SU(3) group, which has eight dimensions. Consequently, there are eight mediators that carry a linear combination of color and anti-color charges themselves, thus enabling self-interaction of these particles. This results in a limited range of QCD interactions. In a semi-classical approach, this range is described by the potential

$$V(r) = \frac{4}{3}\frac{\alpha_\mathrm{s}}{r} + k \cdot r,$$

where r is the distance between two color-charged objects, α_s is the fine-structure constant of the strong interaction and k is a constant of the order of $1\,\mathrm{GeV\,fm^{-1}}$. For large distances or small energies, the potential energy becomes sufficiently large to generate more color-charged particles. As a consequence, only color-neutral objects can propagate freely, which are referred to as hadrons and consist of either three elementary particles with different colors (baryons) or a color-anti-color pair (mesons). This behavior is known as confinement. On the other hand, the coupling strength α_s itself depends on the energy. For small distances and high energies, α_s is small and the color-charged particles are quasi-free. This phenomenon is referred to as asymptotic freedom [50, 51].

Additionally, the SM includes the electroweak (EWK) sector [52–54], which is described by the $\mathrm{SU(2)}_L \times \mathrm{U(1)}_Y$ gauge group. In this theory, the left-handed fermions transform under the $\mathrm{SU(2)}_L$ group as doublets and the right-handed leptons as singlets:

$$\begin{pmatrix} \mathrm{u}_i \\ \mathrm{d}'_i \end{pmatrix}_L, \quad \begin{pmatrix} \nu_i \\ l_i \end{pmatrix}_L, \quad (\mathrm{u}_i)_R, \quad (\mathrm{d}_i)_R, \quad (l_i)_R,$$

where the index i denotes the previously-introduced generations of fermions and d' the eigenstates of the down-type quarks in the flavor basis. These states are superpositions of the physical mass eigenstates following

$$\mathrm{d}'_i = \sum_{ij} V_{ij} \mathrm{d}_j,$$

where V denotes the Cabibbo-Kobayashi-Maskawa (CKM) matrix [55, 56] that expresses the transition probabilities between the quark generations. The diagonal elements of this matrix are close to one, while the off-diagonal elements are much smaller [37], indicating the suppression of transitions between the generations. This has significant implications on some of the experimental techniques explained in Chap. 3.

The $\mathrm{SU}(2)_L$ introduces three gauge bosons $\mathrm{W}^a, (a = 0, 1, 2)$. Within the EWK theory, the coupling strength is denoted as g and the weak isospin I_Z is introduced as a quantum number. Note that right-handed neutrinos have not been observed yet and are not part of this formalism. However, the observation of the neutrino mass may indicate their existence and right-handed neutrinos can be introduced by extensions to the SM. All fermions transform under the $\mathrm{U}(1)_Y$ symmetry, which introduces the gauge boson B^0. The corresponding coupling strength is denoted as g' and the quantum number Y is referred to as hypercharge.

The EWK theory in this form violates unitarity bounds at tree-level and is non-perturbative above approximately 1 TeV. Moreover, the four gauge bosons it introduces are massless, which is in contradiction with the observations of the massive W and Z bosons. These properties necessitate the introduction of a mechanism that generates the masses of the bosons and introduces a scalar boson, which is realized in the Higgs mechanism.

2.1.2 The Higgs mechanism

The Higgs mechanism introduces a complex doublet φ with hypercharge $Y = 1/2$, referred to as Higgs multiplet, and four degrees of freedom. The Lagrangian for this newly-introduced field is

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= \mathcal{L}_{\text{Higgs,kin}} - V(\varphi) \\ &= (D^\mu \varphi)^\dagger (D_\mu \varphi) + \mu^2 |\varphi|^2 - \lambda |\varphi|^4, \qquad \text{with} \quad \varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}. \end{aligned} \tag{2.1}$$

In this equation, the covariant derivative

$$D_\mu \varphi = \partial_\mu \varphi - \frac{1}{2} i g \mathrm{W}^a_\mu \sigma^a \varphi - \frac{1}{2} i g' B^0_\mu \varphi$$

is used for the gauge-invariant description of the kinetic energy of the Higgs multiplet, with the previously introduced fields W^a and B^0, the corresponding coupling strengths g and g' and the Pauli matrices σ^a, which are the generators of the $\mathrm{SU}(2)_L$. The potential $V(\varphi)$ introduces the vacuum expectation value (VEV) $v = \frac{\mu}{\sqrt{\lambda}}$. For values $v \neq 0$, the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry is spontaneously broken based on the Goldstone theorem [57,

58]. The VEV can be taken to be real and in the lower component of the multiplet without loss of generality. Thus, φ can be expanded as

$$\varphi = \exp\left(i\frac{\pi^a\sigma^a}{v}\right)\begin{pmatrix}0\\ \frac{v}{\sqrt{2}}+\frac{\mathrm{H}(x)}{\sqrt{2}}\end{pmatrix},$$

with the real scalar field H and the phase π. Due to the principle of gauge invariance, the phase of the multiplet can be chosen as $\pi = 0$. The kinetic part of Eq. (2.1) involving the EWK bosons can then be written as

$$\mathcal{L}^{\mathrm{EWK}}_{\mathrm{Higgs,\,kin}} = g^2\frac{v^2}{8}\left[\left(\mathrm{W}^1_\mu\right)^2+\left(\mathrm{W}^2_\mu\right)^2+\left(\frac{g'}{g}B^0_\mu-\mathrm{W}^0_\mu\right)^2\right],$$

corresponding to three massive bosons. This equation can be diagonalized with

$$\begin{pmatrix}\mathrm{Z}_\mu\\ \gamma_\mu\end{pmatrix} = \begin{pmatrix}\cos(\theta_W) & \sin(\theta_W)\\ -\sin(\theta_W) & \cos(\theta_W)\end{pmatrix}\begin{pmatrix}\mathrm{W}^0_\mu\\ B^0_\mu\end{pmatrix}, \quad \mathrm{W}^\pm_\mu = \frac{1}{\sqrt{2}}\left(\mathrm{W}^1_\mu \mp \mathrm{W}^2_\mu\right),$$

where θ_W is the Weinberg angle and the resulting fields correspond to the vector bosons introduced in Tab. 2.3. In the end, the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry at high energies is spontaneously broken to the electromagnetic $\mathrm{U}(1)_{\mathrm{EM}}$ symmetry described by quantum electrodynamics (QED) [59–61] and the 4-Fermi theory [62] at low energies.

The Lagrangian describing the couplings of the bosons to the Higgs field can then be written as

$$\begin{aligned}\mathcal{L}^{\mathrm{boson}}_{\mathrm{Higgs}} = &-\frac{1}{2}\mathrm{H}(\Box+\mathrm{m}^2_\mathrm{H})\mathrm{H} - g\frac{\mathrm{m}^2_\mathrm{H}}{4\mathrm{m}_\mathrm{W}}\mathrm{H}^3 - \frac{g^2}{32}\frac{\mathrm{m}^2_\mathrm{H}}{\mathrm{m}^2_\mathrm{W}}\mathrm{H}^4\\ &+2\frac{\mathrm{H}}{v}\left(\mathrm{m}^2_\mathrm{W}\mathrm{W}^{\mu+}\mathrm{W}^-_\mu+\frac{1}{2}\mathrm{m}^2_\mathrm{Z}\mathrm{Z}^2_\mu\right)+\left(\frac{\mathrm{H}}{v}\right)^2\left(\mathrm{m}^2_\mathrm{W}\mathrm{W}^{\mu+}\mathrm{W}^-_\mu+\frac{1}{2}\mathrm{m}^2_\mathrm{Z}\mathrm{Z}^2_\mu\right),\end{aligned} \tag{2.2}$$

where $\Box$ denotes the d'Alembertian operator and

$$\mathrm{m}_\mathrm{W} = \frac{v}{2}g \quad \text{and} \quad \mathrm{m}_\mathrm{Z} = \frac{v}{2}\sqrt{g^2+(g')^2} = \frac{\mathrm{m}_\mathrm{W}}{\cos(\theta_W)}.$$

Consequently, the masses m of the particles emerge within the Higgs mechanism and the coupling of the bosons to the Higgs field H is proportional to the squared masses.

The interactions of the Higgs boson and the fermions after the symmetry breaking are described by Yukawa-type interactions with

$$\mathcal{L}_{\mathrm{Yukawa}} = \sum_i y_i\frac{v}{\sqrt{2}}\psi_i\left(1+\frac{\mathrm{H}}{v}\right)\bar{\psi}_i = \sum_i \mathrm{m}_i\psi_i\left(1+\frac{\mathrm{H}}{v}\right)\bar{\psi}_i, \tag{2.3}$$

where i indexes the fermion flavors, ψ_i denotes the fermion spinors, m_i the fermion masses and y_i is the Yukawa coupling for the respective fermions. As Eq. (2.3) shows, the coupling strength of the fermions to the Higgs field is expected to be proportional to the fermion mass.

As indicated by Eq. (2.3), the strongest coupling is expected with the heaviest particles. Therefore, the measurement of the top-Higgs-Yukawa coupling is especially suited to construct powerful tests of the SM. While the direct decay of the Higgs boson into two top quarks is kinematically forbidden, the production of the Higgs boson in association with a top-quark-anti-quark pair ($\mathrm{t\bar{t}H}$) can be used as a direct probe for the coupling, which is the focus of this thesis. The following explanations will therefore be given in the context of the $\mathrm{t\bar{t}H}$ process.

Table 2.4: **SMEFT operators of interest for this thesis**. Shown are the operator names used throughout this thesis, the corresponding parameter in the Warsaw basis and the Lagrangian used to extend the SM. Definitions taken from Ref. [65]

Operator	Warsaw Basis	Lagrangian
$\mathcal{O}_{t\varphi}$	$\mathcal{O}_{uH}$	$\left(\varphi^\dagger\varphi - \frac{v^2}{2}\right)\bar{Q}t\tilde{\varphi} + h.c.$
$\mathcal{O}_{tG}$	$\mathcal{O}_{uG}$	$igs\left(\bar{Q}\frac{1}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)T_A t\right)\tilde{\varphi}G^A_{\mu\nu} + h.c.$
$\mathcal{O}_{\varphi G}$	$\mathcal{O}_{HG}$	$\left(\varphi^\dagger\varphi - \frac{v^2}{2}\right)G^{\mu\nu}_A G^A_{\mu\nu}$

2.2 Standard Model Effective Field Theory

As indicated in the review of the SM, the theory shows some shortcomings such as the correct description of neutrinos. Moreover, there are observations in other fields of physics such as the measurement of the rotational speed of galaxies [8] and the analysis of the mass distribution of the Bullet cluster [9], that are attributed to non-radiative dark matter not described by the SM. These observations are evidence for physics beyond the SM (BSM). This gives rise to two strategies employed in the search for BSM physics: the direct search for New Physics, e.g. outlined in Ref. [63], or the high-precision measurements which aim to find deviations from the SM predictions.

At the time of writing this thesis, no convincing deviation above five standard deviations from the SM prediction has been found. The absence of indications for new particles suggests that the energy scale for BSM physics might be significantly higher than the EWK scale as characterized by the masses of the heaviest particles in the SM or the Higgs VEV. The low-energy effects of the high-energy BSM physics can then be approximated by integrating-out the BSM particles to obtain higher-dimension interactions as described in Ref. [64]. Being approximations, these effective field theories (EFT) are valid only for energies far below the characteristic energy scale of new physics, which is denoted as Λ in the following. This is also apparent in the corresponding Lagrangian, in which the higher-dimensional operators need to be normalized by powers of Λ to be compatible with the four-dimensional Minkowski space:

$$\mathcal{L}_{\text{EFT}} = \sum_{d=1}\sum_{o}\frac{c_o}{\Lambda^d}\mathcal{O}_o^{4+d}, \tag{2.4}$$

where d denotes the additional dimensions to four-dimensional space-time and o the possible operators for the corresponding dimension. The contribution of the respective operators is quantified with the dimensionless factor c_o. As indicated by Eq. (2.4), higher-dimensional operators are increasingly suppressed as a function of the energy scale of BSM physics Λ. The EFT framework is a powerful tool to interpret observed data from different experiments and can be used to construct new theories by matching the coefficients of the corresponding operators at different dimensions.

In this thesis, the Standard Model Effective Field Theory (SMEFT) is used. The operators defined in Eq. (2.4) are added to the SM Lagrangian, thus giving the theory more flexibility to account for BSM contributions. Thorough reviews of SMEFT can be found in Refs. [26–28]. Since the contribution of the operators decreases with their dimension, most efforts concentrate on the leading EFT operators of up to six dimensions. Due to requirements such as gauge invariance, there is one dimension-five operator, referred to as the Weinberg operator, which violates the conservation of the lepton number by two units [66] and is therefore explicitly excluded from most analyses in high-energy physics. The remaining

dimension-six operators can be expressed in different bases. For this thesis, the Warsaw basis is used, which is discussed in more detail in Ref. [67]. As shown in Ref. [17], the $\mathrm{t\bar{t}H}$ process is expected to increase the sensitivity to the contributions of three operators in particular, which have recently become available for automated calculation as described in Refs. [65, 68]. The corresponding extensions to the SM Lagrangian are listed in Tab. 2.4 and are briefly introduced in the following.

The operator $\mathcal{O}_{t\varphi}$ rescales the top-Higgs Yukawa coupling in the SM and also contributes to the pairwise Higgs production in association with a $\mathrm{t\bar{t}}$ pair. $\mathcal{O}_{tG}$ represents the chromomagnetic dipole moment of the top quark, which modifies the $\mathrm{gt\bar{t}}$ vertex and adds new four-point vertices $\mathrm{ggt\bar{t}}$ and $\mathrm{gt\bar{t}H}$ as well as the five-point interaction $\mathrm{ggt\bar{t}H}$. The operator $\mathcal{O}_{\varphi G}$ introduces additional loop-induced interactions between the gluon and the Higgs fields. The operators are consistently implemented in the `SMEFTatNLO` tool presented in Ref. [68] and further described in Chap. 5.

The contributions of these operators are compared to the expected sensitivity of the measurement presented in this thesis. For this comparison, the current constraints for the operators are used as presented in Ref. [69]. This allows to gauge the potential gain in constraining power this analysis can yield to the currently available sensitivity, which is further described in Sec. 5.3.

2.3 Physics in proton-proton collisions

After briefly discussing the underlying theoretical concepts, some concepts of the calculations needed for this thesis are introduced here. This thesis presents an analysis of data recorded at the Large Hadron Collider (LHC), where protons are collided at the highest energies available in the laboratory today. Consequently, the LHC belongs to the class of hadron colliders, which will be the focus of this section. The following explanations introduce general concepts relevant to this thesis. More thorough reviews of collider physics can be found in Refs. [37, 70].

Particle collisions are generally characterized by two quantities. The first is the center-of-mass energy $\sqrt{s}$ which denotes the energy available during the collision. At collider experiments such as the LHC, the center-of-mass energy equates to $\sqrt{s} = 2E$, where E is the energy of each of the colliding objects, in this case the protons. The other quantity is the number of collision events, which is calculated as

$$N = \sigma_X \cdot L_{\mathrm{int}} = \sigma_X \cdot \int_t^{t+\Delta t} \mathrm{d}t' L_{\mathrm{inst}}, \tag{2.5}$$

where Δt is the time span of data taking. In this equation, σ_X denotes the cross section, which is a measure of the probability of a given final state X after the collision, while the instantaneous luminosity L_{inst} quantifies the particle flux density during the collisions. The integrated luminosity L_{int} is a measure of the total amount of recorded data. More information about the luminosity is< given in Chap. 3.

Being hadrons, protons are composite particles consisting of some of the elementary particles introduced in Sec. 2.1, which are referred to as partons in this context. For sufficient center-of-mass energies $\sqrt{s}$, these partons are asymptotically free and are in fact the objects that participate in the collision. This has to be accounted for in the calculation of the cross section σ_X, which can be written as

$$\sigma_X = \sum_{ij} \int \mathrm{d}x_i \mathrm{d}x_j \mathrm{PDF}\left(x_i, \mu_{\mathrm{F}}^2\right) \mathrm{PDF}\left(x_j, \mu_{\mathrm{F}}^2\right) \sigma_{ij\to X}\left(x_i p_1, x_j p_2, \mu_{\mathrm{F}}^2, \mu_{\mathrm{R}}^2\right). \tag{2.6}$$

The indices i and j denote the partons originating from the two protons, respectively. The factor x is referred to as Bjorken scaling variable [71] and quantifies the fraction of the given parton momentum with respect to the total momentum of the corresponding proton, which is denoted as p_1 or p_2 for the two separate protons. The parton distribution function (PDF) is the probability density function for the given parton as a function of its momentum fraction x, which cannot be calculated from first principles and therefore has to be extracted from data. Finally, the cross section $\sigma_{ij\to X}$ quantifies the probability for the final state X given the partons i and j, and is a function of the parton momenta as well as the renormalization scale μ_R and factorization scale μ_F, which are introduced shortly.

The cross section σ_{ij} can be calculated using Fermi's golden rule, which requires a calculation of the available phase space and the matrix element $|\mathcal{M}|^2$. The latter can be calculated using the Lagrange formalism and perturbation theory if the relevant couplings are small enough. Consequently, cross section calculations are classified according to the order in perturbation theory they are calculated in, i.e. as leading order (LO), next-to-leading order (NLO), etc. Using perturbation theory gives rise to two characteristic energy scales. Since the calculation stops at a given order, it is possible that the resulting cross section contain unphysical divergences. In order to ensure a meaningful interpretation of the theory predictions, such divergences can be eliminated by renormalizing the calculation, which is characterized by the renormalization scale μ_R. Furthermore, since the coupling constants in the underlying theory are energy-dependent themselves, the perturbative calculation is only valid until a certain threshold. The non-perturbative parts of the calculation factorize at this characteristic energy scale, referred to as the factorization scale μ_F, and are absorbed into the PDFs which are extracted from observed data. More details about the concepts of renormalization and factorization scales can be found in Refs. [34, 72–74]. The matrix element $\mathcal{M}$ describes the physics of the hard interaction of the initial partons and can be translated into Feynman diagrams using the rules outlined in Ref. [61, 75], which allows for a well-defined depiction and calculation of the physics processes.

These calculations are very successful due to their high predictive power. However, they can entail various mathematical challenges that arise due to effects such as gluon splitting functions, which are difficult to calculate and can lead to additional divergences. There are different techniques to avoid the ensuing deviations in the calculation, one of which being the exact definition of what to absorb into the PDFs as previously mentioned. In the scope of this thesis, two of these schemes are used. For matrix elements derived in the five-flavor scheme (5FS), only the top quark is considered a heavy quark, while the masses of the other quarks are negligible. The corresponding contributions to the cross section calculation associated with the kinematics of the light quarks are factorized into the PDF set, thus yielding the calculation of inclusive cross sections stable without a loss in accuracy. Matrix elements calculated in the four-flavor scheme (4FS) consider the top quarks and the bottom quarks as heavy and account for them directly in the calculation. As previously mentioned, these calculations are generally harder but are expected to yield a more accurate description of the kinematic behavior of bottom quarks than derivations in the 5FS. A comparison of these calculation can be found in Ref. [76].

The final state particles themselves can in turn emit additional radiation, which triggers a chain of emissions depending on the energy of the particles. This parton shower is modeled by a stepwise parton splitting criterion until the remaining energy is too low for perturbation theory to be applicable. As a result, it is possible to create the same final state from LO matrix elements paired with parton shower emissions and NLO matrix elements that directly consider additional radiation. To avoid double-counting of these final states, the matrix element calculation needs to be matched to the parton shower, which is referred as ME-PS matching and is implemented in the tools described in Chap. 5.

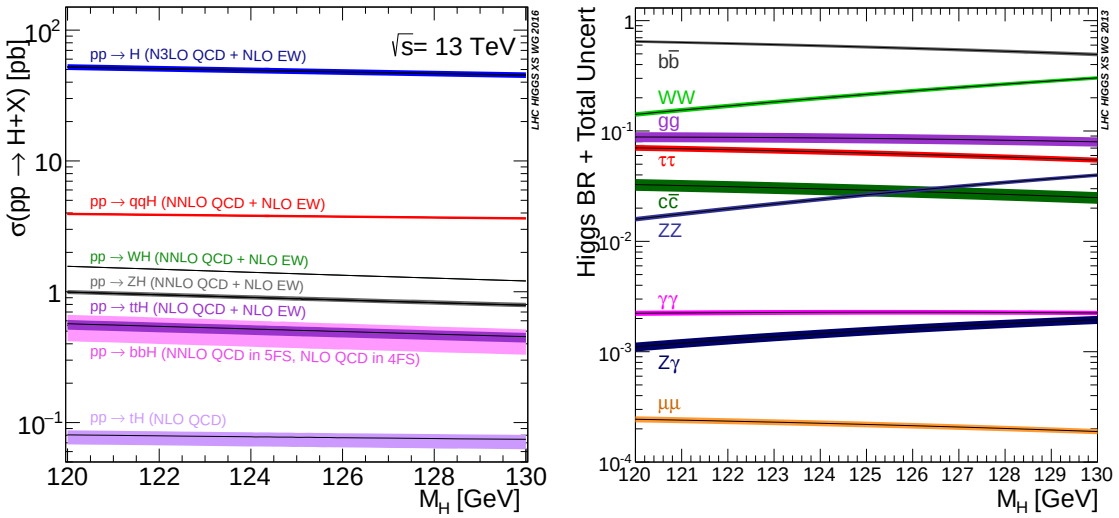


Figure 2.1: **Higgs production cross sections and branching ratios**. Shown are the production cross sections at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ (left) and the branching ratios of the different Higgs decay modes (right) as a function of the Higgs boson mass. Figures taken from Ref. [19].

As outlined in the discussion about QCD in Sec. 2.1, there are no free color-charged particles due to the principle of confinement. Consequently, these particles emit radiation until they are all bound in color-neutral hadrons, a process that is referred to as hadronization. The energies of the particles at this stage are now in a regime where non-perturbative effects play an important role. Therefore, this process is modeled with phenomenological models implemented in the tools referred to later in this thesis.

The partons participating in the original collision are removed from the bound proton state, which is thus no longer color-neutral. As a result, these particles hadronize or otherwise interact with each other, which yields additional hadronic activity in the event. This activity is referred to as underlying event, which is described by multiple parton-parton interactions within one collision. These interactions are calculated using perturbative QCD where possible or by phenomenological models where non-perturbative effects cannot be neglected.

The characteristic time scales for the interactions involved in these steps are much longer than the time scales of the involved decays. Therefore, the long-distance and short-distance interaction factorize in the calculation of the cross section, which is referred to as the factorization theorem [77]. Nowadays, the chain of calculations outlined by the discussions of the matrix element, the parton shower, the hadronization and the underlying event is implemented in dedicated tools based on the Monte Carlo technique, which are further discussed in Chap. 5.

2.4 Cross sections and branching ratios for the Higgs boson

Using the concepts introduced in Sec. 2.3, it is now possible to characterize the probabilities for the different production channels of the Higgs boson. The production cross sections at the currently available fixed order in perturbation theory are shown in Fig. 2.1. Additionally, the figure shows the branching ratios of the Higgs boson, which are the probability for a given decay mode. These figures will be used to define the relevant processes for this thesis in Chap. 4.

Table 2.5: **SMEFT contributions to the total cross section**. Listed are the NLO contributions to the total cross section defined in Eq.(2.7) for a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. All values taken from Ref. [17]

Contribution	Value in fb
σ_{SM}	507.1
$\sigma_{t\varphi}$	−62.0
$\sigma_{\varphi G}$	872.0
σ_{tG}	503.0
$\sigma_{t\varphi,t\varphi}$	1.9
$\sigma_{\varphi G,\varphi G}$	1021.0
$\sigma_{tG,tG}$	674.0
$\sigma_{t\varphi,\varphi G}$	−53.0
$\sigma_{t\varphi,tG}$	−31.0
$\sigma_{\varphi G,tG}$	859.0

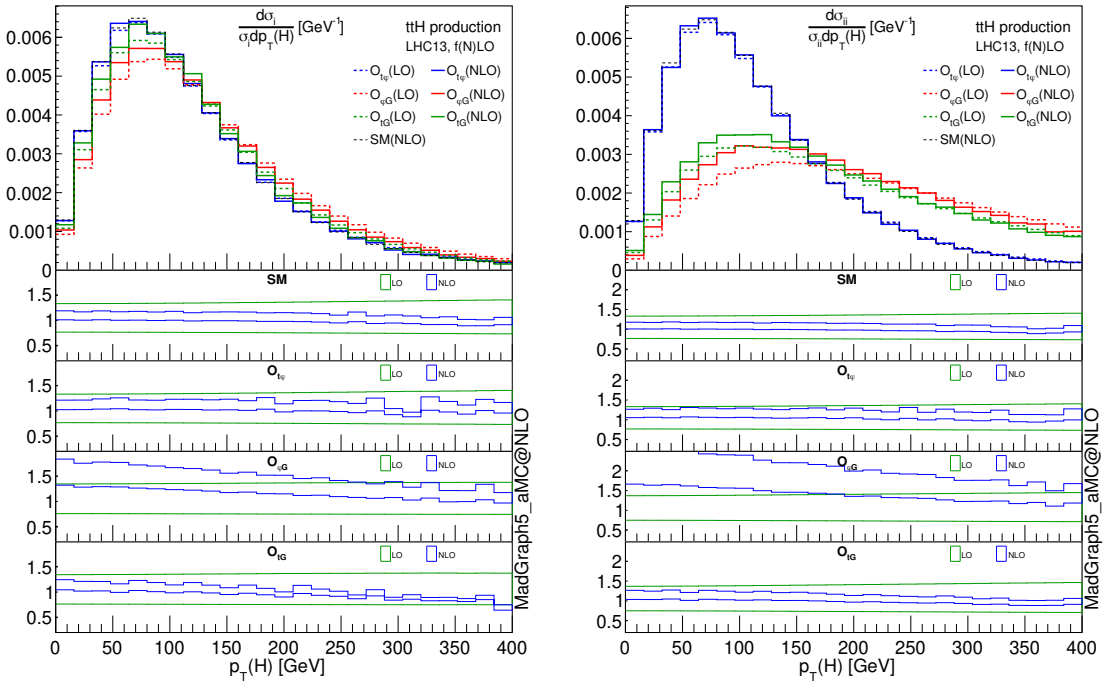


Figure 2.2: **Contributions of the SMEFT operators to the Higgs momentum distribution**. Shown are the contributions as introduced in Eq. (2.7) normalized to unity. The effect of the operators of interest is shown as the lines in the corresponding color at leading-order (dashed line) and next-to-leading-order (solid line) perturbation theory. The lower panels show the ratio between the LO (green) and NLO (blue) predictions and the associated $\mu_{\text{R}}/\mu_{\text{F}}$ uncertainties. Figure taken from Ref. [17].

Moreover, it is now possible to further quantify the contributions to the $t\bar{t}H$ channel originating from the higher-dimensional operators introduced by the SMEFT framework as discussed in Sec. 2.2. As shown in Ref. [17], the total cross section can be calculated with

$$\sigma = \sigma_{\mathrm{SM}} + \sum_i c_i \cdot \sigma_i + \sum_{ij} c_i \cdot c_j \cdot \sigma_{ij}, \tag{2.7}$$

where i and j denote the three operators of interest and the interference effects between the operators are accounted for by the contributions with $i \neq j$. For convenience, the contributions are listed in Tab. 2.5 and are shown in Fig. 2.2. As indicated by the figure, especially the σ_{ii} contributions in Eq. (2.7) of the monochromatic dipole moment and the effective gluon-Higgs coupling are expected to modify the shape of the transverse momentum of the Higgs boson $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum for high values. This observation is important for the comparison to the measurement presented in this thesis as shown in Sec. 10.4.

3 Experimental setup

After the review of the theoretical foundations in Chap. 2, the experimental setup is briefly discussed in this chapter. The experimental data used for the statistical inference presented in this thesis is recorded with the Compact Muon Solenoid (CMS) detector [78, 79] which is situated at the Large Hadron Collider (LHC). Both are introduced in the following.

First, the LHC and the acceleration of the protons is reviewed. Subsequently, the geometry and the components of the CMS detector are explained, before the experimental techniques and definitions used as a basis for this thesis are summarized. The final part of this chapter briefly reviews the planned upgrades of the LHC and the CMS detector. The information presented in this chapter is based on Refs. [33, 78, 80–83] as well as the *Accelerator Physics of Colliders* review in Ref. [37].

3.1 The Large Hadron Collider

The LHC is a circular collider situated at the European Organization for Nuclear Research (CERN). It is a synchrotron designed to accelerate protons as well as heavy ions, which are used in collision experiments located at dedicated points along the ring. The LHC is the last part in the accelerator chain shown in Fig. 3.1, which is briefly discussed in the following. Since this thesis considers data collected in proton-proton collisions, the discussion focuses on these particles.

The protons are created by removing the electrons from hydrogen atoms using a strong electric field. Subsequently, they are injected into the linear accelerator (LINAC), which accelerates the particles to approximately 50 MeV. In the following steps, the protons are accelerated in circular machines, to 1.4 GeV in the Booster, 25 GeV in the Proton Synchrotron (PS) and 450 GeV in the Super Proton Synchrotron (SPS), respectively, before they are finally injected into the LHC ring. During this injection, the protons are separated into two beams that are accelerated into opposite directions in separate pipes with ultra-high vacuums. The particles are accelerated to their final energy of 6.5 TeV, thus yielding a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ for the proton-proton collisions.

The acceleration in the circular accelerators is realized by employing high-frequency cavities that generate strong, oscillating electric fields. The frequency of these fields is adjusted depending on the current particle energy. As a consequence of this technique, the particles are arranged in packages within each beam, which are referred to as bunches. The energy

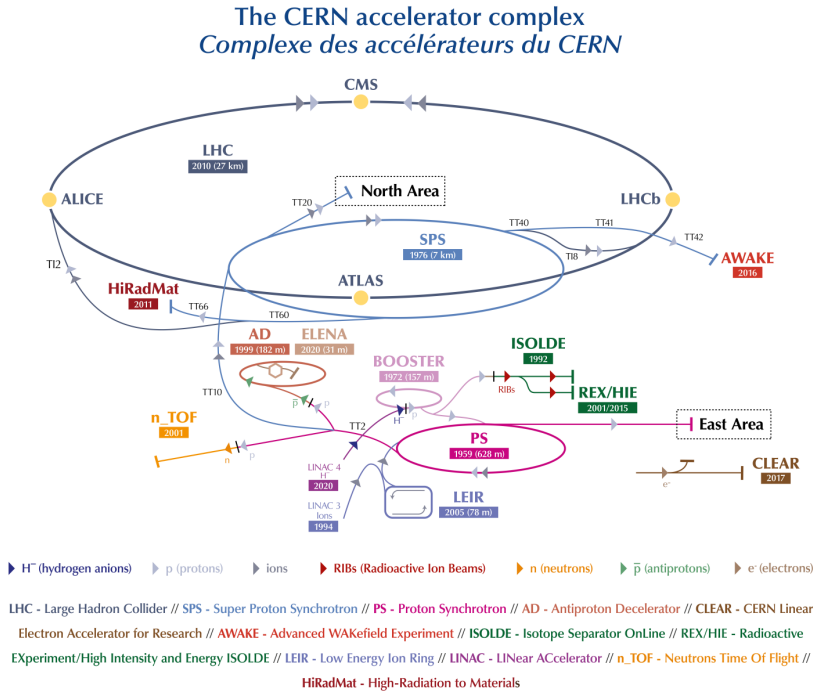


Figure 3.1: **Accelerator complex at CERN**. Shown is the accelerator chain the particles traverse before their final injection into the LHC ring. Figure taken from Ref. [84].

transferred to the protons both gradually accelerates them as well as replenishes the kinetic energy lost due to radiative emissions common in synchrotrons. The cavities in the LHC operate at a final frequency of 40 MHz, which dictates the time between collisions of 25 ns.

The LHC is not perfectly circular but is a combination of straight and bent parts distributed across the whole circumference of the ring. Segments with curvature house dipole magnets with fields of up to 8 T that deflect the particles onto the approximately-circular design trajectory. The straight sections contain quadrupole or higher-order magnets to focus and stabilize the beams. The magnets at the LHC are superconductors operated at approximately 2 K. This is sufficient to dissipate the heat originating from the synchrotron radiation of roughly 7 kW and to accommodate the large electric currents needed for the strong magnetic fields.

After the particles reach their final energies, they are brought to collision at dedicated points along the LHC. As already mentioned in the discussion of the calculation of the number of collision events presented in Eq. (2.5), the particle flux at these points is characterized by the instantaneous luminosity

$$L_{\text{inst}} = f \frac{n_1 n_2}{4\pi \sigma_x^* \sigma_y^*} \mathcal{F}, \tag{3.1}$$

where f is the collision frequency of 40 MHz, n_1 and n_2 denote the particle numbers in the colliding proton bunches and σ_x^* and σ_y^* characterize the transverse beam sizes in the horizontal and vertical directions. $\mathcal{F}$ is a factor of order one introduced to account for losses arising from geometric effects, like an angle between the trajectories of the crossing bunches and the finite bunch length, and dynamic effects, such as the mutual focusing of the two beams during the collision. The collisions are recorded with specialized detectors

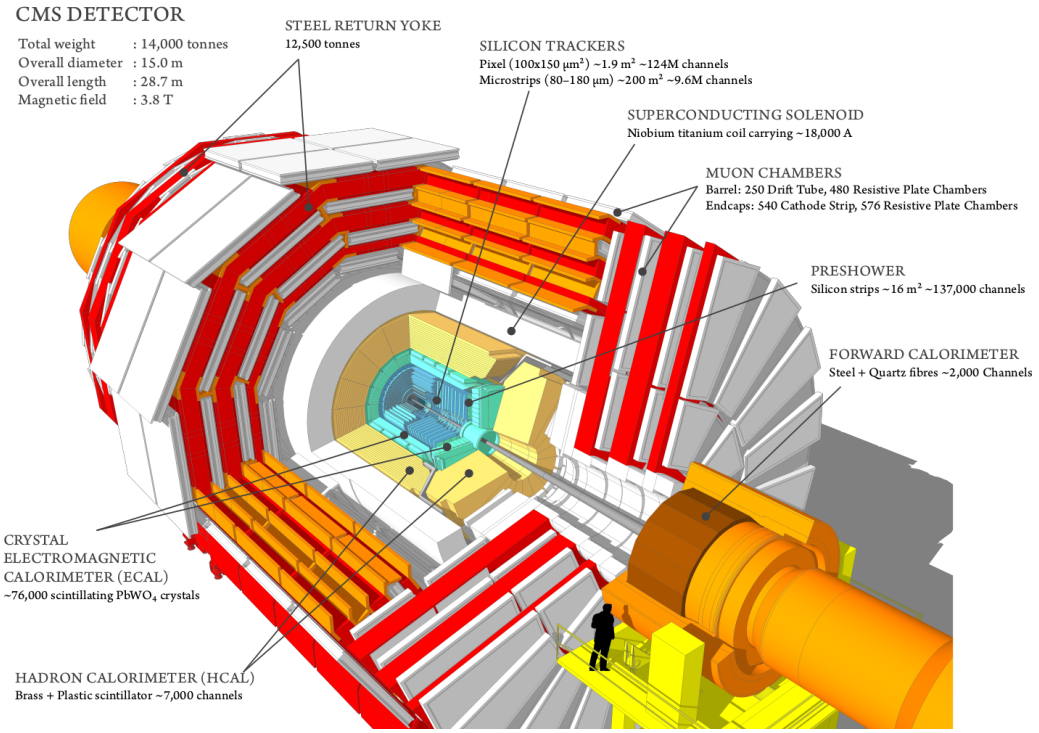


Figure 3.2: **Overview of the CMS detector**. Shown is the schematic overview of the different components of the CMS detector after the upgrade of the pixel detector in 2017. Figure taken from Ref. [85].

located at the collision points. One of these is the CMS detector, which is discussed in the following section.

3.2 The Compact Muon Solenoid Detector

The detector at the CMS experiment is a multi-purpose detector suitable for a broad range of studies of the physics described by the Standard Model (SM) as well as possible contributions beyond. The data analyzed in this thesis was recorded with this machine in the years between 2016 and 2018. Therefore, the detector and its components as shown in Fig. 3.2 are introduced here. First, the detector geometry and relevant observables are introduced. Subsequently, the different detector systems are discussed following the trajectory of the particles emerging from the collisions at the center of the CMS detector. The introduction presented in this section is based on Ref. [78], which discusses the CMS detector in more detail.

3.2.1 Geometry of the detector

The geometry of the CMS detector is generally described within a right-handed Cartesian coordinate system, where the x-axis points towards the center of the LHC, the y-axis points vertically upward and the z-axis is parallel to the beam pipe containing the proton bunches. The origin of this system is in the center of the detector, which is also the nominal collision point for the bunches.

As discussed in Sec. 2.3, the objects participating in the collision events are not the accelerated protons themselves, but their respective partons. For the high energies used at the LHC, the components of the parton momenta perpendicular to the z-axis are negligible.

This momentum component is referred to as the transverse momentum in the following. However, since the exact state of the partons is unknown at the time of the collision, the momentum p_z of the colliding particles is not known. Therefore, observables that describe the collision and are invariant under Lorentz-transformations along the z-axis are of special importance. The most important observables used throughout this thesis are introduced in the following.

One of the important observables is the aforementioned transverse momentum defined as

$$p_{\mathrm{T}} = |\vec{p}_{\mathrm{T}}|, \quad \vec{p}_{\mathrm{T}} = (p_x, p_y, 0)\,. \tag{3.2}$$

Since the transverse momenta of the particles in the initial state are negligible and well-defined, the sum of the transverse momenta of the final state has to be zero. This is a central aspect in the reconstruction of the event, as will be used in Sec. 3.3.

In addition, the rapidity y defined by

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.3}$$

is used, where E denotes the particle energy and the difference between two rapidities is invariant under Lorentz-transformation along the z-axis. Additionally, it can be shown that for large energies E or negligible particle masses, y is equal to the pseudo rapidity, defined by

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{3.4}$$

where θ denotes the polar angle in the y-z-plane. Together with the azimuthal angle ϕ in the x-y-plane, the approximately Lorentz-invariant distance ΔR between two objects can be defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{3.5}$$

Here, $\Delta\phi$ and $\Delta\eta$ denote the difference in azimuthal angle and pseudo rapidity between the two objects, respectively. These definitions are used at different points throughout this thesis.

3.2.2 Detector components

The different subsystems of the CMS detector are discussed in this section. Each system is dedicated to the measurement of different properties of the particles emerging from the collision point at the center of the detector. These measurements are the basis for the reconstruction of the collision system. The components are discussed in order of increasing distance from the interaction point.

Tracking system

After the collision event, the emerging final state particles first traverse the tracking system of the CMS detector, which is further discussed in Refs. [86–88]. Due to the strong homogeneous magnetic field of 3.8 T permeating the CMS detector, the trajectory of charged particles is bent proportional to their momenta and charges. The tracking system is designed to exploit this behavior. It is constructed in layers around the interaction point that register the points where charged particles traverse, thus allowing for the reconstruction of the trajectories. In order to preserve the original trajectory, it is essential that these measurements happen with a minimal amount of interactions of the particles with the

detection material. Therefore, this detector system is closest to the interaction point and the material budget is especially low. The sensors in the tracking system of the CMS detector consist entirely of silicon, which yields a good compromise between the material budget and the radiation hardness required due to the hostile environment close to the beam pipe. Similar to the principle of diodes, these sensors are operated with high voltage in the reverse bias direction, which results in a large area depleted of charge carriers. As charged particle traverse this area, they induce a current in the sensor that is measured. More information about the detection principles used in silicon detectors can be found in Ref. [89].

The tracking system is arranged in cylindrical layers around the center of the CMS detector, referred to as the barrel region, that are complemented by disk-like structures in regions with high $|\eta|$ values, called the endcaps. This subdetector is radially divided into two parts. The layers closest to the interaction point belong to the pixel detector, which provides high granularity with pixel sensors capable of 2D position measurements. The high granularity of this system is required due to the large particle flux in this region and thus reduces the occupancy of the individual sensors. The pixel detector was upgraded in 2017, where an additional layer of pixel sensors was introduced and the material budget was further reduced, thus improving the detection capabilities as presented in Ref. [90]. The outer layers of the tracking system consists of silicon strip detectors. The reduced particle flux in this region allows for a sparser granularity, which is why this detector system consist of modules capable of 1D position measurements or double-sided modules capable of 2D measurements where necessary. A review of the performance of the tracking system can be found in Ref. [88].

The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [91, 92] engulfs the tracking system and is designed to measure the energy of electromagnetically-interacting particles, namely photons, electrons and to a small extent muons and charged hadrons. The detection principle is based on the interaction of these particles with matter, specifically the deposit of energy due to bremsstrahlung, the creation of electron-positron pairs and finally ionization processes. As high-energy charged particles enter the electromagnetic field of the nuclei in the detector material, they emit bremsstrahlung. If the emerging photons have enough energy, they convert to electron-positron pairs by interacting with the very same electromagnetic field. Both processes are characterized by the radiation length X_0, where the number of particles produced in these electromagnetic showers is proportional to the energy of the initial particle. The shower continues until the energy of the photons is not sufficient anymore to produce an e^+e^- pair or the probability for energy loss due to ionization is larger than for radiative losses.

The ECAL is a homogeneous calorimeter that consists entirely of lead-tungsten crystals that lead to large energy deposits due to its high density and small radiation length of $X_0 = 0.9\,\text{cm}$. Additionally, the material acts as detection material due to its scintillation properties, which converts the energy deposited by the particles in the electromagnetic shower into light via ionization processes. The emerging photons are directed out of the ECAL and detected with avalanche photodiodes in the barrel region and vacuum phototriodes in the endcap sections. Furthermore, a more granular preshower detector is installed in front of the endcap regions in order to resolve neutral pions with large $|\eta|$, which decay into two photons, improving the π^0/γ separation. The performance of the ECAL is reviewed in Refs. [93, 94].

The hadronic calorimeter

Charged and neutral hadrons are more probable to interact with the nuclei of the detection material directly, which is characterized by the interaction length λ_{int}. Since λ_{int} is usually much larger than the radiation length X_0, the construction of a system dedicated to the energy measurement of hadrons is necessary, which is referred to as hadronic calorimeter (HCAL) [95]. In order to ensure that a sufficient amount of energy is deposited in the confines of the calorimeter, alternating layers of brass and scintillation material are used. The high density of the brass layers with a corresponding short interaction length λ_{int} leads to hadronic showers, which also contain an electromagnetic component arising from electromagnetically-decaying particles, such as neutral π^0. The detection material shows a different response to the components of the hadronic showers, which are later corrected for in software. The detection principle in the scintillating layers of the HCAL is the same as introduced in the discussion for the ECAL. A detailed discussion of the HCAL can be found in Refs. [96–99].

The solenoid magnet

The calorimeters and the tracking system are surrounded by the solenoid magnet system [100] of the CMS detector, which creates the magnetic field of 3.8 T. The field is stabilized using an iron yoke, which also absorbs most particles that managed to traverse the calorimeter system. Additionally, the yoke is used to accommodate the muon system discussed in the following.

The muon system

The muons in the relevant energy regime are referred to as minimum ionizing particles, meaning that it is highly improbable for these particles to deposit much of their energy in the aforementioned detector systems. This enables them to traverse well into the iron yoke of the magnet system. Since only very few other particles are expected to reach this region of the detector, the muon system [101] is located here. The chambers embedded into the iron yoke employ three different technologies, depending on the area in the detector and the associated expected particle flux, and are designed to reconstruct muon trajectories and thus measure their transverse momentum p_{T}. More specifically, the muon system employs drift tube chambers in the central region, cathode strip chambers based on the multi-wire proportional chamber technology and resistive plate chambers in the $|\eta| < 1.6$ region. The latter of these technologies has a good time resolution and is therefore used for the trigger system, which is discussed in the following section. The performance of the muon system is reviewed in Ref. [102].

Trigger system

Thanks to the high frequency of bunch crossings, the LHC produces a vast amount of data. Since it is technically impossible to save all of the collision events due to limitations in read-out speed and available disk space, it is necessary to select events with potentially interesting physics during the data taking. This is realized in the trigger system [103, 104] of the CMS detector, which is divided into two stages further introduced in the following.

The first stage is referred to as the L1 trigger and is comprised of dedicated electronics employing the information of the calorimeter and muon systems to perform a simple reconstruction of electrons, photons, muons, τ leptons and jets. The L1 trigger system offers different operation modes to identify potentially interesting physics, which are referred to as menus. If the criteria defined by one of these menus are fulfilled by an event, the event is saved for further processing. This decision is made within a 4 µs time window, during

which the information of the individual subdetectors is stored in a pipeline. This way, the L1 trigger is able to reduce the initial rate of events of 40 MHz down to approximately 100 kHz.

The subsequent stage of the trigger system is the high-level trigger (HLT). Similar to the L1 trigger, the HLT offers different options to further classify events, referred to as trigger paths. The decisions defined for these paths are based on a reconstruction employing the same techniques used in the offline processing of the events, which are more sophisticated than the L1 reconstruction and which will be briefly discussed in Sec. 3.3. Using this procedure, the final event rate is further reduced to the order of 100–1000 Hz. The trigger system is further discussed in Refs. [105, 106].

3.3 Object reconstruction and experimental techniques

After the introduction of the different detector components and their basic functionality, the experimental techniques used for the object reconstruction as well as some additional definitions are briefly presented in this section. Finally, the last part of this section covers known issues during the operation of the CMS detector that have a potential impact on the statistical inference presented in this thesis.

3.3.1 Track and vertex reconstruction

As mentioned in the discussion of the tracking system, the signals in this part of the detector are used to reconstruct the trajectories of charged particles. These tracks are calculated using an iterative fit procedure based on the Kalman filter, which is further discussed in Refs. [107, 108]. This requires excellent knowledge of the exact position of the individual modules in the tracking system, which is achieved with the alignment procedure further described in Refs. [109, 110].

The tracks are used to reconstruct the interaction points of particles. Collision points of partons emerging from the initial protons are referred to as primary vertices, while decay points of long-lived particles such as bottom quarks are labeled as secondary vertices. Under consideration of the cross section for inelastic proton-proton scattering processes of $\sigma_{\text{inel}} = 69.2\,\text{mb}$ [111], the luminosity provided by the LHC leads to approximately 30 interactions per bunch crossing. The vertex with the largest squared sum of transverse momenta $\sum p_{\text{T}}^2$ is identified as the primary vertex of interest, while the other interactions are referred to as pileup. In order to reconstruct the physics processes correctly, it is crucial to assign the observed tracks to the correct vertex, which is especially challenging considering the high track multiplicity associated with the collision of hadrons. The performance of the track and vertex reconstruction is reviewed in Ref. [88].

3.3.2 The particle flow algorithm

The particle-flow (PF) algorithm [112, 113] is designed to optimally reconstruct electrons, muons, photons and hadrons that emerge from the proton collisions and reduce the impact of the pileup events. To this end, it utilizes information from all subdetector systems by connecting the information using a dedicated linking algorithm. The data analysis presented in this thesis is based on the identification of the particles by the PF algorithm. The identification processes for the different physics objects are briefly discussed in the following. A thorough review of the PF algorithm and its performance can be found in Refs. [113, 114].

Muon identification

Muons are reconstructed using the information of the inner tracker and the muon system. During the reconstruction, the PF algorithm considers three types of muons, which are distinguished by the requirements for the reconstruction. Standalone muons are reconstructed based entirely on the muon system. Global muons are constructed by matching a standalone-muon track to a track in the inner tracker if the parameters of the two tracks propagated onto a common surface are compatible. Tracker muons are constructed by extrapolating the tracks obtained in the inner tracker to the muon system. If at least one muon segment matches this extrapolation, the track is associated with a tracker muon. Using these definitions, the PF algorithm applies several selection criteria, such as the number of hits in the different subsystems and the particle activity in the vicinity of the muon candidates which is referred to as the isolation of the particles. As a result, the algorithm is able to account for the different production mechanisms of muons in the collision events, such as the direct production of muons in the initial parton collision or as a result of electromagnetic or hadronic activity as the particles propagate. Tracks identified as PF muons are removed from further consideration in the reconstruction of the other particles. The reconstruction criteria are combined to several working points that are provided centrally within the CMS collaboration and are further discussed in Refs. [115]. The muons used in this thesis are selected using the tight working point, which is designed to suppress muons originating from particles decaying in flight and from hadronic punch-through.

Electron and photon reconstruction

As discussed in Sec. 3.2.2, electrons and photons are expected to deposit most of their energy in the ECAL. Additionally, the electrons interact with the tracking layers, which allows the reconstruction of their trajectory. Consequently, the information provided by the tracking system and the ECAL is used by the PF algorithm to reconstruct electrons and photons. The signals in the ECAL crystals are reconstructed by fitting the signal pulse with multiple template functions to subtract the contributions of pileup events occurring in adjacent bunch crossing before or after the bunch crossing of interest. Similar to the detection principle used in the ECAL, an electron may produce bremsstrahlung photons and photons may convert to an electron-positron pair by interacting with the detection material in the tracker system. Consequently, these particles may no longer be a single particle by the time they arrive at the ECAL, thus impacting the trajectory reconstruction and energy measurement of the initial electrons and photons. These effects are accounted for by the application of dedicated algorithms that reconstruct these processes in superclusters which are enlarged in the ϕ-η plane with respect to the ECAL crystals. Similar to muon identification, the criteria of the electron and photon reconstruction are combined in different working points, which include a set of additional requirements on the signals in the tracker system, the shape of the energy deposition in the ECAL and the isolation of the particles. The reconstruction of electrons and photons is further described in Ref. [116]. In the scope of this thesis, the tight working point is used, which has an identification efficiency of approximately 90 %. The elements used for the reconstruction of electrons and photons are removed from further consideration of the PF algorithm.

Hadron and jet reconstruction

The remaining elements used for the PF algorithm are attributed to particles from hadronization processes, specifically charged and neutral hadrons and non-isolated photons that can arise from π^0 decays. First, ECAL and HCAL entries in the barrel region of the detector that are not linked to any tracks are attributed to neutral particles, where ECAL

entries are identified as non-isolated photons and HCAL signals as neutral hadrons. In the region $|\eta| \geq 2.5$, charged and neutral hadrons can no longer be clearly separated due to the missing tracker information. Since the charged hadrons also leave energy deposits in the ECAL, the exact origin of ECAL clusters is no longer clear, thus rendering the assumption that these clusters arise primarily from photons no longer justified. Therefore, linked ECAL and HCAL entries are attributed to the same hadron, while standalone ECAL clusters are identified as photons. The remaining calorimeter entries linked to trajectories in the tracking system are attributed to charged hadrons. More information about the reconstruction of these particles can be found in Ref. [113].

The objects arising from the hadronization of color-charged particles emerge as jets of particles, all moving in a similar direction corresponding to the conservation of the momentum of the initial parton. In the scope of this thesis, these jets are reconstructed using the anti-k_T algorithm presented in Ref. [117], which has been shown to be robust against soft radiation and collinear splitting effects. The algorithm performs a sequential combination of neighboring particles with high momenta in clusters using the distance metrics

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.6}$$

$$d_{iB} = p_{\mathrm{T},i}^{-2}. \tag{3.7}$$

In these equations, $p_{\mathrm{T},i,j}$ denotes the transverse momentum of particle i and j, ΔR_{ij}^2 denotes the angular distance as defined in Eq. (3.5) between the particles and d_{iB} is the distance of object i to the beam line. The variable R is the size parameter attributed to the jet, which is set to $R = 0.4$ for the jets used in this thesis. These jets are commonly referred to as AK4 jets. The algorithm computes the distances as specified above for all particles. If the minimum value in this list is the distance d_{ij}, the corresponding objects are combined into a jet-like structure. If the minimum distance is d_{iB}, the object i is declared a jet. The performance of this jet algorithm is reviewed in Ref. [118].

Missing transverse energy reconstruction

As mentioned in Sec. 3.2.1, the transverse momentum of the colliding protons and associated partons prior to the collision event is negligible compared to the momentum along the beam direction, which is also true for the observed final state particles due to momentum conservation. However, there are very weakly interacting particles, such as neutrinos, that do not interact with the detector and escape without detection. As a consequence, an imbalance in the observed transverse momentum in the detector arises, which is referred to as the missing transverse energy or momentum and is defined as

$$\vec{\not{E}}_\mathrm{T} = -\sum_{\text{visible}} \vec{p}_\mathrm{T} = \sum_{\text{invisible}} \vec{p}_\mathrm{T}. \tag{3.8}$$

This observable can be used to reconstruct neutrinos indirectly, and is also commonly used in searches for new particles such as dark matter as presented for example in Ref. [63].

Pileup mitigation

The PF algorithm allows for powerful mitigation of the effects of pileup events. Charged hadrons with reconstructed tracks associated to pileup vertices are removed from the aforementioned steps for the reconstruction of the physics objects. This technique is referred to as pileup-charged-hadron subtraction (CHS) and is applied for the objects used in the scope of this thesis.

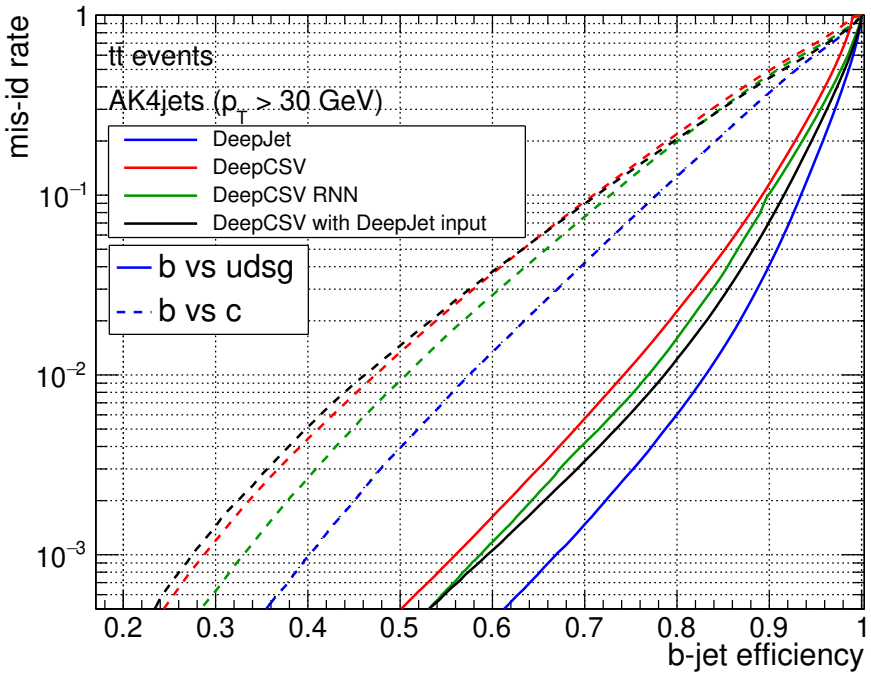


Figure 3.3: **Efficiency of the DeepJet tagger**. Shown are the tagging efficiencies for the discrimination of bottom quarks against light-flavor partons (solid line) and charm quarks (dashed line), respectively, of the DeepJet (blue) and other tagging algorithms. Figure taken from Ref. [120].

The CHS technique cannot be applied to photons and neutral hadrons as well as reconstructed particles outside the tracker acceptance. In these cases, the average p_T contribution of pileup events is estimated based on the assumption that the corresponding p_T density is uniform per unit area in the (η,ϕ)-plane. The CHS method as well as different estimation techniques for the p_T density from pileup interactions are further described in Ref. [113, 119].

3.3.3 B-jet identification

As will be shown in Chap. 4, this analysis relies heavily on the correct identification of jets emerging from bottom quarks. The flavor of the original parton producing the reconstructed jets is derived with dedicated algorithms, which are referred to as tagging algorithms. The current algorithm used within CMS is the DeepJet tagger [120], which is based on a deep neural network consisting of layers with different architectures. This network uses several sources of information, such as low-level properties of charged and neutral PF jet constituents as well as properties of secondary vertices arising from the longevity of bottom quarks due to their CKM-suppressed transitions into other quark generations. The goal of the DeepJet tagger is to classify jets as one of six flavor classes. In the scope of this thesis, the sum of the b-jet-related classes is used as an observable to identify jets originating from bottom quarks.

Figure 3.3 depicts the tagging efficiency for events related to $t\bar{t}$ pair production, which is one of the relevant topologies for this data analysis. The b-tagging performance is quantified in the form of working points, which are defined using the misidentification rate of light-flavor partons and are summarized in Tab. 3.1. In the scope of this thesis, the

Table 3.1: **B-tagging working points and the corresponding discriminant values**. Shown are the working points, the associated minimal values of the DeepJet discriminant and the corresponding misidentification rates of light-flavor jets. Values taken from Refs. [121–123].

Working Point	Min. b-tag discriminant value			Misidentification rate in %
	2016	2017	2018	
loose	0.0614	0.0521	0.0494	10.0
medium	0.3093	0.3033	0.2770	1.0
tight	0.7221	0.7489	0.7264	0.1

medium working point is used to identify b-jets, which has a misidentification rate of 1 % and a b-tagging efficiency of approximately 80 %.

3.3.4 Known detector issues that impact the event selection and reconstruction

During the recording of the data with the CMS detector, two issues arose that impact the analysis presented in this thesis. These problems are briefly summarized in this section.

In the course of the detector operation in 2016 and 2017, a gradual timing shift of the ECAL was not correctly propagated to the L1 trigger inputs, which resulted in a significant fraction of the high-$|\eta|$ candidate objects to be associated to the previous bunch crossing by mistake. As a result, events with a significant amount of ECAL energy in the region of $2 < |\eta| < 3$ are able to veto themselves since the L1 rules forbid two consecutive bunch crossings to fire. This effect is referred to as "L1 prefiring" [124]. The corresponding simulated samples are corrected and a dedicated uncertainty in the statistical inference is applied, which is further explained in Secs. 5.2.6 and 8.1, respectively.

Furthermore, a region of the detector was completely depleted of HCAL energy deposits during the 2018 operation due to a broken HCAL submodule, which is referred to as the HEM issue [125, 126]. As a result, the number of reconstructed isolated electrons is significantly higher since information provided by the HCAL is used to distinguish between real electrons that leave energy deposits in the ECAL and jets that would additionally deposit energy in the HCAL. The impact on the observables relevant for the analysis presented in this thesis is small. Nevertheless, this effect is considered by including a dedicated uncertainty in the statistical inference, which is discussed in Sec. 8.1.

3.4 Future operation at the LHC

In the search for physics beyond the SM, the LHC is a powerful tool that nonetheless has its limits. The maximum collision energy cannot be increased beyond 14 TeV due to the limits set by the circumference of the tunnel and the strength of the dipole magnets. Since physics at higher energy scales is therefore not directly accessible, the only way to find new physics at the LHC is to increase the statistical sensitivity of the analyses by either developing novel techniques, gaining better control of systematic uncertainties or by analyzing a larger data set. For a fixed collision energy and thus fixed cross section, the latter necessitates increasing the luminosity recorded with the detectors. The current estimates for the LHC operation and the size of the collected data sets are shown in Fig. 3.4. Based on these expectations, it would take approximately ten years of data taking with the current version of the LHC to collect enough data to double the statistical precision of future analyses.

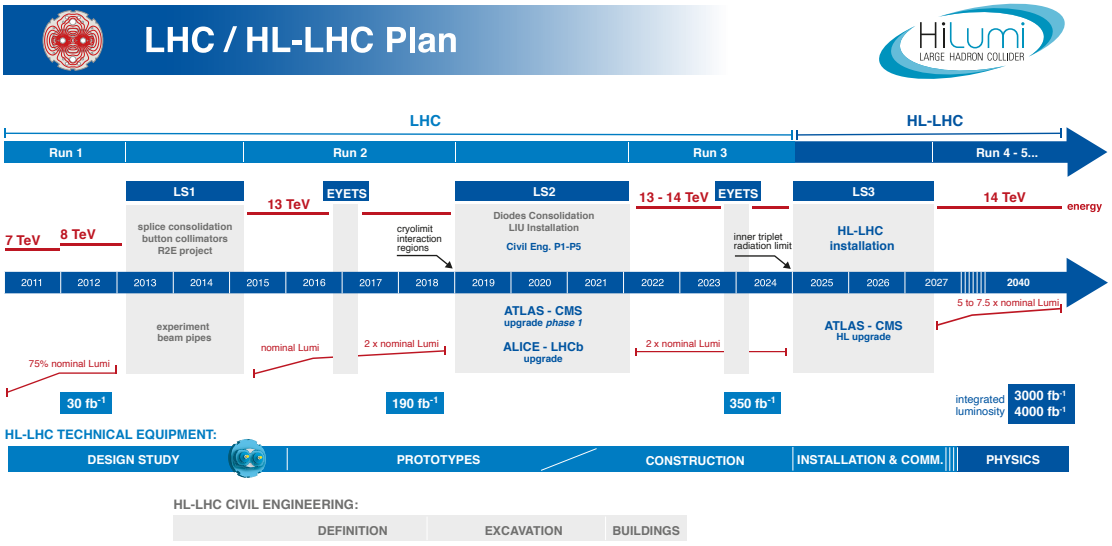


Figure 3.4: **Future plans for the operation of the LHC ring**. Shown are the time estimates for the operational phases of the LHC. At the time of writing this thesis, Run 3 is about to begin. Figure taken from Ref. [32].

In order to exploit the full potential of the LHC at the current energy scale, the High Luminosity Large Hadron Collider (HL-LHC) project aims to upgrade the ring such that the luminosity increases by an order of magnitude [31, 32, 37]. Considering Eq. (3.1), the luminosity can be increased by improving the control over the beams such that $\mathcal{F}$ is close to one and reducing the beam sizes. Improving this highly complex system is challenging and necessitated years of research, which are summarized in Ref. [31]. The necessary improvements are made possible by the utilization of novel technologies, such as the usage of Nb_3Sn superconducting magnets, superconducting compact "crab" cavities and improved beam control near the designated interaction points. As shown in Fig. 3.4, this upgrade is planned to be completed by the year 2027, which will enable the detectors to record 3000–4000 fb^{-1} of integrated luminosity by the end of the HL-LHC time of operation, depending on the scenario.

While the HL-LHC offers new opportunities for future analyses, it also poses new challenges to the detectors in the form of increased levels of radiation in all detector regions and the forward regions in particular, and the mean number of pileup events, which is expected to increase from its current value of 30 to 140 or even 200 events per bunch crossing. The relevant upgrades to address these new challenges are presented in Ref. [80] and are briefly summarized in the following for convenience.

The tracking system will be replaced completely by a similar system with extended geometrical coverage of $|\eta| < 4$, improved radiation hardness, higher granularity and compatibility with higher data rates. More information will be provided to the L1 trigger system and the threshold for tracks will be reduced to approximately 3 GeV, which will allow to keep the trigger rate at a sustainable level without sacrificing physics potential. More information can be found in Refs. [80, 127].

In the barrel region, the ECAL will be operated at lower temperatures to mitigate readout noise due to radiation damage. The readout electronics will be improved to cope with

the new bandwidth requirements and will also provide precision timing information. Additionally, the information from single crystals will be provided to the L1 trigger. The scintillating tiles in the HCAL will be replaced and the readout electronics will be upgraded to silicon photomultipliers (SiPMs) with better efficiency. The ECAL and HCAL in the endcap regions will be replaced by high-granularity sampling calorimeters (HGCal) based on silicon pad sensors that are read out by SiPMs. The resulting detectors will provide high-precision transverse and longitudinal as well as timing information that will lead to improved pileup rejection and identification of electrons, photons, tau leptons and jets. More information can be found in Refs. [80, 128, 129].

The front-end electronics in the muon chambers will be replaced with improved versions offering improved radiation tolerance, readout speed and performance. Improved resistive plate chambers and new chambers based on the gas electron multiplier (GEM) technique in the forward region will add redundancy, improve triggering and the reconstruction performance and will increase the acceptance in the forward detector region up to $|\eta| \approx 2.8$. Additionally, a minimum ionizing particle timing detector will be added between the tracker and the ECAL, providing timing measurements for $|\eta| < 3.0$ at an estimated timing resolution of 30 ps, allowing for a four-dimensional reconstruction of vertices that improves the robustness against pileup events. More information can be found in Refs. [80, 130, 131].

Finally, the trigger system will be upgraded. The capabilities of the L1 stage will be greatly enhanced by the extensive usage of FPGAs as well as the additional information provided by the upgraded detector systems, which will allow for a full PF reconstruction and more sophisticated techniques for pileup mitigation. These upgrades will allow for a higher precision and increased trigger rate of about 750 kHz. Additional upgrades in the HLT will increase its trigger rate to 7.5 kHz. More information can be found in Refs. [80, 132].

As a result of these upgrades, the large number of pileup events is expected to be controlled similarly well as is currently the case and the acceptance for leptons (jets) will extend to $|\eta| < 3.0$ (4.0), which is expected to impact the future capabilities of the analysis presented in this thesis. A projection based on the analysis of the 2016 data set as discussed in Ref. [133] to the future setup is presented in Chap. 11.

4 Physics process definition

After the review of the experimental setup in Chap. 3, the processes relevant for the analysis are introduced in this chapter. As was already motivated in Chap. 2, the production of a Higgs boson in association with a top-quark-antiquark pair ($t\bar{t}H$) is of special interest to test the Standard Model (SM). Within the CMS collaboration, measurements of the $t\bar{t}H$ process are divided according to the decay modes of the Higgs boson and the $t\bar{t}$ system. The H decay relevant for this thesis ($H \to b\bar{b}$) is motivated in Sec. 4.1. The $t\bar{t}$ decay channel is briefly outlined here.

The top quarks decay into W bosons and bottom quarks almost exclusively, as is indicated by the corresponding CKM matrix element which is close to one [37]. Therefore, the decay modes of the $t\bar{t}$ system are usually identified via the decay modes of the W bosons into quarks qq' (hadronic) or a lepton and neutrino pair (leptonic). A schematic depiction of this classification is shown in Fig. 4.1. The different decay modes are analyzed in dedicated measurements within the CMS collaboration. This thesis focuses on the semileptonic (SL) decay mode, which yields a good compromise of statistical accuracy and selection efficiency compared to the other decay modes.

First, the processes and their final states are briefly introduced. The last section of this chapter describes the selection criteria for the objects in simulated samples and the observed data to identify semileptonic final states of the $t\bar{t}H(b\bar{b})$ system.

4.1 Signal processes

Since the Higgs boson is unstable, the only possibility to measure the properties of the Higgs boson is by identifying its decay products. This thesis is dedicated to the measurement of the $t\bar{t}H$ system where the Higgs boson decays into a bottom-quark-antiquark pair ($t\bar{t}H(b\bar{b})$). As shown in Sec. 2.4, this decay mode has the largest branching fraction of all Higgs boson decays, thus yielding the largest amount of data that can be used for statistical data analysis. Figure 4.2 depicts an exemplary leading-order Feynman diagram for the $t\bar{t}H(b\bar{b})$ process. The b-jets are identified using the DeepJet algorithm outlined in Sec. 3.3.3. Since this technique has a misidentification probability for the different quark flavors, all decay products of the Higgs boson are considered as signal in the scope of this analysis to correctly account for small contributions from other decay channels.

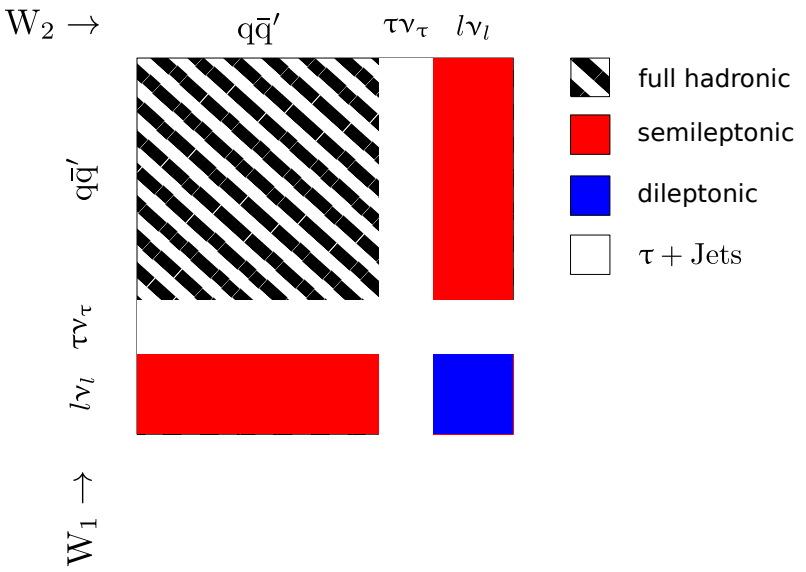


Figure 4.1: **Schematic overview of the $t\bar{t}$ decay modes**. The figure shows the different $t\bar{t}$ decay channels. Due to the challenging reconstruction of the τ leptons, this decay channel is separated from the other leptons $l = e, \mu$. The areas in the figure are not proportional to the actual branching fractions. Figure adapted from Ref. [134].

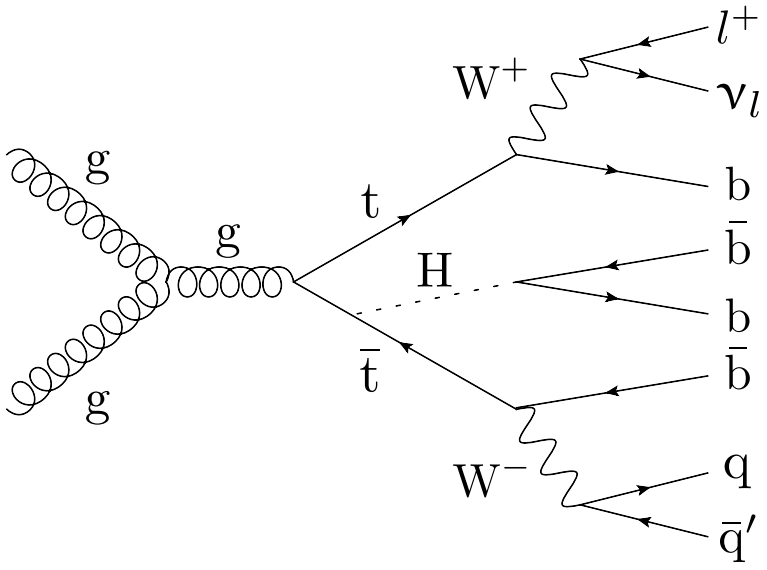


Figure 4.2: **Leading-order Feynman diagram for the $t\bar{t}H(b\bar{b})$ process in the semileptonic channel**. The leptons $l = e, \mu$ denote the leptons considered in this analysis.

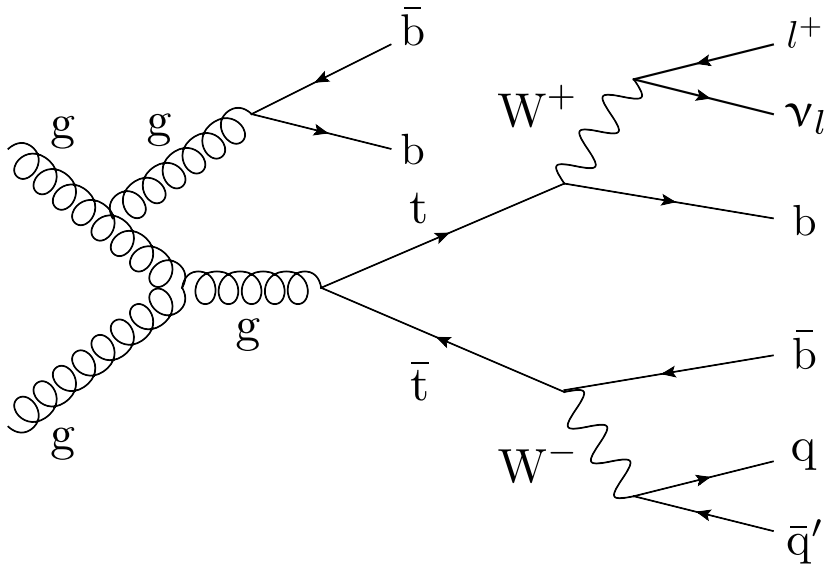


Figure 4.3: **Leading-order Feynman diagram for the $t\bar{t} + b\bar{b}$ process in the semileptonic decay channel**. The leptons $l = e, \mu$ denote the leptons considered in this analysis.

4.2 Background processes

The separation of signal and background events is crucial to correctly infer information about physics processes of interest. Ideally, this is realized by applying a selection on an observable that separates the processes perfectly. However, this is not possible since the measurement at the experiment relies exclusively on the observation of particles in the final state that are also subject to detector effects such as inefficiencies and limited acceptance. Consequently, it is challenging to separate the signal processes from background processes that produce the same final state at the matrix-element level or due to detector effects. While the contribution of the background processes can be reduced with more sophisticated selection criteria and analysis techniques, it is still vital to correctly account for them in any measurement. The following will introduce the relevant background processes for the $t\bar{t}H(b\bar{b})$ measurement.

4.2.1 Dominant background processes

The dominant background processes in any measurement yield a very similar or the same final state as the signal process, thus making a discrimination of signal and background events challenging. In case of the $t\bar{t}H(b\bar{b})$ measurement, the dominant background processes are $t\bar{t}$-pair production processes in association with additional jets, which is referred to as $t\bar{t} + \text{jets}$ in the scope of this thesis. This background is further divided into groups according to the flavor of the additional jets. Technically, this division is performed with information provided during the simulation of the events, which is further described in Ref. [135]. The resulting groups are briefly discussed in the following.

An exemplary Feynman diagram for $t\bar{t}$ production in association with two bottom quarks ($t\bar{t} + b\bar{b}$) is depicted in Fig. 4.3. The treatment of this background process is challenging in both experiment and theoretical calculation. On the experimental side, it yields the same final state as the $t\bar{t}H(b\bar{b})$ signal process and is thus hard to discriminate against the signal. Furthermore, the kinematic properties of the two processes are very similar, which makes the construction of a $t\bar{t} + b\bar{b}$-enriched and signal-depleted control region impossible. Consequently, the $t\bar{t} + b\bar{b}$ process is an irreducible background in the $t\bar{t}H(b\bar{b})$ analysis.

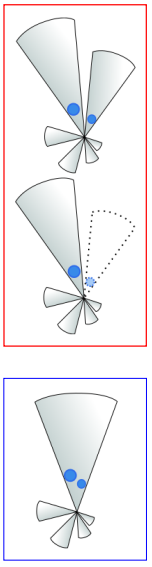


Figure 4.4: **Sub-classification of the $t\bar{t} + b\bar{b}$ process**. Shown are the sub-classes for the $t\bar{t} + b\bar{b}$ processes defined in the text. The cones indicate reconstructed jets, while the blue dots show bottom quarks in the system. The boxes indicate the distinction between the $t\bar{t}$mb (red) and the $t\bar{t} + 2b$ (blue) classes used in this thesis. Figure adopted from Ref. [136]

Additionally, the energy scales involved in the $t\bar{t} + b\bar{b}$ process range from 4 GeV (bottom quark mass) to approximately 173 GeV (top quark mass). This yields the calculation of this process particularly challenging, thus rendering the current theory predictions of this process imprecise. Overall, the control of the $t\bar{t} + b\bar{b}$ process is limited and one of the driving factors of the $t\bar{t}H(b\bar{b})$ measurement. It necessitates advanced techniques in both the statistical data analysis as well as the analysis strategy itself. This is detailed in Chap. 8.

In order to gain more control over the $t\bar{t} + b\bar{b}$ process, it is further divided into three groups based on resolution and acceptance effects in the detector, which are schematically depicted in Fig. 4.4.

$t\bar{t} + b\bar{b}$ (resolved) contains the events for which both additional b-jets are correctly resolved.

$t\bar{t} + b$ contains events in which one of the b-jets is out of the detector acceptance.

$t\bar{t} + 2b$ contains events in which the two b-jets are resolved as one jet.

In this scheme, the calculation of the $t\bar{t} + 2b$ process is particularly affected by additional collinear gluon splitting effects, which is currently challenging to calculate. Therefore, the treatment of the sub-classes in this thesis is as follows. The $t\bar{t} + b\bar{b}$ (resolved) and $t\bar{t} + b$ categories are treated as one process which is referred to as $t\bar{t}b\bar{b}$/b or $t\bar{t}$mb in the scope of the thesis. The $t\bar{t} + 2b$ class is treated separately and is the origin for an additional source of uncertainty to account for the challenging calculation of the additional gluon splitting effects. The construction of the additional source of uncertainty and the treatment in the analysis strategy are discussed in Secs. 8.1 and 8.2, respectively.

The next class of $t\bar{t}$ + jets processes is the $t\bar{t} + c\bar{c}$ process, where the additional jets originate from charm quarks. In principle, these events should be clearly separable from the $t\bar{t}H(b\bar{b})$ process by requiring b-jets in the final state that do not stem from the $t\bar{t}$ decays. However, as indicated in Sec. 3.3.3, the misidentification rate for charm quarks with the b-tagging

algorithm is non-zero. Therefore, this processes is expected to contribute to the final analysis phase space and needs to be accounted for.

The final class of the $t\bar{t}$ + jets processes contains all processes in which the additional jets arise from other quark flavors, i.e. up, down and strange quarks as well as additional gluon emissions. Similar to the $t\bar{t} + c\bar{c}$ process, this process class contributes to the analysis phase space due to the misidentification probability of the b-tagging algorithm. Although this probability is much smaller than for charm quarks, the high branching ratio for these processes yields a significant contribution to the final analysis phase space that needs to be considered in the analysis strategy. In the scope of this thesis, these processes are referred to as $t\bar{t} + \text{lf}$.

4.2.2 Minor background processes

Additional to the $t\bar{t}$ + jets background that can yield the same final state at leading order, there are other processes that can produce a similar result when considering higher-order effects, such as additional emissions and detector inefficiencies. Since the probability for the selection of such processes decreases with every higher-order effect needed to produce the relevant final state, these processes are referred to as minor backgrounds. This class of backgrounds are discussed in the following.

The tH process can produce the same final state as $t\bar{t}H(b\bar{b})$ if additional radiation is reconstructed as part of a $t\bar{t}$ system. In the scope of this thesis, the tH process is divided into the tHq and tHW processes. This separation enables additional measurements of the sign of the top-Higgs Yukawa coupling and the CP properties of the Higgs boson, which are performed in Ref. [137]. These measurements are not covered in the scope of this thesis.

Similarly, the production of a single top quark without the additional Higgs boson can lead to the relevant final state. Although these prerequisites make the selection of this process less probable, the contribution to the final analysis phase space are non-negligible since the overall production cross section is of $\mathcal{O}(400)$ times larges than the $t\bar{t}H$ cross section [138]. This process is referred to as single t in the following.

The production of other vector bosons in association with a $t\bar{t}$ pair ($t\bar{t} + V$) are also relevant for this analysis. In particular, the production of W bosons ($t\bar{t} + W$) and of a Z boson or a virtual photon ($t\bar{t} + Z$) can produce the same or a similar final state via the decays $W \to bq'$ and $Z/\gamma^* \to b\bar{b}$, respectively. Moreover, the production cross section for these processes are of the same order of magnitude as the cross section for the $t\bar{t}H$ process [139]. However, the branching ratio for the relevant Z/γ^* decays are much smaller than for the Higgs boson, and the shown W boson decay is suppressed as indicated by the CKM matrix [37]. Therefore, the contributions of these processes are severely reduced in the final analysis phase space.

Furthermore, the production of the above mentioned vector bosons in association with additional jets (V + jets) can produce a similar final state and is further divided into the production of W bosons (W + jets) and Z bosons (Z + jets). These processes can produce final states with an isolated lepton and a similar jet structure as the signal processes. However, apart from the low probability of selecting these processes in the first place, their cross sections diminish quickly with rising jet multiplicities [140]. Therefore, their contribution in the final analysis phase space is expected to be small.

The final contribution to the analysis phase space arises from the production of vector boson pairs, which is referred to as Diboson production. Specifically, these processes can yield a similar final state as the $t\bar{t}H(b\bar{b})$ process if the Z bosons decay into b pairs and the light-flavor jets originating from the decay of the W bosons are reconstructed as b-jets.

Table 4.1: **Selection criteria for the semileptonic $t\bar{t}$ decay mode**. Shown are the values for each year of data taking. If only one number is present, the value is applied for all three years.

Observable	Selection		
	2016	2017	2018
Min. value of largest electron p_T in GeV	29	30	30
Min. value of largest muon p_T in GeV	26	29	26
Max. $\|\eta\|$ of leptons		2.4	
Min. number of jets		4	
Min. number of b-jets		3	
Min. p_T of jets in GeV		30	
Max. $\|\eta\|$ of jets		2.4	
Min. $\vec{\not{E}}_T$ in GeV		20	

However, as this requires multiple higher-order effects, the contribution to the final analysis phase space is expected to be minor.

4.3 Selection criteria for the semileptonic channel

As already mentioned above, only the final state of the physics processes can be observed with the detector setup. Therefore, it is crucial to define selection criteria for the properties of the different objects that can help to identify the processes of interest, in this case the $t\bar{t}H(b\bar{b})$ signature. These criteria need to be applied in both observed data and the simulated events to ensure a consistent treatment and therefore a meaningful comparison. As stated before, this thesis focuses on the SL decay mode of the $t\bar{t}$ system. Therefore, the relevant physics objects are electrons and muons, reconstructed jets and the missing transverse energy $\vec{\not{E}}_T$ originating from the neutrinos in the system. These objects and their quality requirements as provided by the experts within the CMS collaboration are introduced in Chap. 3.

In general, only events that pass the CMS trigger algorithm are considered. For the SL channel, the single-electron and single-muon triggers are chosen. In the single-muon channel, the selection is based on the p_T of isolated muons. In the single-electron channel, a combination of triggers on the sum of the p_T of the jets in addition to a selection of the p_T of electrons is used. The technical details can be found in Tab. A.1. Moreover, the selected events are required to contain exactly the lepton that the trigger selected, i.e. an event that was chosen by the single-electron trigger is required to contain exactly one electron and no muon. Afterwards, further selection criteria are applied to objects of interest to further purify the signal-to-background ratio. The selection criteria applied to these objects to identify the SL decay channel are listed in Tab. 4.1. The higher threshold for the p_T of muons in 2017 data is chosen to account for problems with the muon trigger, which lead to a higher turn-on threshold. The selections are especially effective for the processes listed in Sec. 4.2.2 as well as events originating from pure QCD interactions. The suppression of the latter is effective enough to completely neglect the QCD background in the scope of this thesis.

5 Simulated samples

One of the core concepts of statistical data analysis in high-energy physics that is commonly used is to compare observables measured at the collision experiment with the corresponding predictions given by a theory, such as the Standard Model (SM). As described in Chap. 2, the theory-based simulated data for such a comparison is commonly generated using the Monte Carlo method which is described in more detail in Refs. [141–143]. Using this technique, it is possible to generate the predictions for these observables with random numbers which follow the probability density described by the calculation of the matrix elements (MEs), the hadronization process as well as the simulation of the detector response itself. As mentioned in Sec. 2.3, these steps of the simulation can be performed separately using the factorization theorem. For each part of the simulation chain, dedicated tools are available and used to simulated the theory prediction at varying degrees of accuracy.

This chapter briefly summarizes the details of the simulated samples used in this thesis. First, the samples used in the measurement of the $t\bar{t}H(b\bar{b})$ process in the scope of the SM are discussed. Subsequently, corrections applied to these simulated samples to improve their compatibility with observed data are briefly described. In the final part of this chapter, the simulation of the transverse momentum of the Higgs boson p_T^H in the context of the Standard Model Effective Field Theory (SMEFT) is discussed.

5.1 Standard Model samples

This section briefly summarizes the event generation of the samples used in the $t\bar{t}H(b\bar{b})$ analysis which were centrally produced by a dedicated group within the CMS collaboration. Generally, the MEs for the signal and background processes are generated at next-to-leading order (NLO) of perturbation theory with POWHEG [144–147] or MADGRAPH5_aMC@NLO [148], or at leading order (LO) using PYTHIA 8.2 [149], depending on the process. For these simulations, the value of the Higgs boson mass m_H is set to 125 GeV and the top quark mass m_t is set to 172.5 GeV. The simulation of the detector response is performed with GEANT4 [150–152].

As mentioned in Chap. 2, the probability density function for the momenta of the partons in the initial-state protons is described by the parton distribution functions (PDFs), which are provided by the NNPDF [153]. The exact choices of the PDF for the different samples are listed in Sec. 8.1. In all simulated samples, the parton showering and the hadronization processes are simulated with PYTHIA 8.2, which relies on the phenomenological Lund

Table 5.1: **Tuneable parameters used in the simulated samples**. The table shows the tuneable parameters used in the simulation of the parton shower and the hadronization using Pythia 8.2. The ME-PS matching values for the CUETP8M1 are extracted from Ref. [156]. The values for the CP5 are taken from Ref. [157].

Year	Samples	UE Tune	ME-PS matching in GeV
2016	$W + jets$, $Z + jets$, $t\bar{t} + W$, $t\bar{t} + Z$, Diboson	CUETP8M1	172.5
	all others	CP5	237.9
2017	all	CP5	237.9
2018	all	CP5	237.9

string model [154, 155]. Being phenomenological, this model entails tuneable parameters, which have to be adjusted using observed data. One of these parameters is the energy scale used for the matching of the ME and the parton-showering diagrams (ME-PS matching). The other parameters are used to describe the underlying events (UE) of the collisions that are mentioned in Chap. 2 and are referred to as the UE tune. The values used for the different simulated samples are shown in Tab. 5.1.

The overview over the different simulated samples used in this thesis is shown in Tabs. 5.2, 5.3 and 5.4. The following paragraphs briefly discuss the details of the simulation specific to the respective simulated samples.

5.1.1 Higgs-related samples

The $t\bar{t}H$ signal sample was generated in two parts according to the decay channel of the Higgs boson: one part with decays into bottom-quark-antiquark pairs and another with every other decay. Both samples are generated using Powheg at NLO accuracy and allowing for one additional jet. The inclusive $t\bar{t}H$ cross section is normalized to 507.1^{+35}_{-50} fb, which is taken from a calculation with NLO QCD and electroweak (EWK) corrections. More information about the calculation of the cross section and the branching ratios of the Higgs bosons decays can be found in Refs. [19, 158, 163].

The samples for the tHq and tHW processes are generated using MadGraph5 at LO, such that the simulated events can be reweighted to correspond to different values of the top-Higgs- and vector-boson-Higgs-couplings. The tHq sample is generated in the four-flavor scheme (4FS), which enables a better description of the additional bottom quark from the initial-state gluon splitting. In principle, the 4FS is also desired for the tHW sample. However, the 4FS prediction at LO and the five-flavor scheme (5FS) prediction at NLO show interference effects with the $t\bar{t}H$ predictions. These interference effects cannot be accounted for in the reweighting procedure. Therefore, the tHW sample is generated in the 5FS. More information on the generation of these samples can be found in Ref. [164]. The calculation of the cross section for both processes is done at NLO QCD. The tH processes have little effect in the fiducial phase space of the $t\bar{t}H(b\bar{b})$ analysis and are therefore considered as minor backgrounds.

For all Higgs-related samples, the full phase space of decay modes was considered as shown in Sec. 2.4. More information about the decay modes and the calculation of the branching fractions can be found in Refs. [19, 163]. Subsequent to the generation of these samples, the CMS collaboration published the measurement of the Higgs boson mass m_H in the diphoton final state using the full Run-II data set presented in Ref. [45]. The combination

Table 5.2: **Simulated SM samples for the $t\bar{t}H(b\bar{b})$ analysis**. The table shows the cross section and the number of generated events in the simulated samples produced for each year of data taking. The indented lines indicate dedicated samples for the corresponding production channels that were generated for the indicated sub channels.

Sample	Cross Section in pb	Number of Events		
		2016	2017	2018
$t\bar{t}H$, $H \to b\bar{b}$	0.2953 [19]	10 000 000	8 000 000	12 000 000
$t\bar{t}H$, $H \to b\bar{b}$, $t\bar{t}$ SL*	0.1298 [19]	10 000 000	9 558 675	10 000 000
$t\bar{t}H$, $H \to b\bar{b}$, $t\bar{t}$ DL*	0.0314 [19]	10 000 000	9 794 620	10 000 000
$t\bar{t}H$, $H \to$ nonb$\bar{b}$	0.2118 [19]	10 000 000	7 966 779	8 000 000
tHq	0.0743 [158–160]	10 000 000	10 000 000	30 000 000
tHW	0.0152 [158–160]	5 000 000	5 000 000	15 000 000
inclusive $t\bar{t}$ (5FS)	831.7600 [161]			
$t\bar{t}$ (SL)	365.4574 [161]	110 000 000	159 889 425	114 520 000
$t\bar{t}$ (DL)	88.3419 [161]	70 000 000	69 155 808	70 960 000
$t\bar{t}$ (FH)	377.9607 [161]	70 000 000	130 797 650	134 808 000
$t\bar{t} + b\bar{b}$ (4FS)	reweighted to $t\bar{t} + b\bar{b}$ (5FS)			
$t\bar{t} + b\bar{b}$ (SL)	–	6 494 400	5 667 979	6 123 100
$t\bar{t} + b\bar{b}$ (DL)	–	3 749 700	3 697 000	3 615 000
$t\bar{t} + b\bar{b}$ (FH)	–	6 077 198	6 727 500	5 914 500
single t				
s-channel ($W \to l\nu$)	3.3600 [138]	10 000 000	19 846 893	20 000 000
s-channel ($W \to$ had)	6.9600 [138]	5 000 000	10 000 000	10 000 000
t-channel	136.0200 [138]	32 000 000	5 982 064	160 000 000
$\bar{t}$-channel	80.9500 [138]	18 000 000	3 984 790	80 000 000
tW-channel	35.8500 [138]	5 000 000	7 945 242	10 000 000
W-channel	35.8500 [138]	5 000 000	7 977 430	7 745 276
$t\bar{t} + W$	0.5500 [139]			
$W \to l\nu$	0.1792 [139]	2 160 168	9 994 543	5 000 000
$W \to$ had	0.3708 [139]	833 298	811 306	830 000
$t\bar{t} + Z$	0.8600 [139]			
$Z \to l\bar{l}$	0.2589 [139]	1 992 438	11 141 490	14 000 000
$Z \to$ had	0.6012 [139]	749 400	750 000	750 000
Diboson				
WW	118.7000 [162]	7 982 180	7 779 630	8 000 000
WZ	65.5443 [162]	3 997 571	3 928 630	4 000 000
ZZ	15.8274 [162]	1 988 098	1 925 931	2 000 000

*: Samples only used for training of classifiers described in Secs. 8 and 10.

Table 5.3: **Simulated SM samples for other background processes**. The table shows the cross sections as well as the number of generated events in the simulated samples produced for each year of data taking. The indented lines show dedicated samples for the corresponding production channels that were generated for the indicated kinematic range. H_T denotes the some of the transverse momenta of the jets.

Sample	Cross Section in pb	Number of Events		
		2016	2017	2018
$W + \text{jets}, W \to l\nu$				
$70\,\text{GeV} \leq H_T < 100\,\text{GeV}$	1595.9900 [162]	10 127 405	20 000 000	20 000 000
$100\,\text{GeV} \leq H_T < 200\,\text{GeV}$	1627.4500 [162]	79 477 953	40 483 322	30 000 000
$200\,\text{GeV} \leq H_T < 400\,\text{GeV}$	435.2370 [162]	39 693 758	20 000 000	20 000 000
$400\,\text{GeV} \leq H_T < 600\,\text{GeV}$	59.1811 [162]	7 759 701	8 084 390	8 000 000
$600\,\text{GeV} \leq H_T < 800\,\text{GeV}$	14.5805 [162]	18 687 480	20 000 000	20 000 000
$800\,\text{GeV} \leq H_T < 1200\,\text{GeV}$	6.6562 [162]	7 830 536	17 676 257	8 000 000
$H_T \geq 1200\,\text{GeV}$	1.6081 [162]	9 510 262	35 385 974	10 000 000
$Z + \text{jets}, Z \to l\bar{l}$				
$m_{l\bar{l}} < 50\,\text{GeV}$, $70\,\text{GeV} \leq H_T < 100\,\text{GeV}$	370.7994 [162]	10 084 984	–	8 988 863
$m_{l\bar{l}} < 50\,\text{GeV}$, $100\,\text{GeV} \leq H_T < 200\,\text{GeV}$	275.0934 [162]	9 591 846	9 029 117	6 794 838
$m_{l\bar{l}} < 50\,\text{GeV}$, $200\,\text{GeV} \leq H_T < 400\,\text{GeV}$	45.6444 [162]	3 123 007	1 929 923	1 939 010
$m_{l\bar{l}} < 50\,\text{GeV}$, $400\,\text{GeV} \leq H_T < 600\,\text{GeV}$	4.3939 [162]	3 088 204	3 062 244	2 008 779
$m_{l\bar{l}} < 50\,\text{GeV}$, $H_T \geq 600\,\text{GeV}$	1.3791 [162]	3 038 457	2 931 736	1 975 490
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $70\,\text{GeV} \leq H_T < 100\,\text{GeV}$	215.0931 [162]	9 691 660	9 344 037	10 019 684
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $100\,\text{GeV} \leq H_T < 200\,\text{GeV}$	180.8598 [162]	11 017 086	11 197 488	11 530 510
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $200\,\text{GeV} \leq H_T < 400\,\text{GeV}$	40.9900 [162]	9 754 170	10 728 447	11 225 887
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $400\,\text{GeV} \leq H_T < 600\,\text{GeV}$	6.9669 [162]	10 152 204	10 219 524	9 697 098
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $600\,\text{GeV} \leq H_T < 800\,\text{GeV}$	1.6773 [162]	8 292 957	8 743 640	8 862 104
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $800\,\text{GeV} \leq H_T < 1200\,\text{GeV}$	0.7735 [162]	2 673 066	3 114 980	3 138 129
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $1200\,\text{GeV} \leq H_T < 2500\,\text{GeV}$	0.1858 [162]	596 079	625 517	536 416
$m_{l\bar{l}} \geq 50\,\text{GeV}$, $H_T \geq 2500\,\text{GeV}$	0.0044 [162]	399 492	419 308	427 051

Table 5.4: **Simulated SM samples for the estimation of the ME-PS matching and the underlying event uncertainties**. The table shows the number of generated events in the simulated samples produced for each year of data taking. The indented lines indicate dedicated samples for the corresponding decay modes of the $t\bar{t}$ system. The cross sections are the same as shown in Tab. 5.2. The "Up" and "Down" varied samples were generated using the values corresponding to the uncertainties derived in Refs. [156, 157].

Sample	Number of Events		
	2016	2017	2018
$t\bar{t}$ (SL)			
ME-PS matching Up	30 000 000	27 500 000	27 496 000
ME-PS matching Down	30 000 000	27 130 991	26 468 000
Underlying Event tune Up	30 000 000	20 122 010	27 482 000
Underlying Event tune Down	30 000 000	22 911 672	27 189 000
$t\bar{t}$ (DL)			
ME-PS matching Up	15 000 000	3 405 716	5 284 000
ME-PS matching Down	15 000 000	5 476 459	5 488 000
Underlying Event tune Up	15 000 000	5 500 000	5 494 000
Underlying Event tune Down	15 000 000	5 500 000	5 452 000
$t\bar{t}$ (FH)			
ME-PS matching Up	30 000 000	27 260 880	25 305 000
ME-PS matching Down	30 000 000	27 164 554	27 315 000
Underlying Event tune Up	30 000 000	27 301 272	24 064 000
Underlying Event tune Down	30 000 000	26 200 472	27 500 000
$t\bar{t} + b\bar{b}$ (SL)			
ME-PS matching Up	2 061 000	2 061 000	2 061 000
ME-PS matching Down	1 991 500	2 098 500	2 019 000
$t\bar{t} + b\bar{b}$ (DL)			
ME-PS matching Up	2 394 000	2 394 000	2 394 000
ME-PS matching Down	2 135 700	2 137 500	2 137 500
$t\bar{t} + b\bar{b}$ (FH)			
ME-PS matching Up	2 152 500	2 152 500	2 152 500
ME-PS matching Down	2 389 290	2 391 790	2 389 493

Table 5.5: **Scaling factors to account for updated Higgs boson mass**. The table shows the factors used to update the production cross section (first block) and branching ratios (second block) to a Higgs boson mass of $m_H = (125.38 \pm 0.14)\,\text{GeV}$. All values taken from Refs. [19, 158, 163].

	Value at $m_H = 125\,\text{GeV}$	Scale Factor
$t\bar{t}H$ XS	0.5070 pb	0.9928
tHW XS	0.0150 pb	1.0000
tHq XS	0.0770 pb	0.9987
$B_{H\to b\bar{b}}$	0.5824	0.9896
$B_{H\to W^+W^-}$	0.2137	1.0293
$B_{H\to gg}$	0.0819	0.9953
$B_{H\to l\bar{l}}$	0.0627	0.9903
$B_{H\to c\bar{c}}$	0.0289	0.9898
$B_{H\to ZZ}$	0.0262	1.0352
$B_{H\to\gamma\gamma}$	0.0023	1.0000
$B_{H\to Z\gamma^*}$	0.0015	1.0209

with the $H \to ZZ^* \to 4l$ analysis yielded a value of $m_H = (125.38 \pm 0.14)\,\text{GeV}$, which is currently the most precise measurement. In order to account for these updated results, the cross sections and branching ratios of the Higgs boson are reweighted. The factors used for this reweighting procedure are shown in Tab. 5.5.

5.1.2 $t\bar{t}$ and single top background samples

The $t\bar{t}$ background processes are simulated using POWHEG. The inclusive cross-section is normalized to $831.76\,^{+19.77}_{-29.20}\,(\text{scale})\,^{+35.06}_{-35.06}\,(\text{PDF}+\alpha_s)\,\text{pb}$, which has been calculated at next-to-next-to-leading-order (NNLO) in QCD including resummation of next-to-next-to-leading-logarithmic (NNLL) soft gluon terms with `top++2.0` [165–171].

As discussed in Chap. 4, the $t\bar{t}$ events are divided into different classes based on the flavor of the additional jets which do not stem from the top-quark decays in the event. Specifically events with additional b-jets ($t\bar{t}+b\bar{b}$ events), which constitute the irreducible background process, are generated with the dedicated POWHEG program presented in Ref. [172]. With this tool, the $t\bar{t}+b\bar{b}$ process is simulated at NLO accuracy in the 4FS and both additional b-jets are considered in the ME calculation. This method is expected to better describe the hard interaction processes in the generation of the $t\bar{t}+b\bar{b}$ process, and is hence the preferred mode of generation. This simulation has been studied by the LHC Higgs Cross Section Working Group [173, 174] and in the scope of the full Run-II $t\bar{t}H(b\bar{b})$ analysis [137] in detail. All other $t\bar{t}+\text{jets}$ events are generated using a POWHEG $t\bar{t}$ simulation at NLO accuracy in the 5FS. In these samples, only one of the additional jets is described by the ME calculation. The other jet is generated during the PS and is thus subject to less stringently defined uncertainties.

The samples are merged to fill the entire $t\bar{t}+\text{jets}$ phase space. To this end, the $t\bar{t}+b\bar{b}$ 4FS prediction is normalized to the corresponding fraction that contributes to the inclusive $t\bar{t}$ cross section in the 5FS. Consequently, the differential information for the $t\bar{t}+b\bar{b}$ process is obtained from the 4FS prediction, while the precise NNLO + NNLL inclusive $t\bar{t}$ cross section is required to stay the same. In order to ensure that no bias is introduced by this requirement, the $t\bar{t}+b\bar{b}$ normalization is estimated in situ in the final measurement.

Additionally, dedicated samples are used to estimate the effect of the underlying events and the ME-PS matching on the $t\bar{t}$ + jets simulation, which are summarized in Tab. 5.4. The variation of the underlying event tune affects the soft-particle regime and should not depend on the details of the simulation of the hard process. Therefore, no dedicated samples are produced for the $t\bar{t} + b\bar{b}$ 4FS samples. The samples are used to derive uncertainties on these effects for the $t\bar{t}$ + jets processes, which is further described in Sec. 8.1.

The single top quark background processes in the t- and tW-channels are simulated with POWHEG [175, 176], while the s-channel single top quark processes are simulated with MADGRAPH5_aMC@NLO. The inclusive cross-sections are normalized to approximate NNLO (tW channel [177]) and NLO (t- and s-channels [178, 179]). The cross-section values and uncertainties are taken from [138].

The PDF and α_s uncertainties on the $t\bar{t}$ and the single t cross sections were calculated using the PDF4LHC prescription [180] with the MSTW2008 68 % CL NNLO [181, 182], CT10 NNLO [183, 184], and NNPDF2.3 5f FFN [185] PDF sets, added in quadrature to the scale uncertainty.

5.1.3 Other background samples

As introduced in Sec. 4.2.2, further minor background contributions arise from the production of W and Z/γ^* bosons with additional jets (referred to as V + jets), $t\bar{t}$ production in association with a W or a Z boson (referred to as $t\bar{t}$ + V), and diboson (WW, WZ, and ZZ) processes. The V + jets and $t\bar{t}$ + V processes are simulated with MADGRAPH5_aMC@NLO, where the ME-PS matching is performed using the FxFx [186] prescription. Diboson production is simulated using the PYTHIA 8.2 event generator. The inclusive cross-sections are normalized to NNLO (V + jets except Z + jets 2016, WW [140]), to NLO+NNLL QCD and NLO EWK ($t\bar{t}$ + V [139]), and to NLO (other diboson [187], Z + jets 2016) accuracy. The cross-section values and uncertainties are taken from [139, 162, 188, 189].

5.2 Corrections to the event simulation

Since the final measurement heavily relies on the simulations described above, it is crucial to ensure a high quality of the description of the observed data with the simulated samples. However, this is a challenging endeavor for different reasons, such as the inability to model certain effects from first principles and the evolving understanding of effects such as detector response. Therefore, dedicated control regions are constructed. In these regions, the underlying physics is well-known and -understood, such that the comparison of the simulated events and the observed data yields a powerful tool to validate the event simulation. Differences between the simulated samples and the observed data are mitigated by scale factors (SFs), which are used to reweight the simulation where necessary.

The derivation of these SFs are subject to statistical and systematic uncertainties, which are accounted for in the final measurement. A discussion of the implementation of these uncertainties in the statistical model used for the $t\bar{t}H(b\bar{b})$ measurement can be found in Sec. 8.1. The corrections applied in the scope of this thesis are briefly discussed in the following.

5.2.1 Lepton corrections

The expected final state in the semileptonic decay channel contains muons and electrons. Therefore, SFs for the muon reconstruction, identification and isolation efficiencies are

applied. The corrections for the individual years of data taking are centrally provided by the group within the CMS collaboration dedicated to the reconstruction and measurement of muons. Further information on the SFs can be found in Refs. [190–192]. Additionally, the transverse momenta of the muons is corrected for biases due to a misalignment of the detector. Further information can be found in Ref. [193]. The SFs are provided in bins of muon p_{T} and η and are applied on an object-by-object basis. Similar to the muons, correction factors to improve the agreement of the reconstruction and identification efficiencies of electrons in the simulated samples and the observed data are applied. These corrections are applied in each event, respectively, and are provided in bins of electron p_{T} and η. The SFs are provided by the dedicated electron and photon group within the CMS collaboration and are documented in Refs. [194, 195].

5.2.2 Trigger corrections

As already mentioned in Chap. 4, the trigger decision of the CMS experiment is crucial for the selection of events considered for the $\mathrm{t\bar{t}H(b\bar{b})}$ analysis. Therefore, the trigger performance is corrected in the simulated data in order to account for differences arising e.g. from deviations in the turn-on behavior of the triggers close to the decision threshold. The trigger SFs for the muon triggers are provided centrally by the muon group within the CMS collaboration. More details on these corrections can be found in Refs. [190–192]. Similarly to the previously discussed muon SFs, the trigger corrections are provided in bins of p_{T} and η and are applied on an event-by-event basis. Unfortunately, the electron trigger SFs are not provided centrally. These corrections were derived in the scope of the $\mathrm{t\bar{t}H(b\bar{b})}$ measurement with the full Run-II data set at the CMS experiment and are provided in bins of electron p_{T} and η. The procedure is documented in Ref. [196] and the resulting corrections are approved by the experts of the CMS collaboration. Overall, the scale factors in the most relevant detector regions are found to be close to 1 and show uncertainties below approximately 5 %. This is compatible with the trigger corrections used in previous iterations of the $\mathrm{t\bar{t}H(b\bar{b})}$ analysis [21, 133].

5.2.3 Pileup corrections

During the event simulation, a fixed probability density function to model the pileup during the simulated collisions is used. This pileup distribution is later reweighted to account for the latest measurement at the CMS experiment and match the shape and central value expected for data. The reweighting follows the established procedure within the CMS collaboration and is documented in Ref. [111]. The procedure requires the inelastic cross section σ_{inel}, which is a measure of the probability of proton-proton interaction during the bunch crossings. The recommended value is the latest measurement at the CMS experiment, which is $\sigma_{\mathrm{inel}} = 69.2\,\mathrm{mb} \pm 4.6\,\%$.

5.2.4 Jet Energy Corrections

Due to the nature of the hadronization of color-charged objects, the energy measurement of jets is particularly challenging. Therefore, the detector response is constantly studied and calibrated. Previous measurements and simulations are updated to exploit the full power of the CMS detector. These Jet Energy Corrections (JEC) are divided into two groups, which are briefly discussed in the following.

Jet Energy Scale

Corrections to the Jet Energy Scale (JES) are provided centrally with the CMS collaboration and are derived by a dedicated group. The following explanation is based on Ref. [197].

A more thorough review of these corrections at the CMS experiment can be found in Refs. [198–200]. The corrections for the JES are derived in the following steps.

First, the energy of pileup contributions during the jet clustering are subtracted, which is achieved by comparing simulated QCD dijet samples that were generated with and without pileup overlay. This procedure yields correction factors as a function of offset energy density ρ, jet area A, jet η and jet p_{T}. Additionally, residual differences between data and the detector simulation as a function of η are determined with the random cone method and corrected in the observed data events.

Subsequently, the simulated jet response corrections are calculated. They are derived from a QCD dijet sample by comparing the reconstructed jet p_{T} to the corresponding transverse momentum at particle-level. These corrections are derived as a function of the jet p_{T} and η and render the detector response uniform in these variables.

The response to jets corrected by the previous steps is expected to exhibit small remaining differences of the order of percents between observed data and the simulation. Therefore, the detector response is further corrected as follows. On the one hand, the η dependence of the corrections is further refined by adjusting the response to a jet in a simulated dijet sample with respect to a jet of similar p_{T} in the barrel region of the detector. The absolute scale of the jet energy is determined for barrel jets as a function of p_{T} in simulated $\mathrm{Z} \to \mu\mu(\mathrm{ee})$+jets, photon+jets and QCD samples.

In addition, a dedicated JES source is applied for both observed data and simulated samples corresponding to 2018 to cover the effects of the HEM issue described in Sec. 3.3.4. The application follows the recommendation of the experts within the CMS collaboration and is documented in Refs. [125, 126].

Jet Energy Resolution

Dedicated measurements at the CMS experiment indicate that the Jet Energy Resolution (JER) is worse in data than in the event simulation. Therefore, it is necessary to deteriorate the resolution in the simulation by smearing the p_{T} of the jets. The following briefly describes this procedure. A more thorough review by the experts within the CMS collaboration can be found in Ref. [201].

Two methods to smear the p_{T} of the reconstructed jets are used. If the reconstructed jet can be associated with a jet clustered from generator particles, the four-momentum of the jet can be rescaled with a factor

$$c_{\mathrm{JER}} = 1 + (s_{\mathrm{JER}} - 1)\,\frac{p_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{particle}}}{p_{\mathrm{T}}},$$

where p_{T} is the transverse momentum of the reconstructed jet, $p_{\mathrm{T}}^{\mathrm{particle}}$ is the p_{T} of the corresponding jet on particle-level and s_{JER} is the data-to-simulation core resolution scale factor. If this is not possible, the reconstructed jet p_{T} is scaled with

$$c_{\mathrm{JER}} = 1 + \mathcal{N}(0, \sigma_{\mathrm{JER}})\sqrt{\max(s_{\mathrm{JER}}^2 - 1, 0)}.$$

Here, σ_{JER} and s_{JER} are the relative p_{T} resolution in simulation and data-to-simulation scale factors and $\mathcal{N}(0, \sigma)$ is a random number sampled from a normal distribution with an expected value of zero and a variance of σ^2.

5.2.5 b-tagging Corrections

As discussed in Chap. 4, the expected final state for the crucial processes in the $t\bar{t}H(b\bar{b})$ analysis consist of multiple b-jets. As explained in Chap. 3, jets emerging from bottom quarks are identified using the DeepJet tagging algorithm [120], which is applied on both observed data and simulated events. Different groups within the CMS collaboration have observed differences in the efficiency of this tagger in observed data and the simulation [202, 203]. In order to ensure a correct description of the observed data by the simulation, it is vital to mitigate these differences prior to the final measurement.

To this end, b-tagging SFs are provided by the group within the CMS collaboration dedicated to the maintenance of the b-tagging algorithm, which are parameterized as a function of the jet b-tagging score, p_{T} and η. The corrections are derived separately for udsg- (LF) and b-jets (HF) with a tag-and-probe approach, which is briefly summarized in the following. More details on this method can be found in Ref. [204].

In principle, the tag-and-probe method requires two jets. The "tag" jet is used to identify events that are to be considered in the calculation of the correction in observed data and simulated events. The information of the "probe" jet is subsequently used to actually perform the calculation. The SFs are derived using a dileptonic selection requirement with exactly two leptons and two jets. The derivation for each correction type is performed in dedicated control regions. The HF corrections are derived from a $t\bar{t}$ enriched control sample, which is constructed by rejecting events with reconstructed masses of the lepton pair close to the Z boson mass and by requiring a missing transverse momentum requirement $\vec{p}_{\mathrm{T}} > 30\,\mathrm{GeV}$. The LF SFs are derived in a Z + jets-enriched environment, which is achieved by inverting the aforementioned selection criteria for the HF region. Additionally, the "tag" jet needs to pass the medium working point for the HF region and the loose working point for the LF region, which were introduced in Sec. 3.3.3.

In each region, the distribution of the b-tag score of the "probe" jet is compared between observed data and simulated samples in bins of p_{T} and η. Here, the simulated events are normalized to the observed event yield. In order to account for a possible contamination of the control samples, the data is first cleaned before the actual calculation of the SF. For example, the correction for the HF jets as a function of the jet b-tag score, p_{T} and η is calculated with

$$\mathrm{SF}(\text{b-tag}, p_{\mathrm{T}}, \eta) = \frac{\mathrm{Data} - s_{\mathrm{LF}}}{s_{\mathrm{HF}}},$$

where s_{LF} and s_{HF} designate the number of events arising from light-flavor and heavy-flavor contributions in the simulated samples, respectively. The calculation for the LF jets is performed analogously.

After the derivation of these corrections, the SFs are interpolated across the respective p_{T} and η bins using an iterative fit method. This procedure allows for two degrees of freedom: a linear and a quadratic deformation of the interpolation function. The method provides continuous scale factors for all jet b-tag score, p_{T} and η values. The final SF for the event is the product of the corrections for the individual jets in the event.

After the application of the SFs, changes in the distribution of observables such as the jet multiplicity and the transverse momenta of the jets are observed. This behavior is not expected from the b-tagging SFs, and is therefore corrected by applying an additional correction factor as a function of the jet multiplicity and the sum of the jet p_{T} present in the event. The weights are derived from simulated events passing all selection criteria discussed in Sec. 4.3 except for the b-tagging requirement. This ensures that the phase

Table 5.6: **Settings for the event generation with `SMEFTatNLO`**. The settings were chosen to be as close as possible to the setup used in Ref. [17] and were validated by the `SMEFTatNLO` development team in a private communication [206].

Parameter	Value
Λ	1 TeV
μ_{R}	$\mathrm{m}(\mathrm{t\bar{t}H})/2 \approx 234.5\,\mathrm{GeV}$
μ_{F}	μ_{R}
μ_{EFT}	μ_{R}
PDF set	LHAPDF ID 306000

space for the derivation of the correction factors is as similar as possible to the analysis phase space. The factors are derived for the major processes $\mathrm{t\bar{t}H}$, $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$, $\mathrm{t\bar{t}} + \mathrm{c\bar{c}}$ and $\mathrm{t\bar{t}} + \mathrm{lf}$ separately, while single-top production as well as the processes introduced in Sec. 4.2.2 are corrected with the factors derived for $\mathrm{t\bar{t}} + \mathrm{lf}$. This procedure corrects the observed distortion in the relevant observables and does not impact the original correction of the b-tagging discriminant. The relevant distributions are presented in App. B.

5.2.6 L1 prefire corrections

As already explained in Sec. 3.3.4, there was a gradual timing shift of the ECAL that was not properly accounted for during the 2016 and 2017 operation of the CMS detector. The effect is not well-described in simulation and is therefore accounted for by applying SFs on an event-by-event basis. The event weights are derived using the official recommendations within the CMS collaboration described in Ref. [124]. The procedure is to calculate the probability of an event not to prefire based on the present objects and their kinematic properties. The weights are applied for all simulated samples generated for the 2016 and 2017 data taking years.

5.3 Standard Model Effective Field Theory samples

As described in Chap. 2, the transverse momentum of the Higgs boson $p_{\mathrm{T}}^{\mathrm{H}}$ in the $\mathrm{t\bar{t}H}$ process is potentially sensitive to effects beyond the SM (BSM). Simulated events based on BSM predictions in the context of the Standard Model Effective Field Theory (SMEFT) are generated for $p_{\mathrm{T}}^{\mathrm{H}}$ emerging from the $\mathrm{t\bar{t}H(b\bar{b})}$ process in the scope of a private production for this thesis. This section is structured as follows. First, the general setup for the event generation in the scope of the private production is described. Subsequently, the studies to validate the production are presented. Finally, the generation setup for the SMEFT predictions is discussed. These predictions are later used in a comparison to the Simplified Template Cross Section (STXS) measurement in the $\mathrm{t\bar{t}H(b\bar{b})}$ channel in Sec. 10.4.

5.3.1 Event generation

The generation is performed with MadGraph5_aMC@NLO v3.1.0 [205] using the `SMEFTatNLO` software package [68]. More information about the technical steps and requirements to generate $\mathrm{t\bar{t}H}$ events with this tool can be found in Ref. [65]. This setup allows to generate $\mathrm{t\bar{t}H}$ events at NLO QCD. In the following, some of the technical aspects and settings are briefly discussed.

The values used for the adjustable parameters of the tool are listed in Tab. 5.6 and reflect the recommendations by the developers [206]. During the calculation of the MEs, the

scales are fixed to these values. The PDF set used for the `SMEFTatNLO` simulation is the same as for the $t\bar{t}H$ samples discussed in Sec. 5.1 and is accessed via the LHAPDF interface [207]. The documentation for other tools involved in the event generation can be found in Refs. [208–213]. In the end, the MadGraph5_aMC@NLO tool yields the events in the so-called Les Houches (lhe) format, which is further documented in Refs. [214, 215]. This file contains the information of the particles in the initial state, the virtual processes and the final state particles before the parton shower (PS) and hadronization processes.

Next, the lhe events are propagated to the PS and the hadronization algorithms. In order to ensure the maximum amount of compatibility to the simulated samples generated by the CMS collaboration group, the same tools and configurations are used. Specifically, this means that the PS and hadronization is performed with the Pythia 8.2 tool, which is set up with the parameters used within the CMS collaboration. Additionally, the ME-PS matching is performed in this step to avoid double-counting of generation probabilities.

Finally, the events are further processed to extract the relevant information for the simulated samples, such as the phase-space weights calculated by the ME-PS matching algorithm and kinematic properties of the particles. The most important properties for the studies presented in this thesis are the transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$ and the rapidity y of the Higgs boson, which is further explained in Sec. 10.1. These features are extracted with a Rivet routine [216] that is centrally provided within the Higgs boson group in the CMS collaboration.

5.3.2 Validation of the private production

Before the generation of the BSM contributions to the $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum, the private production is validated. Ideally, the comparison should be done with a corresponding simulated sample generated by the experts within the CMS collaboration. Therefore, the SM prediction generated in the scope of this thesis using the `SMEFTatNLO` setup is compared to a corresponding MadGraph5_aMC@NLO sample produced within the CMS collaboration at NLO QCD. For the reference sample, the lhe events are generated with the centrally-provided setup within the CMS collaboration. The aforementioned steps for the PS, hadronization and the extraction of the final event properties is performed with the setup from the private production to ensure good comparability of the samples.

The comparison of both samples is shown in Fig. 5.1. For this comparison, 100 000 events in each sample have been considered. Both samples are normalized to the cross section mentioned in Sec. 5.1. The samples show good compatibility within the statistical accuracy. Therefore, the setup used for the private production of the simulated SMEFT predictions is considered to be valid.

As described in Sec. 5.1, the SM $t\bar{t}H$ sample used in the $t\bar{t}H(b\bar{b})$ analysis allows for an additional jet at the ME level. Currently, this is not yet supported in the `SMEFTatNLO` setup. Consequently, only one additional jet emerging from the NLO corrections is considered in the calculation of the ME, while all other additional jets emerge in the PS. Since it is known that the description of jets emerging from the PS is worse compared to jets considered at the ME-level, this difference in the generation setup could potentially impact the comparison of the cross sections in Sec. 10.4. In order to gauge the impact of this difference in the $p_{\mathrm{T}}^{\mathrm{H}}$ observable, a centrally-generated sample produced with MadGraph5_aMC@NLO is considered. This sample has the same settings as the simulation used in the initial validation of the private production, but allows for an additional jet in addition to the NLO QCD corrections. Therefore, it is expected to be indicative of the influence of the additional jet description in a comparison with the aforementioned samples.

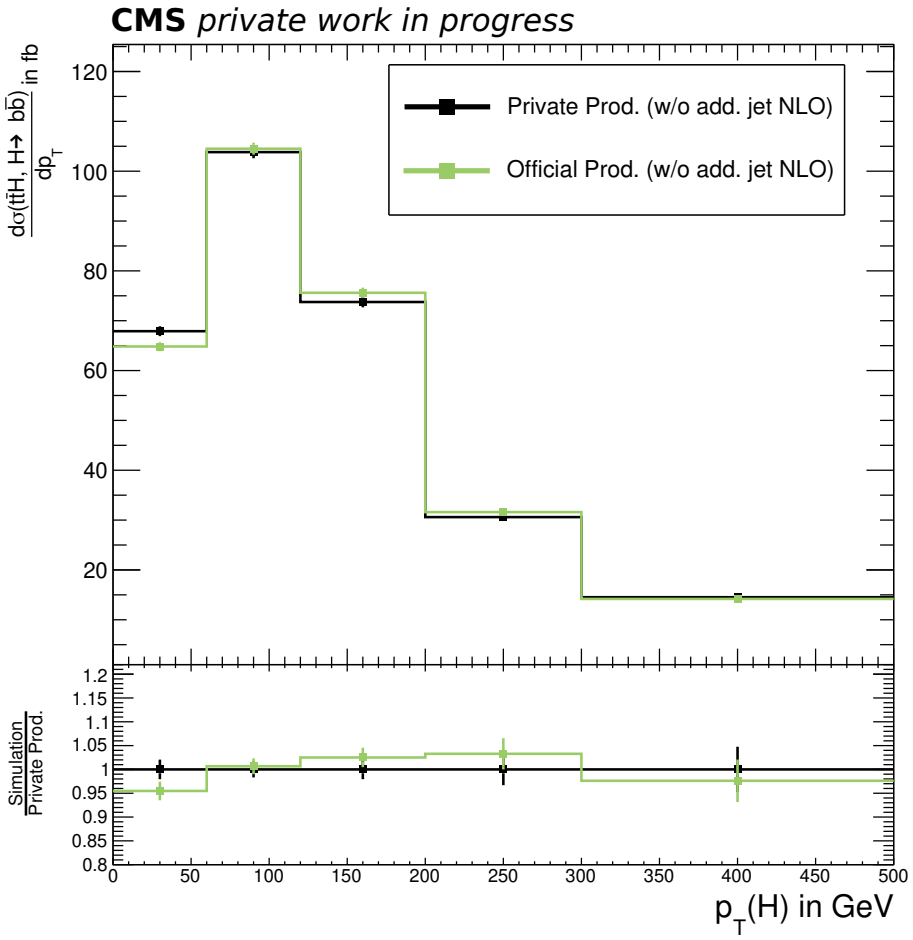


Figure 5.1: **Validation of the private production of the simulated samples**. Shown are the SM predictions for $p_{\mathrm{T}}^{\mathrm{H}}$ generated with the `SMEFTatNLO` tool (black) and the centrally-provided tools within the CMS collaboration (green). The lower part of the figure displays the ratio with respect to the simulation from the private production. Both samples are normalized to the inclusive $\mathrm{t\bar{t}H(b\bar{b})}$ cross section at NLO QCD and EWK.

The comparison of all previously discussed simulated samples is shown in Fig. 5.2. All MadGraph5_aMC@NLO samples show good compatibility within the statistical precision. Consequently, the exact description of the additional jet in the $\mathrm{t\bar{t}H}$ simulation has no significant impact on the transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$. The lower panel of the figure shows the prediction of the MadGraph5_aMC@NLO samples with respect to the Powheg simulation. Here, differences of up to 10 % are visible between the predictions from the two different generators, depending on the $p_{\mathrm{T}}^{\mathrm{H}}$ interval. Such differences between generators are well-known and covered by the scale uncertainties on the calculations, as has been shown for example in Ref. [19]. Nevertheless, these differences have to be considered when comparing these predictions to observed cross sections, as it is done in Sec. 10.4.

In summary, the private production using the `SMEFTatNLO` tool has been validated with respect to comparable samples produced by the experts within the CMS collaboration. Samples from the same generator show good compatibility within the statistical precision of the event generation. The SM predictions of the samples used in the $\mathrm{t\bar{t}H(b\bar{b})}$ analysis and the privately-produced simulations show differences of up to approximately 10 %, which have to be considered when performing the comparison to the measured cross sections

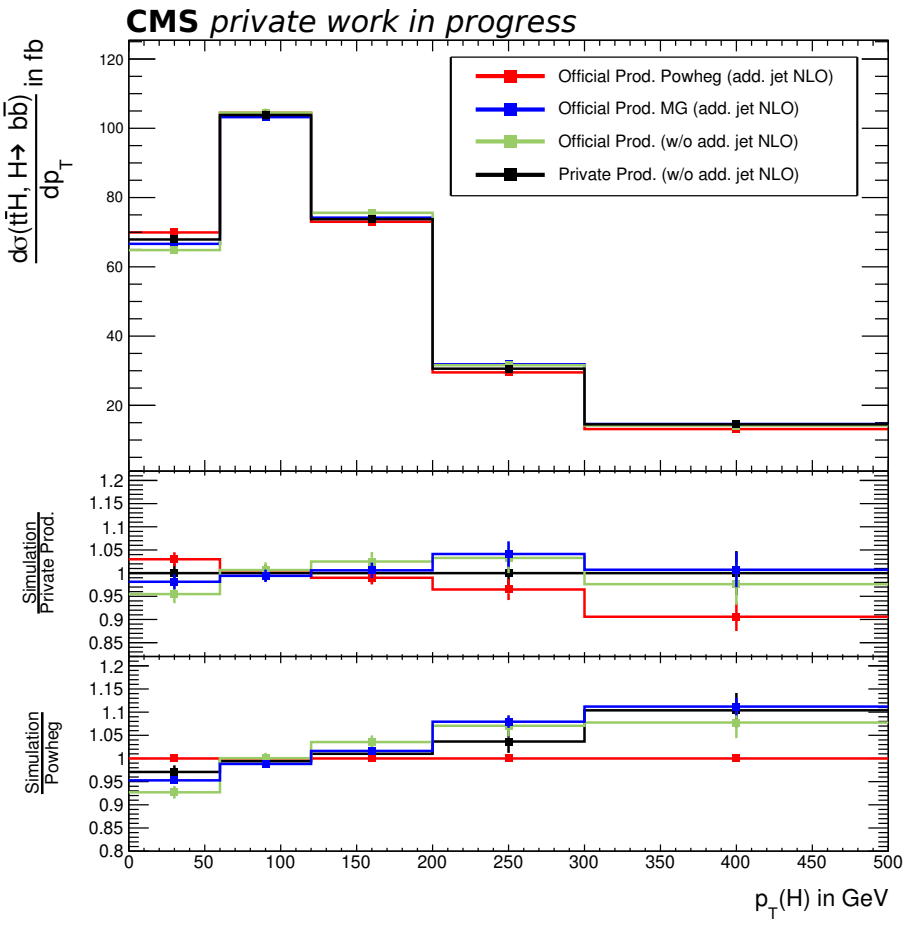


Figure 5.2: **Validation of the impact of the additional jet description in the $t\bar{t}H$ generation**. Shown are the SM predictions for $p_{\mathrm{T}}^{\mathrm{H}}$ introduced in the text. For better comparison, the ratio with respect to the private production (middle panel) and to the POWHEG prediction used in the $t\bar{t}H(b\bar{b})$ measurement (bottom panel) is shown. All samples are normalized to the same inclusive cross section calculated at NLO QCD and EWK.

later in this thesis. The next section will give details about the generation of the BSM contributions to the $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum in the scope of the SMEFT.

5.3.3 Details for the SMEFT generation

As already introduced in Chap. 2, the contribution of three SMEFT operators are considered in this thesis: the shifted top-Higgs Yukawa coupling ($c_{t\varphi}$), the effective gluon-Higgs operator ($c_{\varphi G}$) and the chromomagnetic dipole operator (c_{tG}). During the event generation with MADGRAPH5_aMC@NLO, contributions of all three operators as well as the SM predictions discussed in the previous section can be taken into account. This way, all interference effects are accounted for.

In order to fully gauge the potential contribution of the $t\bar{t}H(b\bar{b})$ channel to the sensitivity of these operators, the simulated events are based on the current constraints on the parameters, which are presented in Ref. [69]. In this publication, different data sets and observables are combined and a simplified likelihood estimation is performed to extract constraints on the SMEFT operators. To this end, the authors present the results in two ways. The first set of results extract the constraints in each operator separately, while disregarding all other

Table 5.7: **SMEFT operator values used for the simulated samples**. The tables show the constraints on the SMEFT operators in the individual (top table) and the marginalized scheme (bottom table) considered in this thesis. Values taken from Ref. [69].

Operator	Best-fit	Upper Bound	Lower Bound
$c_{\varphi G}$	0.0	3.2×10^{-3}	-3.4×10^{-3}
$c_{t\varphi}$	-9.0×10^{-2}	8.4×10^{-1}	-1.0
c_{tG}	-1.0×10^{-2}	8.6×10^{-2}	-1.0×10^{-1}

Operator	Best-fit	Upper Bound	Lower Bound
$c_{\varphi G}$	0.0	8.6×10^{-3}	-1.4×10^{-2}
$c_{t\varphi}$	1.5	5.7	-2.8
c_{tG}	3.6×10^{-1}	6.0×10^{-1}	1.2×10^{-1}

EFT contributions. This scheme is referred to as "individual" scheme in the following. The second set of results marginalizes the other parameters, i.e. integrates the contributions of the other parameters out of the likelihood. This scheme is referred to as the "marginalized" scheme. The parameter values are listed in Tab. 5.7.

The event generation is performed as follows: For the individual scheme, the individual SMEFT operators are set to the values listed in Tab. 5.7. The other two parameters are set to 1×10^{-5} and are therefore not considered. In the marginalized scheme, the bounds for the respective parameters are generated while the other two parameters are set to their best-fit values. This way, the event generation is the most comparable to the integration of the operator contributions in the likelihood. Since the SMEFT is an extension of the SM, the SM contributions are always considered in both schemes. Each parameter combination is generated separately with 100 000 events, which has been observed to yield sufficient statistical accuracy in all $p_{\mathrm{T}}^{\mathrm{H}}$ regions. Furthermore, the samples are normalized to the cross sections reported in Ref. [17] to ensure compatibility. The final cross section is calculated as described in Sec. 2.4.

6 Introduction to statistical data analysis

Due to the probabilistic nature of high-energy physics processes, the inference of the truth-level information from the experimentally observed data is challenging and requires techniques of statistical data analysis. This necessitates a powerful setup that enables the estimation of the properties of interest under consideration of both statistical fluctuations and systematic uncertainties that can arise due to the experimental setup or the theoretical calculations. This chapter briefly introduces the concepts relevant for the data analysis in this thesis. First, the test statistic used for the parameter estimation is introduced. Subsequently, relevant definitions and approximations are discussed. The explanations in this chapter are based on Refs. [217–221] and the *Statistics* review in Ref. [37], which contain more detailed reviews of the concepts introduced here.

6.1 General concepts

Generally, the truth-level information of the underlying physics processes at collider experiments cannot be accessed directly. Instead, it is only possible to measure the final state particles with the detector and compare the predictions made by a particular model, such as the Standard Model (SM) of particle physics discussed in Sec. 2.1, to the observation. The core concept for the statistical inference in many analyses in high-energy physics is therefore the estimation of the parameters used in such models and hypothesis testing.

In principle, there are two philosophies used for the interpretation of such data analyses, which are also employed in the scope of this thesis. The first is the frequentist approach, where the probability of a certain event A is interpreted as the frequency of the outcome of a repeatable experiment:

$$\mathcal{P}(A) = \frac{n(\text{observation of } A)}{n(\text{total})}$$

This approach enables the construction of statistical prescriptions to test hypotheses and of confidence intervals which are constructed such that they cover the true value of a parameter with a specified probability. Both of these concepts are discussed later in this chapter.

The other philosophy is Bayesian statistics, which is based on a more general interpretation of probability. The Bayesian probability for two events A and B can be expressed by Bayes'

theorem

$$\mathcal{P}(A|B) \cdot \mathcal{P}(B) = \mathcal{P}(B|A) \cdot \mathcal{P}(A), \tag{6.1}$$

where $\mathcal{P}(A|B)$ reads as the probability for event A given event B, and $\mathcal{P}(B|A)$ vice versa. One of the key concepts in Bayesian statistics is the degree of belief for an event, denoted as $\mathcal{P}(A)$ and $\mathcal{P}(B)$, which is also referred to as subjective probability and represents knowledge about the events prior to the measurement. As such, this approach offers a natural means to include additional information obtained by e.g. independent measurements, which will be further explained in a subsequent section.

The final concept that is needed for the data analysis presented in this thesis is the representation of the data in the form of histograms. In this approach, events are sorted into intervals $[a, b)$ which are referred to as bins. Using more bins, i.e. small intervals, increases the differential shape information available for the analysis. However, in the case of limited data samples, this decreases the number of events in each bin that can be analyzed. This problem is known as the "curse of dimensions" [222, 223] and has to be circumvented by choosing an appropriate number of bins. The discussion in this chapter is based on the assumption of a binned representation of a data set.

6.2 Parameter estimation and confidence intervals

The parameter estimation is done using the maximum likelihood method. In this approach, the content of each bin is independent from the others and the binned likelihood function

$$\mathcal{L}(\vec{\nu}) = \prod_i \mathcal{P}(n_i|\lambda_i(\vec{\nu})) \tag{6.2}$$

is maximized, where $\vec{\nu}$ denotes the parameters that are to be estimated and i enumerate the bins of the histogram used for the parameter estimation. $\mathcal{P}(n_i|\lambda_i(\vec{\nu}))$ denotes the probability to observe n data events when expecting λ. In high-energy physics, this probability follows the Poisson distribution

$$\mathcal{P}(n_i|\lambda_i(\vec{\nu})) = \frac{\lambda_i^{n_i}(\vec{\nu})}{n_i!} \exp(-\lambda_i(\vec{\nu})), \tag{6.3}$$

where

$$\lambda_i(\vec{\nu}) = \lambda_i\left(\mu, \vec{\theta}\right) = \mu \cdot s_i\left(\vec{\theta}\right) + b_i\left(\vec{\theta}\right). \tag{6.4}$$

In this equation, the number of expected events λ in bin i is the sum of the number of signal events s and the number of background events b as predicted by the model under scrutiny. The set of parameters $\vec{\nu}$ to estimate is divided into two groups. The parameter μ scales the signal contribution in the bins, and is referred to as the parameter of interest (POI). The set of nuisance parameters $\vec{\theta}$ adds degrees of freedom to the statistical model to mitigate differences between the observed data and the expected values that are unrelated to the model under scrutiny but can arise from systematic effects. The full list of nuisance parameters used in this thesis is defined in Sec. 8.1. While these parameters are not of special interest, they are needed for the parameter estimation to avoid a systematic bias.

In order to account for prior knowledge on the systematic effects, the likelihood function defined in Eq. 6.2 is extended by constraint terms that parameterize the degree of belief $\rho\left(\theta|\tilde{\theta}\right)$ where the true values $\vec{\tilde{\theta}}$ of the nuisance parameters should be. Such a frequentist-Bayesian hybrid methodology can be re-interpreted using Bayes theorem defined in Eq. (6.1)

$$\rho\left(\theta|\tilde{\theta}\right) \propto p\left(\tilde{\theta}|\theta\right) \cdot \pi(\theta), \tag{6.5}$$

where $\pi(\theta)$ is a uniform distribution and $p\left(\tilde{\theta}|\theta\right)$ corresponds to an auxiliary measurement of the nuisance parameters. The distributions used for $p\left(\tilde{\theta}|\theta\right)$ in the scope of this thesis are introduced in Sec. 6.4. Using this definition, the likelihood function can be written in a fully-consistent frequentist approach as

$$\mathcal{L}\left(\mu, \vec{\theta}\right) = \prod_i \mathcal{P}\left(n_i|\mu \cdot s_i\left(\vec{\theta}\right) + b_i\left(\vec{\theta}\right)\right) \cdot p\left(\tilde{\theta}|\theta\right). \tag{6.6}$$

Since more degrees of freedom reduce the precision of the parameter estimation, the profile likelihood method is used throughout this thesis. In this approach, the nuisance parameters are calculated as a function of μ, thus reducing the set of estimated parameters. Following the Neyman-Pearson lemma [224], the final test statistic $\tilde{q}_\mu$ used for the parameter estimation is defined as the likelihood ratio

$$\tilde{q}_\mu = -2\ln\left(\frac{\mathcal{L}\left(\mu, \hat{\vec{\theta}}_\mu\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\vec{\theta}}_{\hat{\mu}}\right)}\right), \tag{6.7}$$

where the quantities with a hat denote the parameter values that maximize $\mathcal{L}$, thus minimizing the test statistic $\tilde{q}_\mu$. The covariance matrix V and with it the correlations of the parameters $\vec{\nu}$ can be calculated using the second derivative at the maximum of the likelihood:

$$V_{ij}^{-1} = \left.\frac{\partial \mathcal{L}}{\partial \nu_i \partial \nu_j}\right|_{\vec{\hat{\nu}}}. \tag{6.8}$$

Using Wilks' theorem [225], it can be shown that $\tilde{q}_\mu$ can be used to construct confidence intervals corresponding to the Gaussian standard deviations in the limit of large samples. In this case, the test statistic is (hyper-)parabolic and the confidence interval corresponding to s Gaussian standard deviations for one POI can be calculated as the interval fulfilling

$$\tilde{q}_\mu(s) = \tilde{q}_\mu^{\min} + s^2. \tag{6.9}$$

An example of the test statistic and the construction of the confidence intervals is visualized in Fig. 6.1. The formalism enables the construction of different test suits to validate the parameter estimation. An overview of these tests can be found in Ref. [134].

6.3 Hypothesis tests and the calculation of significances

The fully frequentist formulation of the test statistic $\tilde{q}_\mu$ offers a powerful test suit to reject hypotheses H. In the scope of this thesis, two hypotheses are of special interest. The first is H_1, which corresponds to the SM prediction for the signal and background processes introduced in Chap. 4. The base hypothesis H_0 in the scope of this thesis is the SM prediction for the backgrounds. This ensures an unbiased approach for the test of H_1.

The excess of H_1 over H_0, i.e. the presence of the signal, can be quantified by the p-value which denotes the probability of observing an excess of events as large or larger than the observed one given H_0. The formal calculation of this p-value is as follows. First, the nuisance parameters $\vec{\theta}$ are estimated using the test statistic $\tilde{q}_0$ in the absence of the signal, i.e. for the value $\mu = 0$. Subsequently, pseudo data is generated around the resulting best estimates $\hat{\vec{\theta}}_0^{\text{obs}}$ considering the posterior constraints on the nuisances and fluctuations associated with Poisson-distributed event counts. The resulting distribution of the test

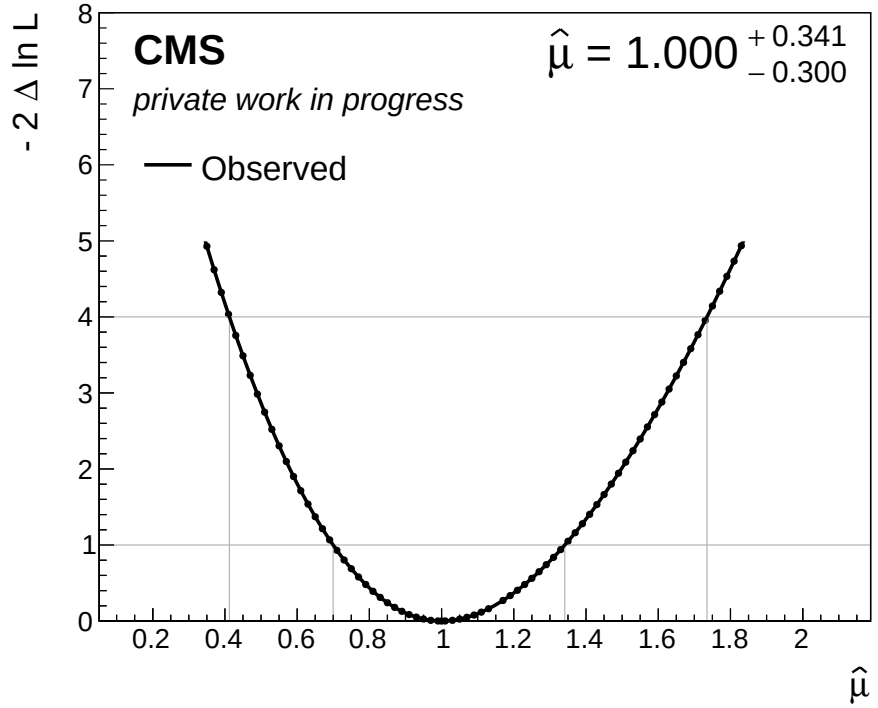


Figure 6.1: **Exemplary visualization of the test statistic for one parameter**. Shown is the test statistic as defined in Eq. (6.7) as a function of the parameter μ. The best estimate for the parameter is located at the minimum of the test statistic. The confidence intervals corresponding to one (two) Gaussian standard deviations are constructed by crossing the test statistic at one (four), which is visualized by the horizontal and vertical lines.

statistic evaluated on this pseudo data is the probability density function $f\left(\tilde{q}_0|0,\vec{\hat{\theta}}_0^{\text{obs}}\right)$ and reflects the expected fluctuation of the background-only hypothesis H_0. Finally, the p-value is defined as

$$p = \mathcal{P}\left(\tilde{q}_0 \geq \tilde{q}_0^{\text{obs}}\right) = \int_{\tilde{q}_0^{\text{obs}}}^{\infty} f\left(\tilde{q}_0|0,\vec{\hat{\theta}}_0^{\text{obs}}\right) \mathrm{d}\tilde{q}_0, \tag{6.10}$$

where $\tilde{q}_0^{\text{obs}}$ corresponds to the minimum value of the test statistic without the signal contribution observed in the given data.

Due to the generation of the pseudo data, this calculation is computationally expensive. Using Wilks' theorem and the resulting asymptotic behavior of the likelihood ratio test statistic $\tilde{q}_0$, the p-value can be estimated as [221]

$$p^{\text{estimate}} = \frac{1}{2}\left[1 - \text{erf}\left(\sqrt{\frac{\tilde{q}_0^{\text{obs}}}{2}}\right)\right], \tag{6.11}$$

where "erf" denotes the Gaussian error function. This method of calculation is used within the scope of this thesis.

The p-value can be converted into a significance s in units of Gaussian standard deviations σ using the convention of a "one-sided Gaussian tail":

$$p = \int_s^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \mathrm{d}x = \frac{1}{2} P_{\chi_1^2}\left(s^2\right), \tag{6.12}$$

where $P_{\chi^2_1}$ denotes the survival function of χ^2 with one degree of freedom. There are two thresholds of importance for the significance in the context of high-energy physics. If the excess of H_1 over H_0 exceeds $3\,\sigma$, there is evidence for H_1. If the significance exceeds $5\,\sigma$, the background-only hypothesis H_0 is rejected in favor for the hypothesis including the signal. If the significance is below the latter threshold, the data is inconclusive and no hypothesis can be rejected. In this thesis, two types of significance are used. The expected significance is evaluated with pseudo data corresponding to the H_1 hypothesis, while the observed significance is evaluated on the data observed at the CMS experiment.

6.4 Treatment of nuisance parameters

Finally, this section briefly introduces the probability density functions used to constrain the nuisance parameters based on external information such as auxiliary measurements that are the basis for the parameterization further discussed in Sec. 8.1. Generally, these parameters are divided into two groups according to their effect on the distributions of the final observable used for the parameter estimation. The first group changes the overall yield of the histograms. The prior constraint for these parameters follows the log-normal distribution

$$p\left(\tilde{\theta}|\theta\right) = \frac{1}{\sqrt{2\pi}\ln\left(\kappa\right)} \exp\left(-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln\left(\kappa\right))^2}\right) \frac{1}{\theta}. \tag{6.13}$$

The standard deviation of this function is characterized by the width parameter κ, which corresponds to a prior uncertainty ϵ with $\kappa = 1 + \epsilon$. For example, a prior uncertainty of $10\,\%$ is characterized by a log-normal distribution with width $\kappa = 1.10$. Technically, the variation for an observable A with best estimate $\hat{A}$ and standard deviation κ_A is modified as $A = \hat{A} \cdot \kappa_A^{\theta}$, where θ follows standard Gaussian distribution

$$p\left(\tilde{\theta}|\theta\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(\tilde{\theta} - \theta\right)^2}{2\sigma^2}\right), \tag{6.14}$$

with central value $\tilde{\theta} = 0$ and standard deviation $\sigma = 1$. It can be shown that A then follows the log-normal distribution defined in Eq. (6.13).

The second group of parameters modifies both the shape and the yield of the distributions used for the parameter estimation. While the yield modification is treated with the log-normal distribution, the shape effect is estimated based on dedicated varied histograms corresponding to the $\pm 1\,\sigma$ effect of the respective uncertainty source. The parameter values θ for these uncertainties follow the Gaussian distribution introduced in Eq. (6.14). The shape effect is interpolated with the spline function

$$\alpha(\theta) = \frac{1}{2}\left((\delta^+ - \delta^-)\theta + \frac{1}{8}(\delta^+ + \delta^-)(3 \cdot \theta^6 - 10 \cdot \theta^4 + 15 \cdot \theta^2)\right), \tag{6.15}$$

where $\delta^{\pm}$ corresponds to the difference between the nominal distribution and the $\pm 1\,\sigma$ variations in each bin, respectively, and $|\theta| < 1$. For $|\theta| > 1$, a linear extrapolation is used. These variations correspond to a continuous vertical morphing of the histograms, which is further described in Refs. [221, 226].

7 Neural Networks

As research in high-energy physics progresses, it is crucial to expand the area of research in order to construct new measurements of physics quantities and find indications for physics beyond the Standard Model. Such tests often require the measurements of processes with a small production probability. Due to the fact that the background processes are often much more dominant in such measurements compared to the physics process of interest, powerful techniques to achieve a good separation of signal and background events are crucial. In high-energy physics, multivariate analysis (MVA) methods have proven to be effective tools for this task. In recent years, neural networks have become the state-of-the-art tool in many aspects of the field.

As will be shown in Sec. 8.2, the $t\bar{t}H(b\bar{b})$ analysis employs neural networks for the separation of signal and background processes. Therefore, this tool is briefly reviewed in this chapter. In the course of this thesis, possible applications of different network architectures were studied in dedicated master theses, specifically the application of Bayesian Neural Networks [227], Adversarial Neural Networks [228, 229] as well as the direct consideration of systematic uncertainties in the neural network training [227, 229]. While these architectures showed great potential for the process discrimination, they did not improve the performance of the $t\bar{t}H(b\bar{b})$ analysis with respect to the previously-used strategy significantly in their current form. Therefore, the established architecture based on feed-forward neural networks are used for the $t\bar{t}H(b\bar{b})$ analysis, which are trained for multi-classification using a supervised-learning approach. The basic principles for such networks are briefly introduced in this chapter. A more complete overview of different architectures and techniques can be found in Ref. [230–232].

7.1 Feed-Forward neural networks

An Artificial Neural Network (ANN) is based on the perceptron as defined in Ref. [233]. The perceptron is modeled after the neural nodes in the human brain that receive inputs x_i and yield an output o if a certain threshold is met. This output value can be used to make a decision based on the given set of input features, such that $o > 0$ could relate to the decision A and $o \leq 0$ to decision B. Using a non-linear function f for the output value

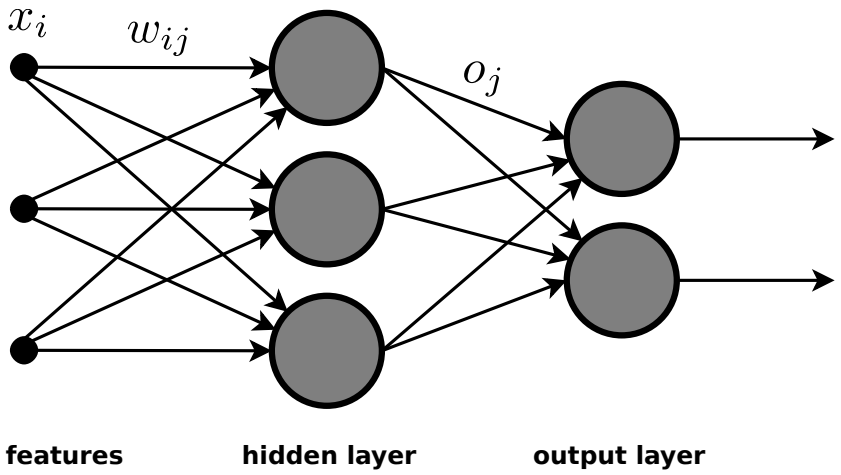


Figure 7.1: **Schematic depiction of the ANN structure**. The input features x_i are multiplied with the corresponding weight w_{ij} and propagated to the neuron j. The output o_j is further propagated in the ANN. Figure adapted from Ref. [234].

enables the construction of decision boundaries for many problems. The output value o can then be written as

$$o = f\left(\sum_i x_i \cdot w_i\right), \tag{7.1}$$

where w_i denotes a trainable weight for the corresponding input feature, which are further discussed in Sec. 7.2.

The ANN uses the perceptron, or "neurons" in the following, as building blocks for structures with more complexity. A schematic example of this structure is shown in Fig. 7.1. The neurons are arranged in layers that propagate the outputs o_j of the individual neurons to the next layer in the ANN structure. Consequently, Eq. (7.1) is extended for node j in layer l to

$$o_j^l = f\left(\sum_m o_m^{l-1} w_{jm}^l\right) = f\left(\sum_m f\left(\sum_i x_i w_{im}\right)^{l-1} w_{jm}^l\right). \tag{7.2}$$

The final layer is commonly referred to as "output layer", while the intermediate layers are the "hidden layers". The number of layers and their neurons are hyper parameters and are a design choice depending on the task of the ANN. In this thesis, multi-classification ANNs are used, which have one node for each considered class k in the output layer. The values in the output layer for the output node k are referred to as y_k in the following. The ANN yields a value for each of these classes for each point in the data set. The values y_k are normalized using the softmax activation function [235]

$$y_k = \frac{e^{z_k}}{\sum_j e^{z_j}}, \tag{7.3}$$

where z_k denote the ANN output values prior to the normalization. Consequently, the sum of the output values y_k is one and these values can be interpreted as probabilistic quantities.

The goal of the ANN is to perform a decision that was previously learned based on the input feature spaces defined by the features x_i. The learning process, commonly referred to as "training", is briefly discussed in the following section.

7.2 ANN training

The ANN is trained by deriving the weight matrices w^l in an iterative fitting procedure using a data set that is representative for the underlying problem. The training can be performed with different strategies. In this thesis, a supervised-learning approach [236] is used, which requires the user to assign labels y_k^{truth} to each point in the training data set. For a given class k, this label is either true or false, i.e. one or zero. The difference of the final output y_k of the ANN to the predefined label is parameterized in the loss function L. In the scope of this thesis, the categorical cross entropy is chosen as the loss function, which considers the difference of all training classes simultaneously:

$$\mathrm{L} = -\sum_k y_k^{\text{truth}} \cdot \log\left(y_k\right) + \left(1 - y_k^{\text{truth}}\right) \cdot \log\left(1 - y_k\right). \tag{7.4}$$

This function is especially suited for classification tasks considering multiple classes [236–238]. The goal of the training is to adjust the weight matrices w such that the output of the ANN is as similar as possible to the true output classes y^{truth}. Consequently, the training is a minimization task of the function L.

In the scope of this thesis, the minimization is performed using the AdaGrad algorithm [239, 240], which is derived from the GradientDescent method. The weights in each layer are updated according to the partial derivative of L:

$$w_{ij}^l \leftarrow w_{ij}^l - \delta \cdot \frac{\partial \mathrm{L}}{\partial w_{ij}^l}. \tag{7.5}$$

The hyper parameter δ is commonly referred to as the "learning rate" and controls the contribution of the derivative to the update of the respective weights in each iteration of the training. The derivatives are derived using the BackPropagation method described in Ref. [241], which is widely used in machine learning applications due to its high efficiency. More information about the training of ANNs can be found in Refs. [231, 232].

A commonly used metric to show the separation power of ANNs is the Receiver Operating Characteristic (ROC). Figure 7.2 illustrates an example of this metric. The ROC is the rate of correctly identified events as a function of the fraction of events that are attributed to the incorrect class. In high-energy physics, this translates to the rate of correctly identified signal events as a function of the incorrectly classified background processes. Additionally, the area under the ROC (ROC-AUC) is used to quantify the separation power. Ideally, all events are correctly classified by the ANN, i.e. the true positive rate is one and the false positive rate is zero. In this case, the ROC-AUC is equal to one. If the ANN cannot perform the classification task, the probability to select true signal events and background events is the same. Consequently, the ROC is a linear function and the ROC-AUC is 0.5. Therefore, the expected range of the ROC-AUC is 0.5–1.

Another metric used in this thesis is the confusion matrix. An example of this metric is shown in Fig. 7.3. For a perfect discrimination between the classes, the diagonal entries of the matrix are one and all other elements are zero. Confusion matrices are useful to gauge the classification efficiency and the misidentification probability for each individual class and are widely used within the field of machine learning.

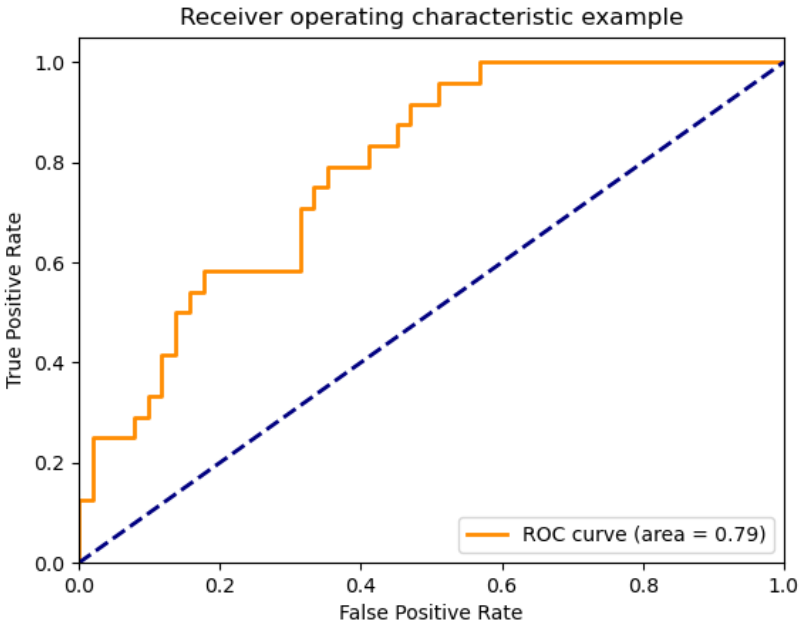


Figure 7.2: **Exemplary Receiver-Operating-Characteristic Curve**. Shown is the ROC metric (orange) for an exemplary ANN setup. For comparison, the metric is shown for an ANN that cannot separate between the classes (dashed line). Figure taken from Ref. [242].

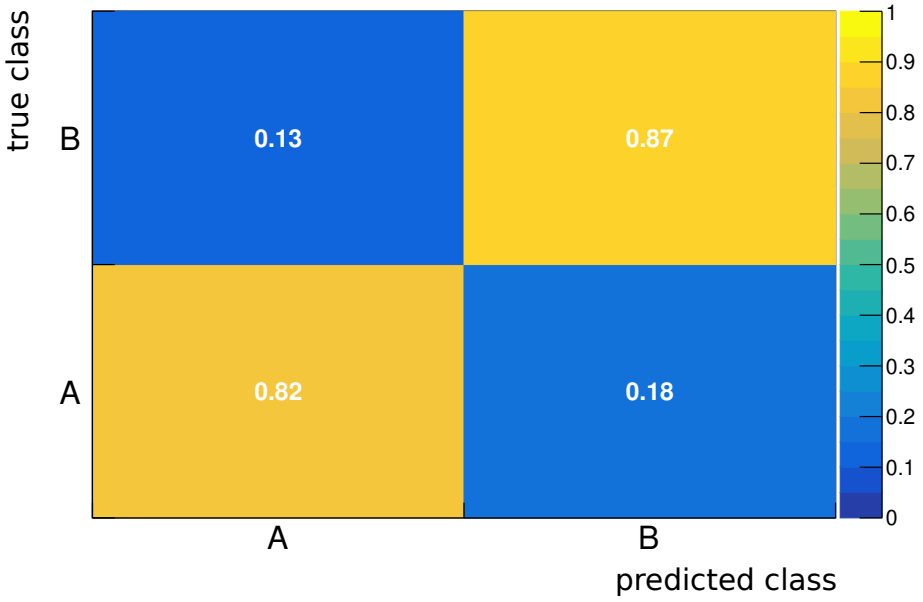


Figure 7.3: **Exemplary confusion matrix for ANN classification**. Shown is the confusion matrix for an exemplary ANN classification with two classes "A" and "B". The values are normalized to one in each row of the matrix, respectively.

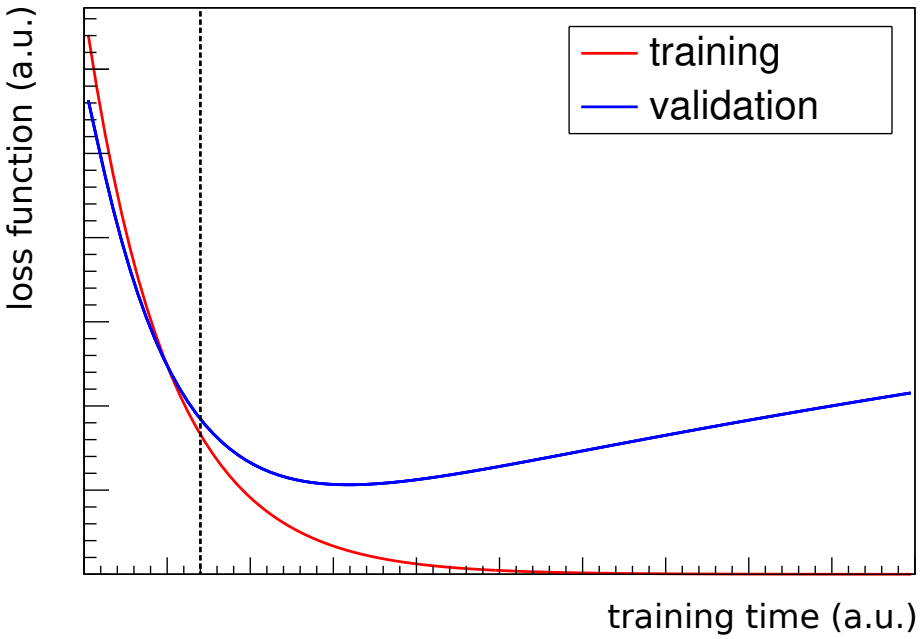


Figure 7.4: **Example for the loss function**. Shown is the loss function during the training of an exemplary ANN. The loss function is evaluated for the training data set (red) and the validation data set (blue). Additionally, the plot indicates a possible time for early stopping (dashed line).

As already discussed, the goal of the ANN training is to learn the underlying characteristics of a given system from data. A common problem within machine learning is the "over-training" effect, where the ANN is not only learning the desired features of the whole system but also learns characteristics specific to the given data set. Since these special characteristics are subject to statistical fluctuations and are in general not representative of the whole system, over-training has to be avoided. The following section briefly reviews the methods used within the scope of this thesis to avoid this problem.

7.3 Methods to avoid over-training

First, the available data is divided into three statistically-independent data sets commonly referred to as training, validation and testing data sets. The first two sets are used within the training, while the latter is used for the final evaluation of the ANN. This treatment ensures that no artificially-augmented separation enters the final discrimination of the ANN, which can arise when data used for the training is evaluated with an over-trained network.

Furthermore, the validation data set can be used as a cross check during the training. An example of such a cross check is shown in Fig. 7.4, which displays the loss function in the different epochs of the training. The loss function falls continuously when evaluated on the training data set, which indicates a successful minimization of the function in the course of the training. However, the loss function increases for the statistically-independent validation data set. This indicates over-training of the ANN, i.e. the network starts to learn specific characteristics of the training data set that are not valid when evaluating the validation data set. In order to avoid this effect, the method of early-stopping [243] is used in this thesis. If the loss function in the two data sets diverges beyond a threshold at a given point during the training, the current state of the ANN is saved. The training is aborted and the saved ANN is used if the difference does not decrease after a given period

of time during the training. Consequently, the training is terminated before the network starts to learn specific characteristics of one data set, which is expected to yield the best description of the general properties of the data sets. The threshold as well as the period of time before the early termination of the training are hyper parameters of the network architecture.

Additionally, the dropout method [244] is used. In this scheme, neurons in the different layers are deactivated at random, which reduces the importance of single neurons. Furthermore, the L2-regularization method [245] is applied, where the loss function is extended to penalize large weights during the training. Both methods are well-established within different fields of machine learning and have been shown to reduce over-training effects.

Finally, the distributions of the predicted values for the different classes are compared between the training and the test data sets. If the shapes of these distributions are not compatible within the statistical precision of the data sets, the ANN is not considered in this analysis.

Part II

ttH Analysis

8 $t\bar{t}H(b\bar{b})$ analysis in the semileptonic channel

After the review of the foundations of this thesis, the $t\bar{t}H(b\bar{b})$ analysis in the semileptonic channel is presented in this chapter. First, the statistical model used for the signal inference is discussed. Subsequently, the analysis strategy and the construction of the final observables used in the measurement are presented. Finally, the results of the inference are discussed.

8.1 Statistical model

As mentioned in Sec. 6, the parameters with respect to which the test statistic is minimized are the parameter of interest (POI) and the set of nuisance parameters (NPs) $\vec{\theta}$. The POI in this analysis is the signal-strength modifier $\mu_{t\bar{t}H}$. It is defined as the fraction of the observed cross section relative to the value predicted by the Standard Model (SM):

$$\mu_{t\bar{t}H} = \frac{\sigma^{\text{obs}}_{t\bar{t}H}}{\sigma^{\text{SM}}_{t\bar{t}H}}. \tag{8.1}$$

This parameter scales the number of signal events in each bin of the discriminants entering the final measurement.

The nuisance parameters are used to represent the systematic uncertainties in the measurement. These parameters can impact both signal and background processes and are correlated where appropriate. The following section briefly summarizes the sources of the uncertainties relevant to the $t\bar{t}H(b\bar{b})$ measurement as well as their parameterization. Further information is provided in Ref. [137].

As introduced in Sec. 6.4, the nuisance parameters are divided into two groups, which are briefly re-introduced here for convenience. Uncertainties that only affect the total yield of processes are called "rate uncertainties". The yield modification introduced by the variation of these parameters follow a log-normal prior probability density function.

Uncertainties that affect both the yield and the shape of the discriminant distributions are referred to as "shape uncertainties". The variations of the yield are internally treated as rate uncertainties. The shape variation is implemented as an interpolation between distributions corresponding to a ± 1 standard deviation (std. dev.) variation of the corresponding uncertainty source, which is discussed in Sec. 6.4.

More information on the modeling of the parameters can be found in Ref. [221]. Table 8.1 gives a summary of the uncertainties used in this analysis, which are discussed in more detail in the following.

Inclusive cross sections: The expected values for the inclusive signal and background cross sections are derived from theoretical predictions of at least NLO accuracy. Uncertainties affecting these normalizations are split into two components due to the choice of α_s. The first component models the effect of this choice on the QCD perturbation theory in the matrix-element calculation (renormalization/factorization scales). The second component is due to this effect in the PDF set used in the calculation of the cross sections (PDF+α_s). These uncertainties are taken into account as rate uncertainties in the final fit. If the sources of these uncertainties are common to multiple processes, they are treated as correlated between the processes. Since the same cross section calculations are used for the simulated samples of the respective processes in the three years of data taking, the uncertainties are treated as correlated between the years.

The specific values for these uncertainties are listed in Tables C.2 and C.3.

Collinear gluon splitting: As mentioned in Sec. 2.3, the calculation of additional gluon splitting effects is complicated and they are not well modeled. Therefore, a rate uncertainty on the $t\bar{t}+2b$ process, which is expected to be especially sensitive to these effects, is constructed to cover uncertainties arising from this issue. It is implemented as a shape uncertainty that assigns a 50 % prior uncertainty to the cross section of $t\bar{t}+b\bar{b}$ events arising from the $t\bar{t}+2b$ process. The challenges in the calculation of these additional splitting effects affect the simulations for the respective years of data taking similarly. Therefore, this uncertainty is treated as fully correlated between the years.

μ_R, μ_F scales: Uncertainties due to the choice of the renormalization scale μ_R and the factorization scale μ_F in the matrix-element generators are modeled by varying the scales independently by factors of 0.5 or 2, respectively. The effects are propagated to the final discriminant distribution in the fit following the recommendations of the experts within the CMS collaboration [246]. Only effects on the shapes of the discriminant distributions are covered by these parameters since the normalization uncertainties of the matrix-element generator are covered by the (N)NLO cross section uncertainties (Table C.3). Additionally, changes due to the μ_R/μ_F variations in the relative fraction of $t\bar{t}+b\bar{b}$ events in the inclusive $t\bar{t}$ phase space are considered with the same merging scheme introduced in Sec. 4.2.1.

The scale-variation uncertainties are implemented as shape uncertainties and are treated as uncorrelated among the $t\bar{t}H$, tHq, tHW, $t\bar{t}+b\bar{b}$ (4FS sample) and other $t\bar{t}$ (5FS sample) processes. The choice of scales is the same for the three years of data taking in the simulated samples for the corresponding processes. Therefore, these uncertainties are correlated across all years.

PDF shape: The shape variation of the final discriminant distributions due to the uncertainty on the PDF set is estimated by reweighting all distributions according to the variations provided with the NNPDF3.0 sets.

The uncertainty is estimated by propagating the effects of the reweighting procedure to the final observables. The final uncertainty is constructed as the RMS or as the quadratic sum of the residuals in each bin of the discriminant distribution, as appropriate for the exact PDF set used for the individual samples. The details of the construction are described in Ref. [247]. Since the overall normalization uncertainties

Table 8.1: **Systematic uncertainties considered in the $t\bar{t}H(b\bar{b})$ analysis**. "Type" refers to rate (R) or shape (S) uncertainties. "Correlation" indicates whether the uncertainty is treated as correlated, partially correlated (as detailed in the text), or uncorrelated across the years 2016–18.

Source	Type	Correlation	Remarks
Renorm./fact. scales	R	correlated	Scale uncertainty of (N)NLO prediction, independent for $t\bar{t}H$, tHq, tHW, $t\bar{t}$, t, V, VV
PDF+α_s (gg)	R	correlated	PDF uncertainty for gg initiated processes, independent for $t\bar{t}H$, tHq, tHW, $t\bar{t}$ + jets and others
PDF+α_s ($q\bar{q}$)	R	correlated	PDF uncertainty of $q\bar{q}$ initiated processes ($t\bar{t}$ + W,W,Z) except tHq
PDF+α_s (qg)	R	correlated	PDF uncertainty of qg initiated processes (single t) except tHW
Collinear gluon splitting	S	correlated	Additional 50% rate uncertainty on $t\bar{t}$ + 2b
μ_R scale	S	correlated	Renormalisation scale uncertainty of the ME generator, independent for $t\bar{t}H$, tHq, tHW, $t\bar{t}+b\bar{b}$ (4FS sample), other $t\bar{t}$ (5FS sample)
μ_F scale	S	correlated	Factorisation scale uncertainty of the ME generator, independent for $t\bar{t}H$, tHq, tHW, $t\bar{t}+b\bar{b}$ (4FS sample), other $t\bar{t}$ (5FS sample)
PDF shape	S	correlated	From NNPDF variations, independent for $t\bar{t}+b\bar{b}$ (4FS sample), tHq, tHW and $t\bar{t}H$/other $t\bar{t}$ (5FS sample)
PS scale ISR	S	correlated	Initial state radiation uncertainty of the PS (Pythia 8.2), independent for $t\bar{t}H$, $t\bar{t}+b\bar{b}$ (4FS sample), other $t\bar{t}$ (5FS sample)
PS scale FSR	S	correlated	Final state radiation uncertainty of the PS (Pythia 8.2), independent for $t\bar{t}H$, $t\bar{t}+b\bar{b}$ (4FS sample), other $t\bar{t}$ (5FS sample)
ME-PS matching ($t\bar{t}$)	R	correlated	NLO ME to PS matching (for $t\bar{t}$ + jets events), independent for additional jet flavours
Underlying event ($t\bar{t}$)	R	correlated	Underlying event (for all $t\bar{t}$ + jets events)
Integrated luminosity	R	partially	Signal and all backgrounds
Lepton ID/Iso	S	uncorrelated	Signal and all backgrounds
Trigger efficiency	S	uncorrelated	Signal and all backgrounds
L1 prefiring correction	S	uncorrelated	Signal and all backgrounds
Pileup	S	correlated	Signal and all backgrounds
Jet energy scale (eleven sources)	S	partially	Signal, $t\bar{t}$ + jets and single t
Jet energy resolution	S	uncorrelated	Signal, $t\bar{t}$ + jets and single t
b-tag HF/LF fraction	S	partially	Signal and all backgrounds
b-tag HF/LF stat (linear)	S	uncorrelated	Signal and all backgrounds
b-tag HF/LF stat (quadratic)	S	uncorrelated	Signal and all backgrounds
b-tag charm (linear)	S	partially	Signal and all backgrounds
b-tag charm (quadratic)	S	partially	Signal and all backgrounds
Size of the simulated samples	S	uncorrelated	Statistical uncertainty of signal and background prediction due to limited sample size

of the PDF are covered by the (N)NLO cross section uncertainties (Table C.3), only the shape and acceptance variations are considered here.

The PDF uncertainties are implemented as shape uncertainties that are correlated according to the used set and the corresponding flavor scheme as introduced in Sec. 2.3. Therefore, this source of uncertainty is treated as uncorrelated between the $t\bar{t} + b\bar{b}$ (4FS sample) on the one hand and the $t\bar{t}H$ and other $t\bar{t}$ (5FS samples) processes on the other hand. The variations are taken from the respective PDF sets `NNPDF31_nnlo_as_0118_nf_4` (4FS sample) and `NNPDF31_nnlo_hessian_pdfas` (5FS samples). Additionally, there are PDF uncertainties for tHq and tHW, separately, which are derived from the quadratic sum of all residuals. The individual PDF uncertainties are correlated across all years since the same set is used for the simulation of the respective processes in all considered years of data taking.

Parton shower scales ISR/FSR: In order to evaluate the impact of the choice of α_s in the parton shower (PS) simulation, the corresponding scale parameters in the shower simulation are varied by factors of 0.5 and 2, as described in Ref. [246]. This is done independently for the initial-state radiation (ISR) and final-state radiation (FSR) showers. Only the effects on the overall shapes of the final observables are considered with these uncertainties since the normalization uncertainties of the matrix-element generator are covered by the (N)NLO cross section uncertainties (Table C.2). Similarly to the μ_R and μ_F scale uncertainties, changes in the relative fraction of $t\bar{t} + b\bar{b}$ events to the inclusive $t\bar{t}$ phase space are considered based on the composition of events in the $t\bar{t}$ (5FS) samples using the merging scheme explained in Sec. 4.2.1.

The ISR and FSR uncertainties are treated as uncorrelated among the $t\bar{t}H$, $t\bar{t} + b\bar{b}$ (4FS sample) and other $t\bar{t}$ (5FS sample) processes to account for the separate generation of the corresponding samples. They are correlated across all years since the parton shower for the corresponding simulations is the same.

ME-PS matching: The uncertainty of the matching between the matrix element generator Powheg and the parton shower Pythia 8.2 is estimated by using the dedicated samples listed in Tab. 5.4. To generate these simulations, the matching parameter ME-PS matching is varied following the methodology presented in Ref. [248]. The values used for the simulations correspond to the uncertainties in the tuning of the ME-PS matching energy scale, which is described in Refs. [156, 157].

Ideally, the ME-PS matching uncertainties would be considered as shape uncertainties by constructing the distributions of the final discriminant from the corresponding dedicated simulations. However, as indicated in Tab. 5.4, only a limited number of events is available for the corresponding $t\bar{t} + b\bar{b}$ (4FS) and $t\bar{t}$ (5FS) samples. As a consequence, the distributions of observables derived from these samples are subject to large statistical fluctuations. The variations between these distributions and the nominal predictions are statistically compatible with a pure rate variation. Therefore, the average variation across the years is estimated per $t\bar{t}$ + jets process and per analysis category, which are introduced in Sec. 8.2. This results in a rate change of typically 3–10 % with rare outliers of up to 20 %, which is comparable to the size of the statistical fluctuations. Finally, these variations are implemented as rate uncertainties in the final statistical model.

The uncertainties are treated uncorrelated among the different $t\bar{t}$ + jets background processes ($t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$, $t\bar{t}$ + lf) and correlated per process between the analysis

categories and channels. Since the variations of these parameters are derived as the average across the years, each of these uncertainties is fully correlated across the years.

Underlying event: The effect on the production cross section of the $t\bar{t}$ + jets processes due to uncertainties in the underlying event tune are estimated using dedicated samples listed in Tab. 5.4. As stated before, the settings for these samples correspond to the uncertainties in the derivation of the tune described in Refs. [156, 157]. The relative effect of the underlying event tune is derived from the $t\bar{t} + b\bar{b}$ (5FS) prediction and applied to the nominal 4FS prediction.

Similar to the ME-PS matching uncertainties, the effect of the underlying event tune precision would ideally be considered as shape uncertainties in the statistical model. However, the same issues arise in the processing of these samples as discussed in the previous paragraph. Therefore, the same procedure as described above for the ME-PS matching scale uncertainties is used to mitigate the effects of statistical fluctuations caused by the limited number of events in the dedicated samples.

The final rate uncertainties are treated as fully correlated among the different $t\bar{t}$ + jets background processes across all categories and channels since the underlying event tune should be independent of the arising final state. The variations of these parameters are averaged across all years of data taking and are therefore treated as fully correlated across the years.

Luminosity: The measurement of the instantaneous and integrated luminosity is performed by a dedicated group within the CMS collaboration and is essential to correctly normalize the simulated samples for the final comparison to the observed data set. The uncertainty in this measurement is considered following the recommendation of the CMS group [249] and is listed in Table C.4. This affects the rates of all processes. The uncertainties are partially correlated across the years, following the recommendation.

Lepton scale factors (tracking, reconstruction, ID, isolation): As discussed in Sec. 5.2, p_T- and η-dependent lepton scale factors are used to improve the compatibility for the tracking (only electrons), reconstruction, ID, and isolation efficiencies of the simulated samples and the observed data set. The uncertainties in the derivation of these scale factors are considered following the recommendation provided by the CMS groups responsible to the reconstruction of muons [190–192] and electrons [195].

The impact of the lepton scale factor uncertainties on the final result is negligible. Therefore, the uncertainties are propagated to the final discriminant distributions as a total shape uncertainty which corresponds to the quadratic sum of the statistical and systematic components of all sources, separately, for electrons and muons. The scale factors are derived for each year individually and are therefore decorrelated across the years.

Trigger efficiency: The uncertainties in the derivation of the trigger scale factors introduced in Sec. 5.2 are also considered in the statistical model. Their impact on the final discriminants is estimated by varying the trigger scale factors within their uncertainties. More details about the trigger efficiencies can be found in Ref. [137]. Similar to the lepton scale factors, the trigger scale factors are derived for each year individually. Consequently, these uncertainties are treated as uncorrelated across the years.

L1 prefiring correction: As mentioned in Chap. 5.2, the effects due to the L1 prefiring issue introduced in Sec. 3.3.4 are corrected by applying dedicated scale factors to

the simulated events. The uncertainties on this correction are taken into account by varying the corresponding scale factor within its uncertainties, which is described in Ref. [124]. The L1 prefiring correction is decorrelated between the 2016 and 2017 data sets and does not exist for 2018.

Pileup: As discussed in Sec. 5.2, the distribution of the number of pileup events in the simulated samples is reweighted to improve the agreement with the observed distribution. Effects due to the uncertainty of this method are evaluated by varying the cross section used to predict the number of pileup interactions in the simulation within its uncertainty of 4.6 % [111]. The changes in the weight factors are propagated to the final discriminant distributions and treated as shape parameters capable of both rate and shape variation in the final fit. The uncertainty is treated as fully correlated among all processes and years since the underlying procedure and its inputs is the same.

Jet Energy Scale (JES): As discussed in Sec. 5.2, the JES is corrected to exploit the latest calibrations of the detector. Each part of the correction procedure is subject to sources of uncertainty. Following the recommendation by the physics analysis group at CMS dedicated to Higgs-boson measurements, there is a total of eleven JES sources considered in this analysis. The uncertainties due to the JES are evaluated by shifting the jet energy scale applied to the reconstructed jets up and down by one std. dev., which are provided centrally by the corresponding group within the CMS collaboration [250]. The events are then re-analyzed to determine varied final discriminant distributions. They are considered as shape uncertainties in the final fit. The correlation of these sources across the different years follows the recommendation by the experts.

In addition, a dedicated uncertainty source is applied simulated samples corresponding to 2018 to cover the effects of the HEM issue described in Sec. 3.3.4, which have a minor impact in this analysis. Following the recommendation by the group of experts in the CMS collaboration, a dedicated jet energy variation

- of 20 % for jets with $-1.57 < \phi < -0.87$ and $-2.50 < \eta < -1.30$, and
- of 35 % for jets with $-1.57 < \phi < -0.87$ and $-3.00 < \eta < -2.50$

is applied to the simulated events. Details can be found in Refs. [125, 126].

Large statistical uncertainties in the resulting varied distributions of the final discriminant can cover possible shape variations. Additionally, such fluctuations are known to artificially constrain the corresponding parameters in the parameter estimation, which is discussed in previous measurements in the $t\bar{t}H(b\bar{b})$ channel at the CMS experiment presented in Refs. [21, 133]. Therefore, these uncertainties are not applied for processes with small contributions in the analysis phase space. It has been shown that this modification does not impact the sensitivity and constraints of the signal-strength modifier and the corresponding nuisance parameters as outlined later in this section.

Jet Energy Resolution (JER): The systematic uncertainty due to the JER described in Sec. 5.2 is evaluated by increasing/decreasing the difference between the reconstructed and particle-level jet energy according to the prescription by the experts within the CMS collaboration, which can be found in Ref. [201]. The events are then re-analyzed to determine varied final discriminant distributions. This source is considered as a shape uncertainty in the final fit which is fully decorrelated across the years following the recommendation by the experts.

Similarly to the JES uncertainties, the magnitude of the statistical uncertainty in the JER sources can be large depending on the considered process. Therefore, the treatment for this set of uncertainties is the same as previously described for the JES uncertainties and later in this section.

b-tag scale factors: The procedure to derive the scale factors (SFs) for the b-tag score described in Sec. 5.2 is subject to three sources of systematic uncertainty on both the heavy flavor (HF) and light flavor (LF) scale factors:

Jet Energy Corrections (JEC): The b-tag uncertainty associated with the different JEC sources is evaluated at the same time as the corresponding source of JEC uncertainty (JES or JER). When the kinematics of the jets are shifted up or down by one std. dev. according to the JEC, the b-tag SF values are also shifted accordingly.

purity (HF/LF fraction): The uncertainties on the purity of the control samples used to derive the scale factors are modeled by a separate nuisance parameter for LF and HF jets. Shifting this parameter to ± 1 std. dev. corresponds to a change of the HF (LF) SF as caused by a higher or lower contamination by LF (HF) jets in the control sample. Due to the iterative nature of the scale factor calculation, this also has a small effect on the LF (HF) SFs.

HF/LF statistics: The impact of statistical uncertainties associated with the size of the samples used to derive the SFs are controlled by means of four nuisance parameters, two for HF and two for LF jets. These parameters correspond to the degrees of freedom in the interpolation of the SFs described in Sec. 5.2. The statistical uncertainties are treated as decorrelated between the years since they originate from the size of independent control samples.

It is not possible to perform the same procedure for charm jets due to the difficulties in constructing an adequate control region. Therefore, the linear and quadratic components of the uncertainties from the heavy flavour SFs are added in quadrature to the purity component, respectively. These values are doubled in size and used to construct two separate nuisance parameters to control these uncertainties. The uncertainties associated with charm jets are treated independently with respect to all the uncertainties for the HF and LF SFs. The purity uncertainties as well as the uncertainties for the charm quark jets are decorrelated between 2016 and the other years of data taking to account for the upgrade of the pixel subsystem in the CMS detector in 2017 mentioned in Sec. 3.2.2.

Size of the simulated samples: The limited size of the simulated samples results in statistical fluctuations of the nominal prediction. This is taken into account via the Barlow-Beeston-lite method [251]. Technically, the implementation is provided with the centrally-provided fitting tool within the CMS collaboration, which is described in more detail in Ref. [252]. For bins with more than ten effective entries, a single Gaussian-constrained nuisance parameter is introduced in each bin that scales the total yield; for bins with fewer entries, a separate Poisson pdf is used for each process. Furthermore, following the recommendations of the tool's authors, the contributions of the signal distributions are not taken into account when evaluating whether to use a single Gaussian nuisance parameter or a Poisson constrained parameter for each process. The rate and shape effects are treated independently in the morphing algorithm introduced in Sec. 6.4 and further described in Refs. [221, 226].

In addition to these sources of uncertainty, the normalizations of the $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ and $\mathrm{t\bar{t}} + \mathrm{c\bar{c}}$ processes are estimated in situ simultaneously with the measurement of the $\mathrm{t\bar{t}H(b\bar{b})}$ process.

This accounts for the challenging calculation of these background processes and enables the statistical model to mitigate differences between simulation and data. Such differences are expected due to previous dedicated measurements as presented in Refs [24, 25, 253].

In order to improve the stability of the final statistical model, the following modifications are applied to the systematic uncertainties. Since the statistical fluctuations of the distributions for the minor background processes tH, $t\bar{t}+V$, $V+\text{jets}$ and Diboson are large in the analysis phase space, the JEC are not considered for these processes. Additionally, each shape uncertainty is checked for a "real" shape effect on the distributions as defined in Ref. [252]. If a given uncertainty has no shape effect, it is transformed into a rate uncertainty. Finally, only rate uncertainties with an effect of at least 0.1 % on a given process are considered for this process. These checks are performed separately for each of the analysis categories in the final fit. It has been shown that this procedure stabilizes the expected results of the analysis without changing the expected sensitivity significantly [254].

As discussed in Chap. 6, the POI and the nuisance parameters constructed from the individual sources of uncertainty are estimated by minimizing the test statistic for the parameter estimation defined in Eq. (6.7). In the following section, the observables entering the final fit are discussed.

8.2 Observables used for the parameter estimation

Choosing the right observable for the parameter estimation is of crucial importance in order to optimize the sensitivity of the $t\bar{t}H(b\bar{b})$ analysis. This section will motivate the strategy chosen for this analysis by illustrating the steps with a small example.

8.2.1 Construction of Analysis Regions

Ideally, the observable for a statistical analysis should be able to perfectly distinguish between signal and background events. Given the $H \to b\bar{b}$ final state that is the focus of this analysis, one would expect the reconstructed invariant mass of the two b-tagged jets emerging from the Higgs boson decay to be a very sensitive observable. For the signal, it should show a narrow resonance around the mass of the Higgs boson, while it should be a continuous distribution for the background processes. Therefore, the probability of finding a jet pair with a mass close to m_H should be much higher for signal events.

In reality, this is not the case, as is illustrated by Fig. 8.1. This is due to many factors, such as the much larger background rate and experimental (in-)efficiency effects in the jet reconstruction and the b-tagging. As a result, there is no strong separation of signal and background events in this observable. This of course means that this discriminator is not suited for a powerful statistical analysis.

The construction of suitable observables was thoroughly studied in the scope of the full Run-II $t\bar{t}H(b\bar{b})$ analysis [137]. In addition to the selection criteria presented in Sec. 4.3, the events are separated by their b-tag multiplicities: one region with exactly three b-tags ($\geq$ 4j, 3t) and another with at least four b-tags ($\geq$ 4j, $\geq$ 4t). The first category is enriched in background events, mostly originating from the $t\bar{t}+\text{lf}$ process. The second is enriched in signal and $t\bar{t}+\text{heavy flavor}$ events. As Fig. 8.2 illustrates, the relative contribution of signal events increases dramatically in the $\geq$ 4j, $\geq$ 4t category. However, this increase in signal purity comes at the cost of distributing the available events into the two analysis regions, thus decreasing the statistical precision. This strategy has proven to yield a higher signal purity while also enabling a good control over the background processes and the relevant uncertainties.

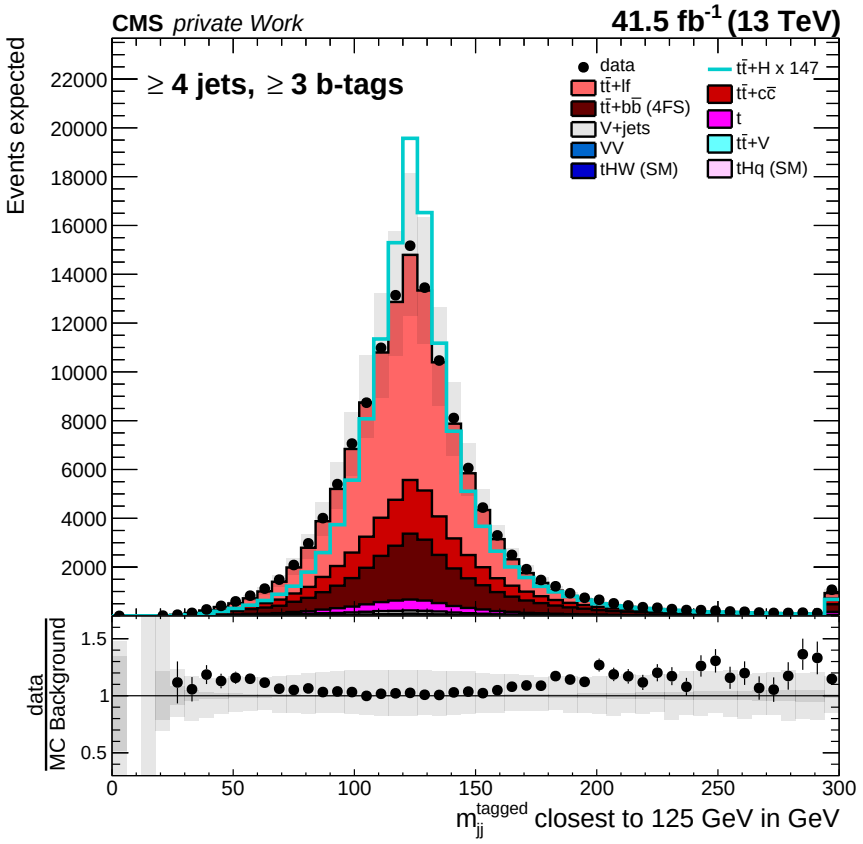


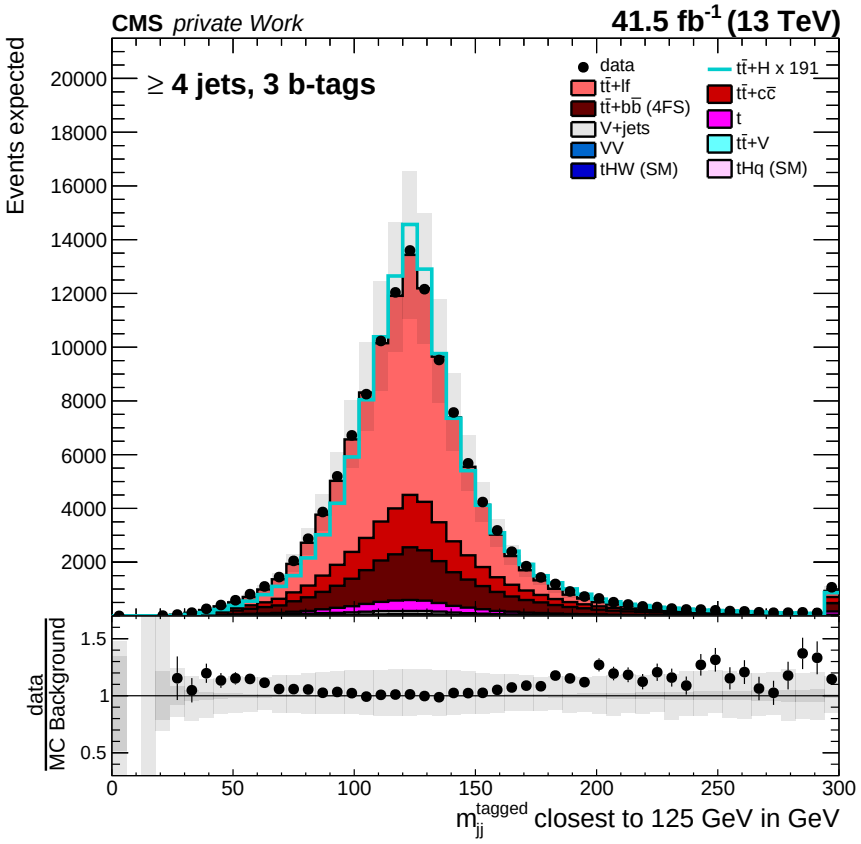
Figure 8.1: **Reconstructed invariant mass of the two b-tagged jets closest to 125 GeV**. The stacked histograms show the distributions of the background processes. For a better comparison, the expected signal process (cyan line) is scaled to the total yield of the background processes. The bottom half of the figure shows the ratio of observed data with respect to the background prediction. The light gray band indicates the systematic uncertainties, while the dark gray band illustrates the statistical uncertainties in each bin of the background prediction. The black dots indicate the observed data.

Nevertheless, even after constructing a dedicated signal region, the reconstructed mass of two b-jets closest to m_H is unable to perform a powerful separation of the signal and background events. Therefore, while the overall purity increased, each bin still has a dominant background contribution, thus limiting the sensitivity of a statistical analysis.

Nowadays, it is very common in high-energy physics to harness the power of multivariate analysis methods to improve the sensitivity of statistical analyses. In this analysis, the final observables are constructed with Artificial Neural Networks (ANNs), as discussed in Sec. 7.

8.2.2 Construction with artificial neural networks

For the purpose of discriminating signal-like events from background-like events, multi-classification ANNs as introduced in Chap. 7 are trained in both analysis regions. The training is based on the simulated predictions for the relevant physics processes. In order to increase the number of events available for training, the networks are trained on the simulations of 2016, 2017 and 2018 simultaneously. This has been shown to improve the overall stability of the ANN training [255].



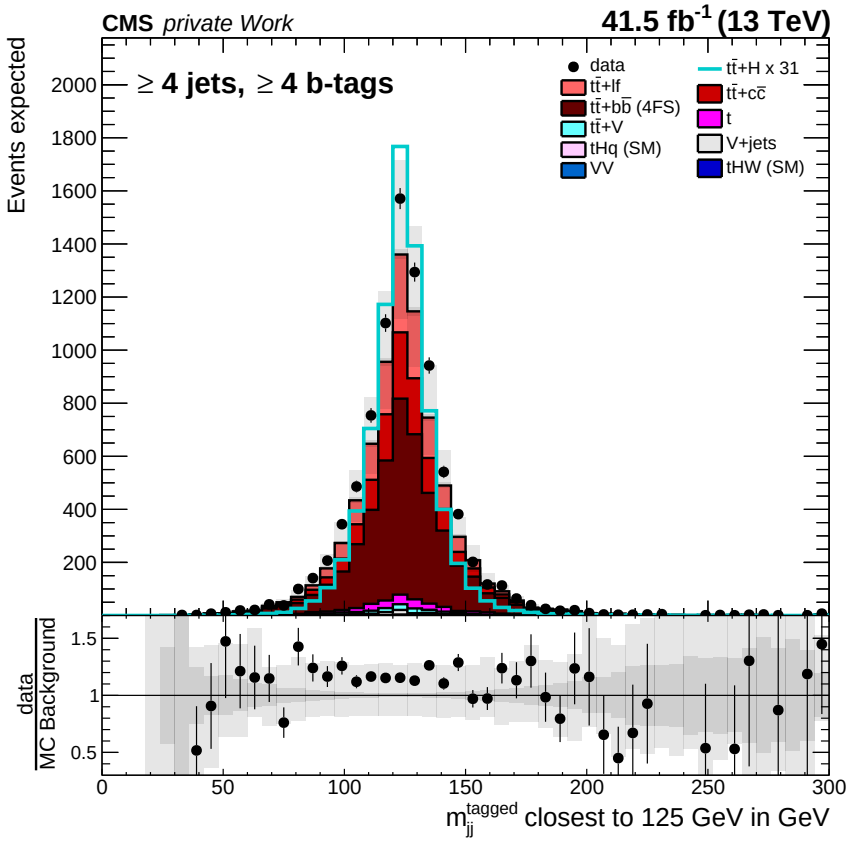


Figure 8.2: **Reconstructed invariant mass of the two b-tagged jets closest to 125 GeV in both analysis regions**. The events in Fig. 8.1 are sorted into the two analysis regions. The description of the figure is the same as for Fig. 8.1.

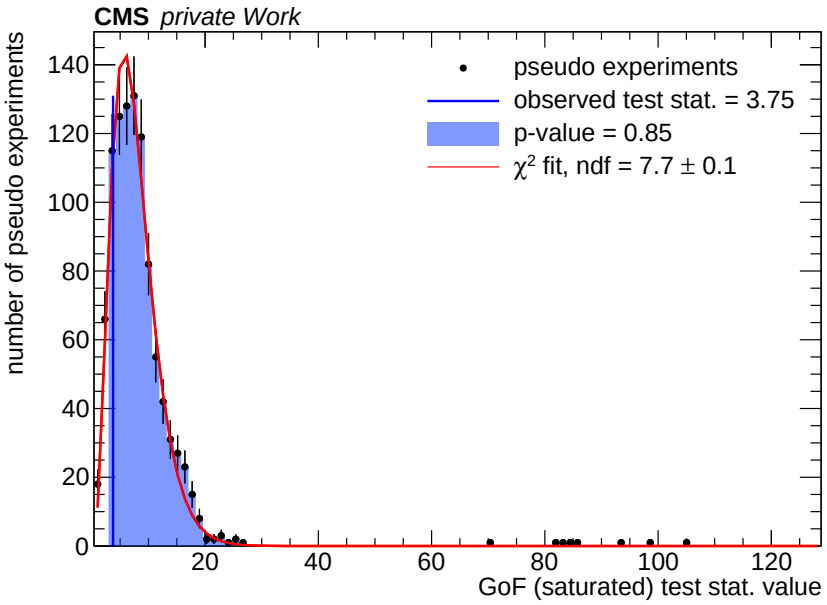


Figure 8.3: **GoF test for the jet multiplicity in the $\geq$ 4j, $\geq$ 4t category in 2016 Data**. First, the probability density is calculated using 1000 a-posteriori pseudo experiments (black dots). Afterwards, the test statistic is calculated for the observed data (blue vertical line). The p-value is the integral of the probability density starting at the value of the observed data (shaded blue area). For comparison, the red line indicates the corresponding χ^2 test statistic.

The training of the classifiers depends entirely on the simulated samples. Therefore, it is of crucial importance to ensure good modeling of the observables used for the training in data. In order to be able to give a quantitative measure of the quality of the description, a Goodness-of-Fit (GoF) test is performed. In this method, a test statistic is constructed using the so-called "saturated" model, which is the generalization of the Chi-square test statistic. More information about the GoF test and the saturated model can be found in Refs. [252, 256].

First, this test statistic is evaluated in observed data. Additionally, the test statistic is evaluated in a posteriori pseudo experiments, which yields the probability density for the test statistic given the statistical model. This strategy allows to quantify the description of the simulated samples in data by providing a p-value. The p-value is the probability to observe the value of the test statistic calculated in observed data or a larger value given the probability density in the statistical model. Consequently, large p-values represent good agreement between the observed data and the simulated samples, while small values indicate a poor description of the observed data. The procedure is illustrated for an exemplary observable in Fig. 8.3. The modeling of observables with a p-value below 5 % is considered too poor. Therefore, these observables are not considered further in the analysis. This test is used to validate both one-dimensional distributions and the correlation between each pair of observables. Only observables passing this GoF test are used to train the ANNs. The summary of the GoF test for the final set of observables used for training can be found in Appendix E.1.

The final set of inputs for the ANNs are quantities related to the characteristic physics of the processes of interest. This includes kinematic variables such as the average transverse momentum of b-tagged jets, but also more sophisticated observerables like the output of

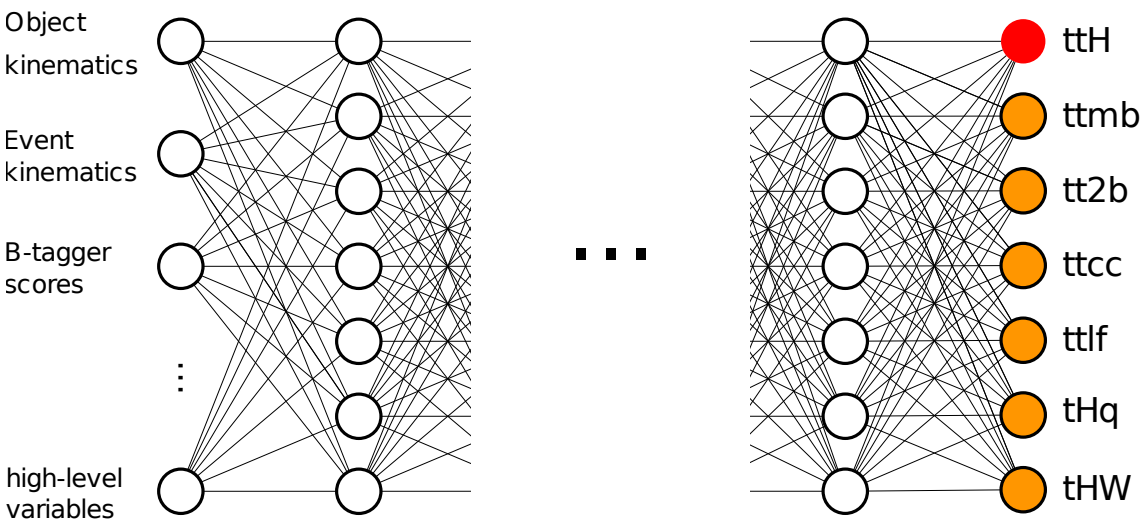


Figure 8.4: **Schematic depiction of the ANN architecture**. Each event is associated with one of the target classes. They contribute information about the physics of the processes with information such as the properties listed on the left-hand side of the illustration. Each event is categorized as one of the target classes listed on the right-hand side.

the Matrix Element Method (MEM) classifier [257]. Additionally, the output of boosted decision trees dedicated to the reconstruction of the $t\bar{t}H$, tHq, tHW and $t\bar{t}$ systems are included in the ANN. The complete list of observables used for the training of the ANNs can be found in Tab. D.7. These well-described observables are the foundation of the ANN training.

The target of these multi-classification networks is to assign a given event to a specific process class. Each of the classes corresponds to physics processes that are of special interest in the analysis phase space. The classes can be grouped in Higgs-related ($t\bar{t}H$, tHq, tHW) and $t\bar{t}$-related classes ($t\bar{t}$mb, $t\bar{t}+2b$, $t\bar{t}+c\bar{c}$, $t\bar{t}+lf$). Consequently, each network assigns the respective events to one of these seven categories. Figure 8.4 illustrates the setup of the ANNs. Events are attributed to the process class with the highest ANN output y.

The hyperparameters of the architecture of the networks are listed in Tab. D.5. The parameters are chosen such that the separation power of the different nodes is good while ensuring that there is no overtraining on the training sample. During the training, the events of the different samples are weighted such that the relative frequency of the different classes is equal. This ensures that the network does not simply attribute the events to the category with the highest occurrence, e.g. $t\bar{t}+lf$.

Figure 8.5 shows the metrics of the ANN in the $\geq$ 4j, 3t region. The loss function indicates a smooth training and, more importantly, no overtraining on the training sample. The confusion matrix indicates a good separation for the tHq and tHW processes. This is expected since the b-tag requirement matches the diagrams for tH production. Furthermore, the confusion matrix indicates a good categorization for the $t\bar{t}+lf$ class. This is also by construction since the relatively low b-tag requirement should filter most of the $t\bar{t}H(b\bar{b})$-like events and should thus be less sensitive to signal-like processes. Due to the low b-tag multiplicity, observables in this analysis region contain information related to especially $t\bar{t}+lf$-like events, which leads to a good separation of this class and the others.

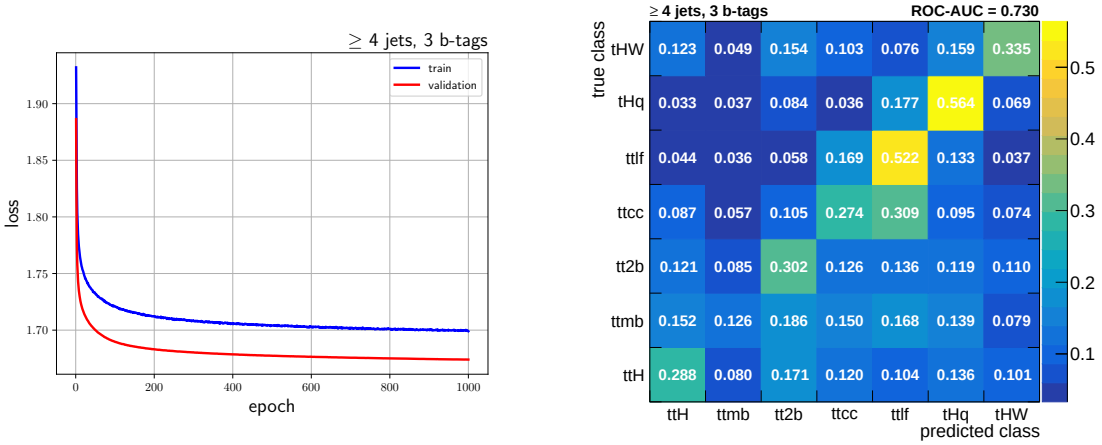


Figure 8.5: **ANN metrics in the ≥ 4j, 3t category**. The loss function (left) shows a smooth decline of the loss for both training and validation sample, thus indicating no overtraining. The confusion matrix (right) shows large values on the diagonal axis, indicating a good separation in the different output nodes.

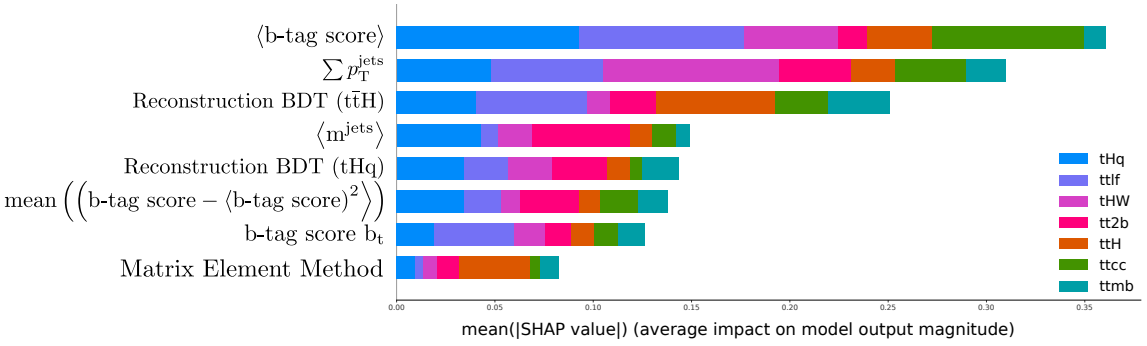


Figure 8.6: **Ranking of the input features in the ≥ 4j, 3t ANN**. The x-axis shows the mean of the absolute Shapley values of 10 000 events. The different colors of the bars indicate the dependence of the given category class on the respective input feature. Large values indicate a high importance.

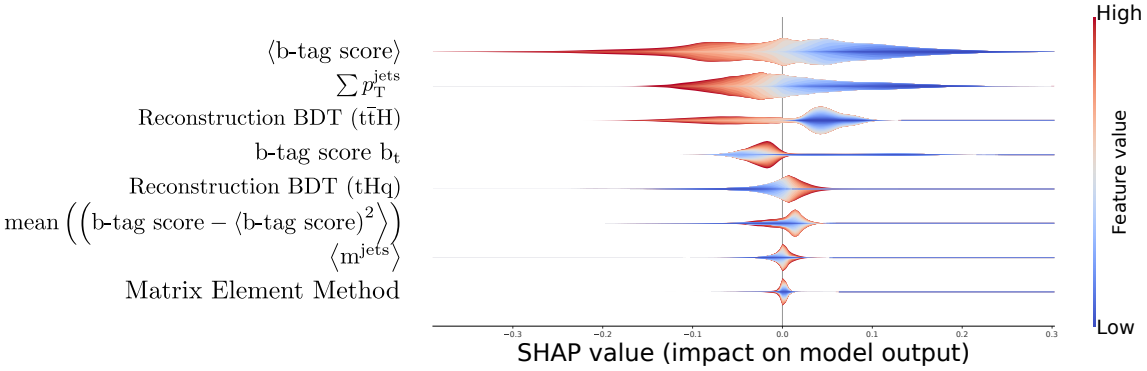


Figure 8.7: **Shapley values for the $t\bar{t}$ + lf node in the ≥ 4j, 3t region**. The values are calculated with 10 000 events. The y-axis shows the input features of the ANN in the ≥ 4j, 3t region. The z-axis indicates low values (blue) or high values (red) of the respective input feature. The x-axis shows the Shapley metric. The frequency of a given Shapley value is indicated by the line width at the corresponding position.

The classes related to $t\bar{t}$ production in association with heavier objects such as bottom quarks, charm quarks and the Higgs boson show a high degree of confusion with other classes. This is expected by construction. As mentioned above, the relatively low b-tag requirement depletes this analysis region of most of the $t\bar{t} + b\bar{b}$ and $t\bar{t}H(b\bar{b})$ events. The classification of the $t\bar{t} + c\bar{c}$ category is especially challenging due to the high mistag probability of charm quarks. Additionally, the kinematic range of the charm quarks is much closer to the additional light-flavor jets than to the much heavier b-jets. This leads to a high confusion of the $t\bar{t} + c\bar{c}$ and $t\bar{t}$ + lf classes.

When using ANNs in a statistical analysis, it is crucial to ensure that the neural network indeed learns the underlying physics of a process class. Two methods to calculate the importance of the input features are used in this analysis. On the one hand, a Taylor expansion of the ANN is performed as proposed in Ref. [258]. In this method, the ranking of each node is based on the first-order derivatives of each class with respect to the individual input features. On the other hand, the ranking is performed with the Shapley metric [259]. This method originated in game theory and calculates the importance of an input feature in the ensemble by calculating the Shapley metric with and without a given observable, thus gauging the impact on the ANN. Both measures show a very similar ranking of the input features. In the following, the Shapley values are discussed in more detail.

Figure 8.6 shows the ranking based on Shapley values for the ANN in the ≥ 4j, 3t category. Overall, the importance of the different observables to the different process classes is reasonable. For example, the most important input feature for the neural network is the average b-tag score of the jets, which is most important to the tH, $t\bar{t}$ + lf and $t\bar{t} + c\bar{c}$ classes. Due to the topology of these processes, they are expected to have no more than three b-tagged jets and more light-flavor jets which ideally have a small b-tagging score. Therefore, the average b-tag score in these classes should be smaller than e.g. for the $t\bar{t} + b\bar{b}$ process. Figure 8.7 illustrates the dependency of the Shapley values on the input features for the $t\bar{t}$ + lf node. Small values of the average b-tag score lead to high Shapley values, and vice versa. As was just explained, this matches perfectly with the expectation and the impact of this observable on the classification of events in the $t\bar{t}$ + lf node indicated by the Shapley metric is reasonable. The importance ranking also shows the poor separation of the $t\bar{t}H$ and $t\bar{t} + c\bar{c}$ classes. There are very few input features with a large impact on these

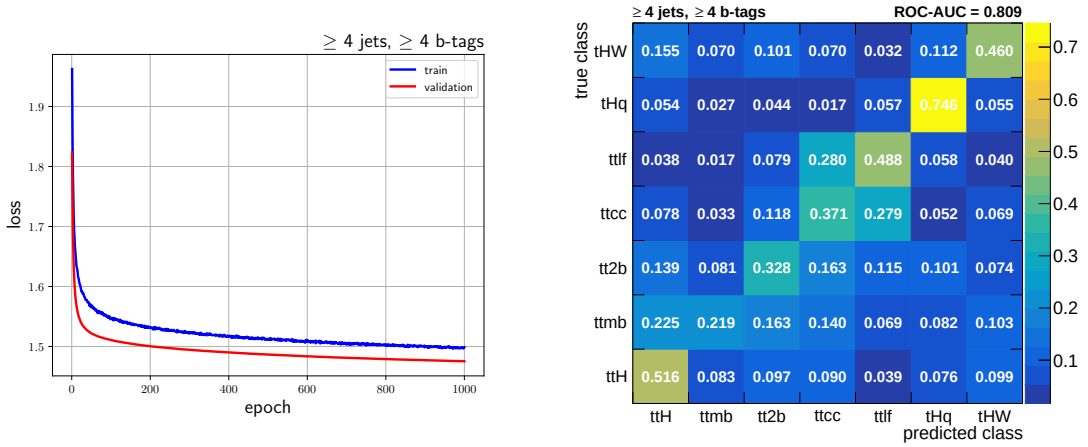


Figure 8.8: **ANN metrics in the ≥ 4j, ≥ 4t category**. The loss function (left) shows a smooth decline of the loss for both training and validation sample, thus indicating no overtraining. The confusion matrix (right) shows large values on the diagonal axis, indicating a good separation in the different output nodes.

classes. This in turn means that none of these variables yield a heightened sensitivity to these classes, thus resulting in the small categorization efficiency indicated in Fig. 8.5.

Figure 8.8 shows the metrics of the ANN in the ≥ 4j, ≥ 4t region. Like the metrics for the ANN in the ≥ 4j, 3t category, the loss function indicates a smooth training and no overtraining on the training sample. The confusion matrix in this region shows excellent identification efficiency for the tH nodes. This is due to the large differences in the topology of these processes with respect to the other classes. For a similar reason, the $t\bar{t}$ + lf class shows good categorization efficiency. However, much like in the ≥ 4j, 3t region, there is a high degree of confusion with the $t\bar{t}$ + $c\bar{c}$ class, which can be understood following the arguments in the previous discussion.

The $t\bar{t}$ + $c\bar{c}$ class shows a higher selection efficiency with respect to the ANN in the ≥ 4j, 3t category. This is due to the fact that the mistag probability for c-jets is much higher than for light-flavor jets, which in turn leads to a higher relative occurrence of $t\bar{t}$ + $c\bar{c}$ events with respect to $t\bar{t}$ + lf events in the ≥ 4j, ≥ 4t region. This in turn allows for a better separation between the $t\bar{t}$ + $c\bar{c}$ and $t\bar{t}$ + lf classes. On the other hand, the kinematic characteristics of the $t\bar{t}$ + $c\bar{c}$ process are different to the heavier final states emerging from the $t\bar{t}$ + $b\bar{b}$ and $t\bar{t}H(b\bar{b})$ processes, thus overall improving the $t\bar{t}$ + $c\bar{c}$ classification.

Since the ≥ 4j, ≥ 4t is enriched in $t\bar{t}H(b\bar{b})$, the identification efficiency for $t\bar{t}H(b\bar{b})$ is much larger with respect to the ≥ 4j, 3t. On the other hand, the confusion matrix also shows a large confusion of the $t\bar{t}H(b\bar{b})$ class with the $t\bar{t}$ + $b\bar{b}$-related classes. As discussed in Chap. 4, this is expected since the $t\bar{t}$ + $b\bar{b}$ process shares many of the kinematic properties with the signal process, thus making it an irreducible background. Therefore, the ANN reflects the underlying physics reasonably well.

Figure 8.9 shows the ranking of the input features based on the Shapley values. Much like in the ≥ 4j, 3t ANN, the ranking indicates that the neural network indeed learns features of the underlying physics of each category class. For example, the most important observables for the $t\bar{t}H(b\bar{b})$ class are the output of a dedicated reconstruction BDT for the $t\bar{t}H$ system and of the MEM. Both methods are designed specifically for identifying the

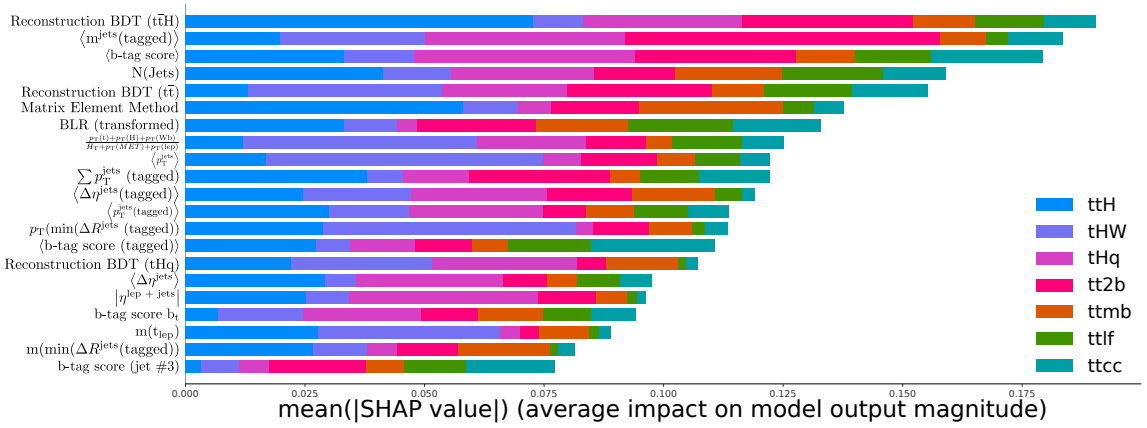


Figure 8.9: **Ranking of the input features in the $\geq$ 4j, $\geq$ 4t ANN**. The x-axis shows the mean of the absolute Shapley values of 10 000 events. The different colors of the bars indicate the dependence of the given category class on the respective input feature. Large values indicate a high importance.

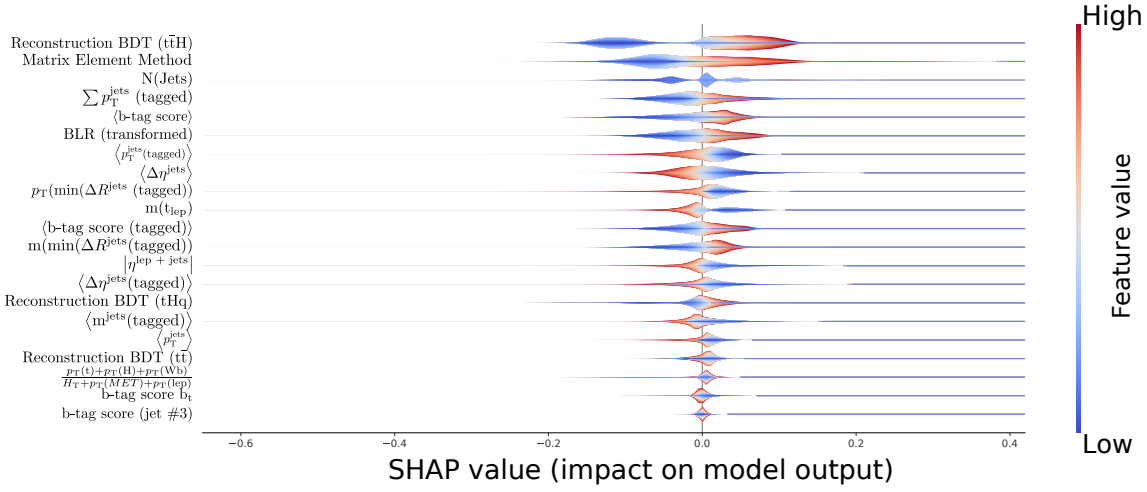


Figure 8.10: **Shapley values for the tt̄H node in the $\geq$ 4j, $\geq$ 4t region**. The values are calculated with 10 000 events. The y-axis shows the input features of the ANN in the $\geq$ 4j, $\geq$ 4t region. The z-axis indicates low values (blue) or high values (red) of the respective input feature. The x-axis shows the Shapley metric. The frequency of a given Shapley value is indicated by the line width at the corresponding position.

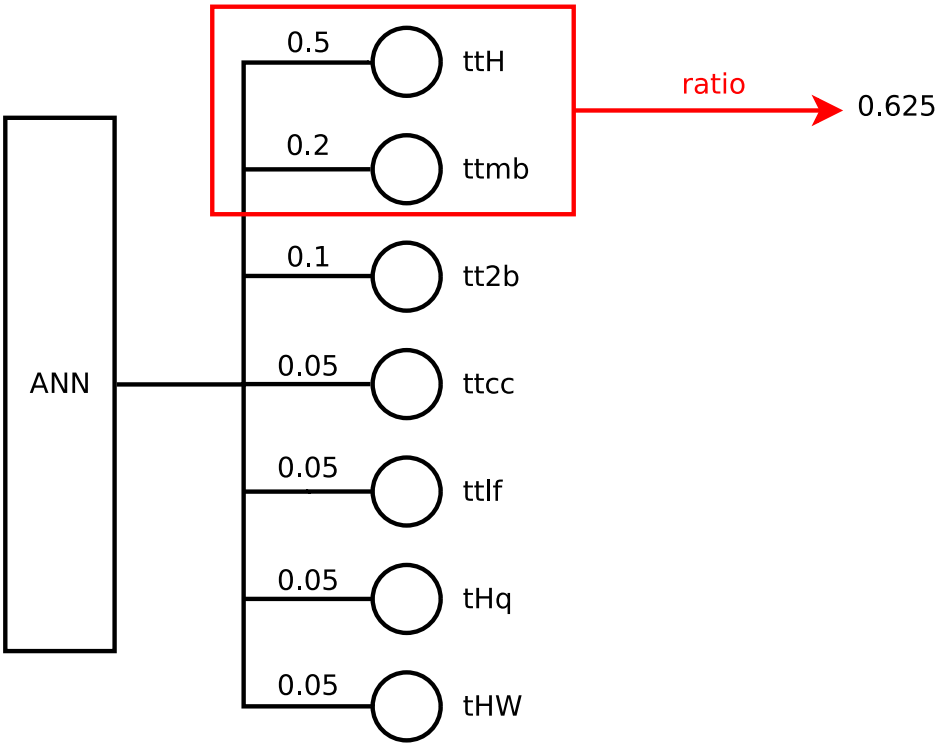


Figure 8.11: **Calculation of the ratio observable in the $\geq$ 4j, $\geq$ 4t region**. The ANN is evaluated for each event. If an event is categorized as either $t\bar{t}H$ or $t\bar{t}mb$, i.e. has the largest output value in either of these classes, the ratio in Eq. 8.2 is calculated. The distribution of these values enter the final fit.

$t\bar{t}H$ system and, in the case of the MEM, to even distinguish between $t\bar{t}H(b\bar{b})$ and $t\bar{t}+b\bar{b}$ events as much as possible. Therefore, it is reasonable that both features have a high impact on the identification of $t\bar{t}H(b\bar{b})$ events. This is also illustrated in Fig. 8.10. Large values of these observables lead to high values in the $t\bar{t}H$ category and vice versa. Most of the other observables describe information about b-tagged jets and their properties, such as transverse momentum, mass or b-tag score. Given the final states of the processes of interest, this ranking of the observables demonstrates an excellent representation of the underlying physics.

As shown in Fig. 8.8, the confusion between the $t\bar{t}H$ and the $t\bar{t}+b\bar{b}$-related classes is high. In order to yield a more powerful separation of these classes, the construction of likelihood ratios based on the ANN outputs in these classes was investigated. The figures of merit were the signal purity in each bin of the observable, the resulting sensitivity in the final fit, the impact of the systematic uncertainties on the signal-strength modifier, the constraints of the parameter estimation as well as bias tests with pseudo experiments using different $t\bar{t}+b\bar{b}$ hypotheses. The studies are summarized in Ref. [260].

In the end, the following ratio of ANN output values y proved to yield the most stable and most powerful results:

$$\text{Ratio} = \frac{y(t\bar{t}H)}{y(t\bar{t}H) + y(t\bar{t}mb) + y(t\bar{t}+2b)}. \tag{8.2}$$

This ratio observable is constructed for events that are classified as either $t\bar{t}H$ or $t\bar{t}mb$ in the signal-enriched $\geq$ 4j, $\geq$ 4t region. The workflow for the calculation is illustrated in Fig. 8.11. The tests to validate the ratio observable are summarized in App. G. The

main benefit of this ratio is the higher signal purity in each bin of the final distribution. This allows for a smaller number of bins in the final fit with higher statistics, which in turn stabilizes the shapes used to construct the previously-discussed shape uncertainties that would otherwise suffer from larger statistical fluctuations. Overall, this improves the sensitivity of the semileptonic analysis by approximately 30 % and makes the parameter estimation more robust.

In the end, 13 event categories enter the final fit. In the $\geq$ 4j, 3t region, the distributions of the seven output nodes of the ANN enter. In the $\geq$ 4j, $\geq$ 4t, the distributions of the $t\bar{t}$ + lf, $t\bar{t}$ + $c\bar{c}$, tHq, tHW and $t\bar{t}$ + 2b nodes as well as the distribution of the ratio observable enter. The binning of these distributions has been chosen such that each bin has at least 15 background events and the relative statistical uncertainty of each bin is below 10 %, which allows for meaningful statistical analysis. The following section will discuss the results of the analysis.

8.3 Results

After discussing the strategy for the semileptonic (SL) channel in the previous sections, the parameter estimation is presented. In order to ensure good scientific practice, analyses using data provided by the CMS collaboration perform the parameter estimation with observed data in the following steps. First, the agreement of the simulated samples and the observed data is quantified. Subsequently, the pulls and constraints of the most important nuisance parameters are analyzed. Finally, the results of the parameter estimation of the signal-strength modifier are reviewed. At the time of writing this thesis, the SL analysis is undergoing the final step in this workflow. Consequently, the final value of the signal strength modifier $\mu_{t\bar{t}H}$ is not known yet.

The discussion in this section reflects this situation. First, the aforementioned agreement between simulation and observed data is reviewed before the nuisance parameter estimation in the fit to observed data is discussed. Since the final value of the signal strength modifier is not known yet, the final part of this section discusses the preliminary expected sensitivity of $t\bar{t}H(b\bar{b})$ analysis in the SL channel, which is obtained using pseudo data corresponding to the SM expectation for the signal and background processes. Unless stated otherwise, the discussion is based on the combined SL statistical model, i.e. the simultaneous parameter estimation in the $\geq$ 4j, 3t and $\geq$ 4j, $\geq$ 4t regions.

8.3.1 Distributions and Goodness-of-Fit test

As mentioned at the end of the previous section, a total of 13 categories enter the final fit in the SL channel. The corresponding pre-fit and post-fit yields and distributions can be found in App. H.1. The binning of the corresponding distributions has been carefully optimized. Several of these categories have only one bin and thus only contribute the information of the process yields to the parameter estimation. Specifically, this is the case for the $t\bar{t}$ + lf, $t\bar{t}$ + $c\bar{c}$ and $t\bar{t}$ + 2b categories in both analysis regions, as well as the $t\bar{t}H$ and $t\bar{t}$mb in the $\geq$ 4j, 3t region. The distributions of all other categories employ the binning scheme as explained at the end of Sec. 8.2, which provide more differential information in the parameter estimation. This strategy provides the fit with a reasonable amount of flexibility to mitigate differences between simulation and observed data, while still providing sufficient sensitivity.

The inputs in the SL analysis with the highest sensitivity to the $t\bar{t}H(b\bar{b})$ signal are the distributions of the ratio observable in the $\geq$ 4j, $\geq$ 4t region, which is shown in Fig. 8.12. The average signal-to-background ratio S/B in this category is approximately 1:15. This

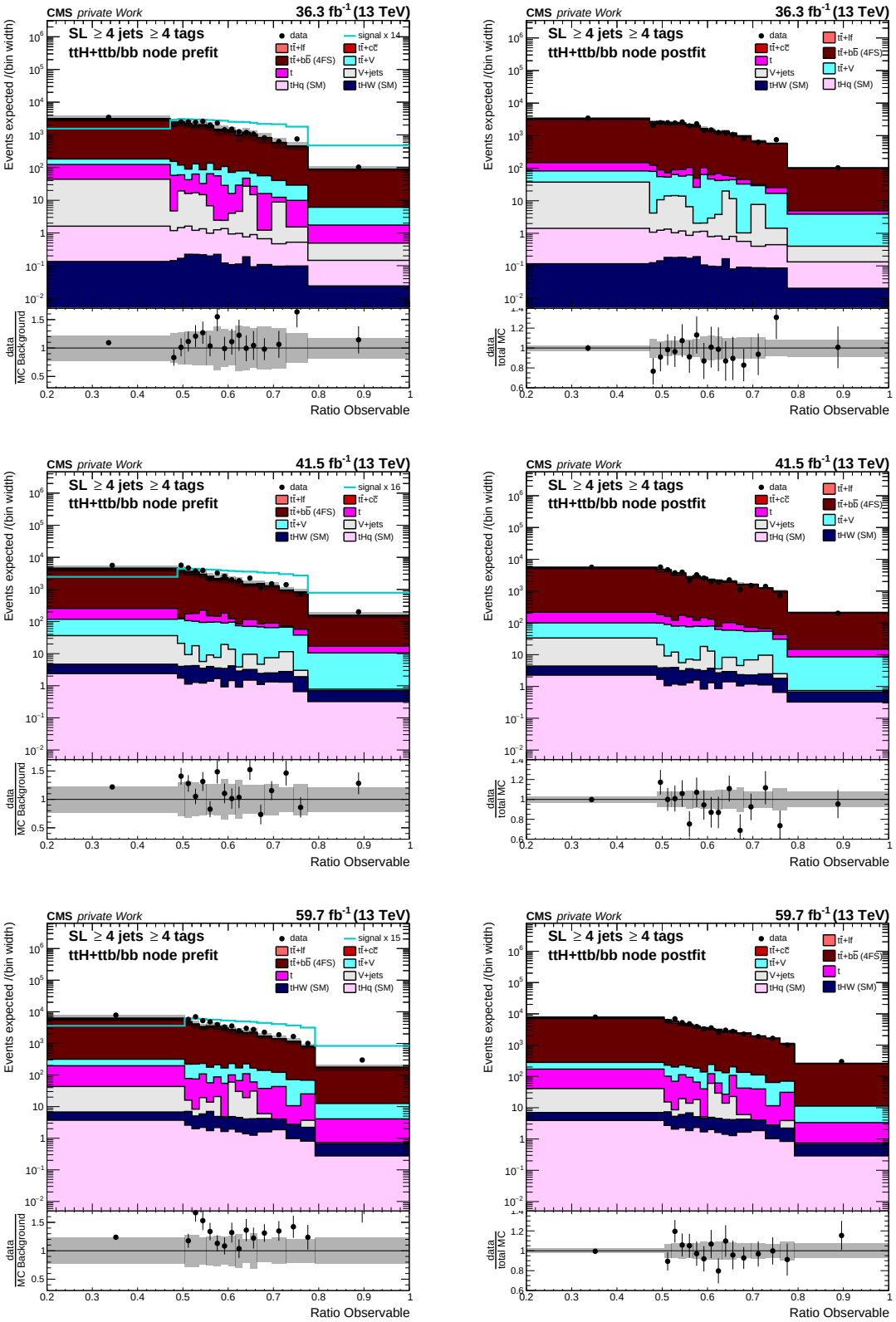


Figure 8.12: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category $t\bar{t}H$ + $t\bar{t}b\bar{b}$/b ratio observable category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

Table 8.2: **Goodness-of-Fit Test for the Final SL Statistical Model**. Each block corresponds to the statistical model for either the $\geq$ 4j, 3t region, the $\geq$ 4j, $\geq$ 4t region or the combination of the both. The values are sorted by decreasing p-values in each block, respectively.

Combination	p value
$\geq$ 4j, $\geq$ 4t (2016)	0.95
Combined (2016)	0.91
$\geq$ 4j, 3t (2016)	0.66
$\geq$ 4j, 3t (2017)	0.99
Combined (2017)	0.95
$\geq$ 4j, $\geq$ 4t (2017)	0.87
$\geq$ 4j, 3t (2018)	1.00
Combined (2018)	0.98
$\geq$ 4j, $\geq$ 4t (2018)	0.92
$\geq$ 4j, $\geq$ 4t (full Run-II)	0.99
Combined (full Run-II)	0.88
$\geq$ 4j, 3t (full Run-II)	0.67

represents a significant improvement in the purity with respect to the whole $\geq$ 4j, $\geq$ 4t phase space mentioned in Sec.8.2, where this ratio is 1:31. Moreover, the shape discrimination between the signal and the background processes is much more pronounced. Both factors illustrate the effectiveness of the neural-network-based classification. Furthermore, the size of the uncertainty bands is reduced in the post-fit distributions, indicating that the fit is sensitive enough to constrain some of the available degrees of freedom. The behavior of the nuisance parameters is analyzed more thoroughly in the subsequent section. Overall, the simulated predictions for the various processes agree well with the observed data points within the uncertainties.

To further validate the description of the observed data in the scope of the statistical model, Goodness-of-Fit (GoF) tests in the separate analysis regions and their combination are performed as described in Sec. 8.2. The results are summarized in Tab. 8.2. All p-values are found to be well above the 5 % threshold, demonstrating the excellent quality of the description of the observed data in the scope of the statistical model.

8.3.2 Nuisance pulls and impacts

To further understand the behavior of the statistical model, the impacts of the individual parameters on the value of the signal-strength inference is calculated. To calculate the impact of a given nuisance parameter, the parameter is set to its $\pm 1\,\sigma$ boundaries. At each point, the signal-strength modifier μ is re-estimated while keeping the nuisance parameter of interest at the current value. The impact is the difference of the thus obtained signal strength modifier and the nominal best estimate of μ.

In order to formulate the expected behavior of the statistical model, the impacts are first calculated using a pseudo data set constructed from the sum of the signal and background predictions. The 20 parameters with the largest expected impact on the signal inference are shown at the top half of Fig. 8.13. Overall, the ranking shows that uncertainties on the $t\bar{t}$ + jets background and $t\bar{t}H$ signal processes are the most important parameters for the estimation of μ. The large impacts of the signal modeling parameters, such as the

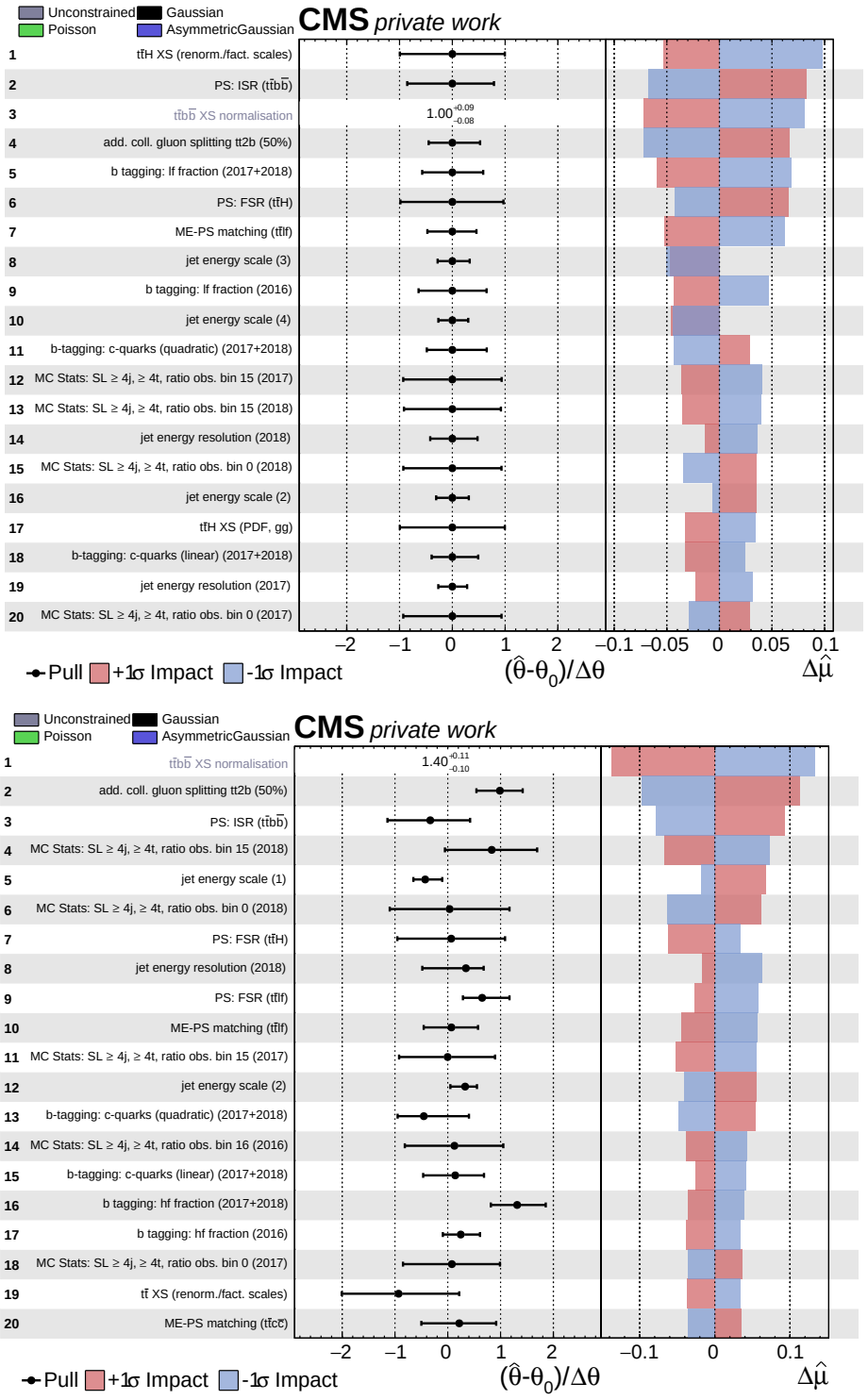


Figure 8.13: **Impacts for the Full Run-II SL Analysis**. Shown are the 20 parameters with the largest impact on the signal inference in the fit to pseudo data containing the signal process expected from the SM (top) and in the fit to the observed data (bottom). The impacts, pulls and constraints are obtained during the parameter estimation with the full Run-II statistical model for the combined SL channel.

uncertainty on the predicted inclusive cross section and the uncertainty on the final-state radiation (FSR), are reasonable since these parameters have a direct influence on the signal process itself. The $t\bar{t}$ + jets processes, and in particular the $t\bar{t} + b\bar{b}$ process, share most of the kinematic properties with the $t\bar{t}H(b\bar{b})$ signal. It is therefore plausible that a modification of the $t\bar{t}$ + jets-related parameters can induce a change in the estimated signal-strength modifier. Consequently, the ranking of the largest expected impacts on the signal-strength inference is reasonable. The individual groups of parameters are discussed in more detail in the following.

The most important theory uncertainties are related to the modeling of the $t\bar{t}$ + jets and the $t\bar{t}H$ processes. The reason for the large impacts of the signal-modeling parameters has already been discussed previously. The background parameters are mainly related to the $t\bar{t} + b\bar{b}$ process, which is expected to yield the final state that is the most similar to the $t\bar{t}H(b\bar{b})$ signal process. It is therefore reasonable that the corresponding uncertainties have large impacts on the signal inference between approximately 5–8 %. Additionally, some of these parameters are constructed with large a priori uncertainties, for example the additional 50 % uncertainty on the collinear gluon splitting. This allows for large variations of the $t\bar{t} + b\bar{b}$ process, which can lead to corresponding changes in the signal process estimation. Therefore, the high position in the impact ranking of these parameters in particular is to be expected.

Among the experimental uncertainties, parameters relating to the b-tagging algorithm, the Jet Energy Scale (JES) and the Jet Energy Resolution (JER) have the highest expected impact on the estimation of the signal strength. The selected SL final state contains at least three b-tagged jets. Therefore, the analysis strategy relies heavily on the b-tagging algorithm. Parameters that impact the measurement of b-tagged objects and thus influence the reconstruction of the different processes entering the analysis are therefore expected to have some impact on the signal inference. Among the b-tagging-related uncertainties, the uncertainty on the purity of the $t\bar{t}$ + lf control region as well as the uncertainties on the charm quark tagging have the highest impact. These parameters also have the largest a priori uncertainty, which can range up to 30 % depending on the category and the processes. These large uncertainties can lead to large variations of the processes. Therefore, it is plausible that these uncertainties also have a large impact on the $t\bar{t}H(b\bar{b})$ measurement. Similar is true for the JES and JER parameters. Moreover, these uncertainties model migration effects of events passing the selections for the analysis phase space. These migration effects can be asymmetric or even one-sided, which leads to the asymmetric impacts that can be seen in Fig. 8.13.

Additionally, uncertainties originating from the limited size of the simulated samples are listed among the parameters with the highest expected impacts on the signal inference. All of these parameters modify bins in the distributions of the ratio observable shown in Fig. 8.12. The bins in question are either the first one in the distribution or are situated at the end of the spectrum. The first bins in the distributions of the ratio observables contain by construction the largest number of simulated events. Therefore, a variation of the yield in these bins is expected to impact the overall normalization of the different processes. This in turn can have a large impact on the estimated value of the signal strength. On the other hand, as shown in Fig. 8.12, the signal process is enriched at the end of the spectrum. The modification of the events in this region therefore has a direct influence on the signal-strength estimation. Therefore, the high impact of these parameters is plausible.

Finally, one of the most important nuisance parameters in the $t\bar{t}H(b\bar{b})$ measurement is the freely-floating background normalization of the $t\bar{t} + b\bar{b}$ process. This parameter directly modifies the yield of the $t\bar{t} + b\bar{b}$ process, the irreducible background in the $t\bar{t}H(b\bar{b})$

measurement. Consequently, the high placement in the impact ranking of this normalization parameter is expected.

The impact ranking in the fit to observed data is shown in the lower half of Fig. 8.13 and is similar to the ranking obtained in the fit to pseudo data. This further validates the compatibility of the simulated predictions for the relevant processes and the observed data. The dominant theory uncertainties modify the $t\bar{t}$ + jets background processes and the $t\bar{t}H$ signal process, which is compatible with the expected impacts. More specifically, the most important theory uncertainties correspond to variations of the irreducible $t\bar{t} + b\bar{b}$ background process. As discussed above, this is expected since any variation of this process has an effect on other processes with similar kinematic properties, such as the signal. Therefore, these uncertainties are expected to have large impacts on the signal strength estimation.

The nuisance parameters are pulled to values of up to $1\,\sigma$, while the a priori values are compatible within an a posteriori confidence interval of approximately $2\,\sigma$. Such pulls are well compatible within the variations expected during statistical inference. Some of these nuisance parameters have a posteriori uncertainties as low as approximately 50 %. As discussed previously, the contribution of the $t\bar{t}$ + jets background processes to the analysis phase space is large. Therefore, the analysis is expected to be sensitive to these processes to some extent. Furthermore, some of these parameters, such as the uncertainty on the collinear gluon splitting, are constructed with a large a priori uncertainty to account for the challenges in the computation of these predictions. Given the expected sensitivity and the large a priori uncertainties, the constraints shown in Fig. 8.13 are reasonable and are well-compatible with the expectation. The post-fit value of the $t\bar{t} + b\bar{b}$ cross section normalization of 1.40 ± 0.10 is well compatible with recent dedicated measurement of the $t\bar{t} + b\bar{b}$ process at the CMS experiment [253, 261].

The JES, JER and b-tagging parameters are the experimental uncertainties with the largest impact on the signal inference. This is well-compatible with the behavior observed when the impacts are calculated with a pseudo data set. As already discussed before, the high impact of these uncertainties is highly plausible given the expected final state of the relevant processes in this analysis. Similar to the theory uncertainties, the nuisance parameters are pulled up to approximately $1\,\sigma$ and are compatible with the expected values within an a posteriori confidence interval of approximately $2\,\sigma$. This is well within the expected fluctuations common in statistical inference.

Finally, the remaining parameters with large impacts on the signal strength estimation are related to the limited size of the simulated samples. Similar to the ranking established using pseudo data, these uncertainties modify the content of bins in the distribution of the ratio observable shown in Fig. 8.12. Moreover, the uncertainties with the largest impacts are either related to the bins with the largest number of simulated events or to the bins at the end of the spectrum, which are expected to have the largest number of signal events. Due to the great importance of these bins in the estimation of the signal strength, the large impact is reasonable.

Overall, the estimation of the nuisance parameters is compatible with the expected behavior. The estimation of the signal-strength modifier is analyzed in the following section.

8.3.3 Parameter of Interest estimation

As already mentioned in the beginning of this section, the following discussion focuses on the preliminary expected sensitivity of the analysis in the SL channel to the $t\bar{t}H(b\bar{b})$ signal. The expected sensitivity is obtained using pseudo data constructed from the sum

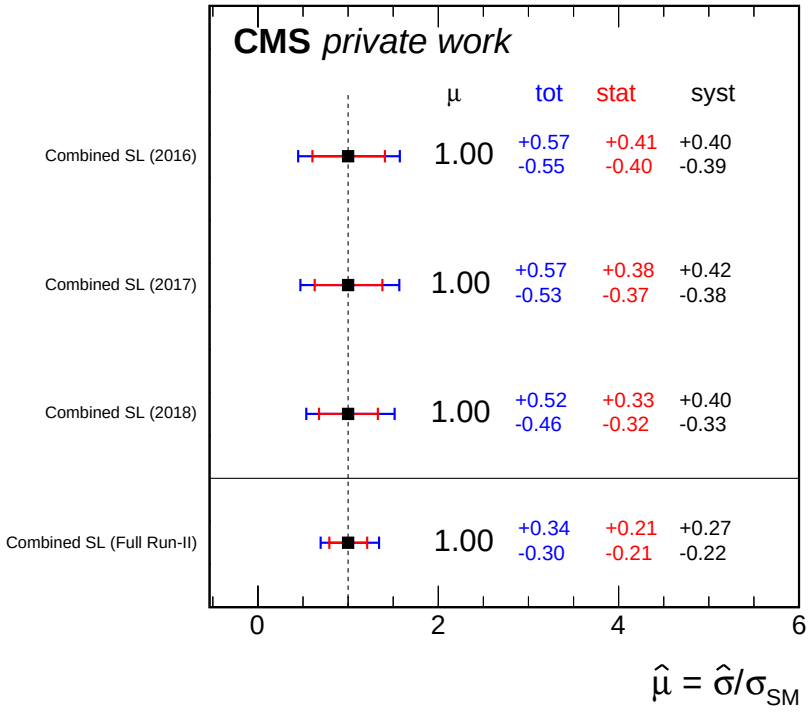


Figure 8.14: **Estimated signal-strength modifiers μ in the SL channel**. Shown are the expected results for the signal-strength estimation in the SL channel for the analyses of the data recorded in the individual years and their combination. The three columns correspond to the total (blue), the statistical (red) and the systematic uncertainty (black) in each measurement.

of the SM predictions for the signal and background processes. Figure 8.14 and Tab. 8.3 summarize the results of the signal-strength estimation for the individual years and the full Run-II combination. The sensitivity to the $t\bar{t}H(b\bar{b})$ signal increases proportional to the integrated luminosity.

In addition to the results obtained in this analysis, Tab. 8.3 shows the results of previous $t\bar{t}H(b\bar{b})$ analyses using the 2016 (HIG-17-026 [133]) and the 2017 (HIG-18-030 [21]) data sets, respectively. First, it is noticeable that the statistical uncertainties in the previous analyses is much smaller than the values quoted for the corresponding measurements in the scope of this thesis. This is due to the definition of the additional cross section uncertainties for the $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ processes. In the previous analyses, these uncertainties were modeled with an a priori uncertainty of 50 %. In this analysis, these normalizations are estimated directly from the observed data. Therefore, they contribute to the statistical uncertainty in the measurement while they were part of the systematic uncertainties previously. Additionally, these parameters are the dominating sources of uncertainty in the measurement. Consequently, the statistical uncertainty in the measurements in the scope of this analysis are much larger than it was in previous iterations. This effect is further discussed in Sec. 9.3

The original measurement of the 2016 data yielded a higher sensitivity to the $t\bar{t}H(b\bar{b})$ process than in this iteration of the analysis. However, the quoted result also considers the dileptonic channel and is thus expected to be more sensitive. Therefore, the measurement described in HIG-17-026 will be compared in more detail to the full Run-II analysis reviewed in Sec. 9.3. The analysis of the data recorded in 2017 in the SL channel yields a similar sensitivity as the measurement for the same year of data taking presented in this thesis.

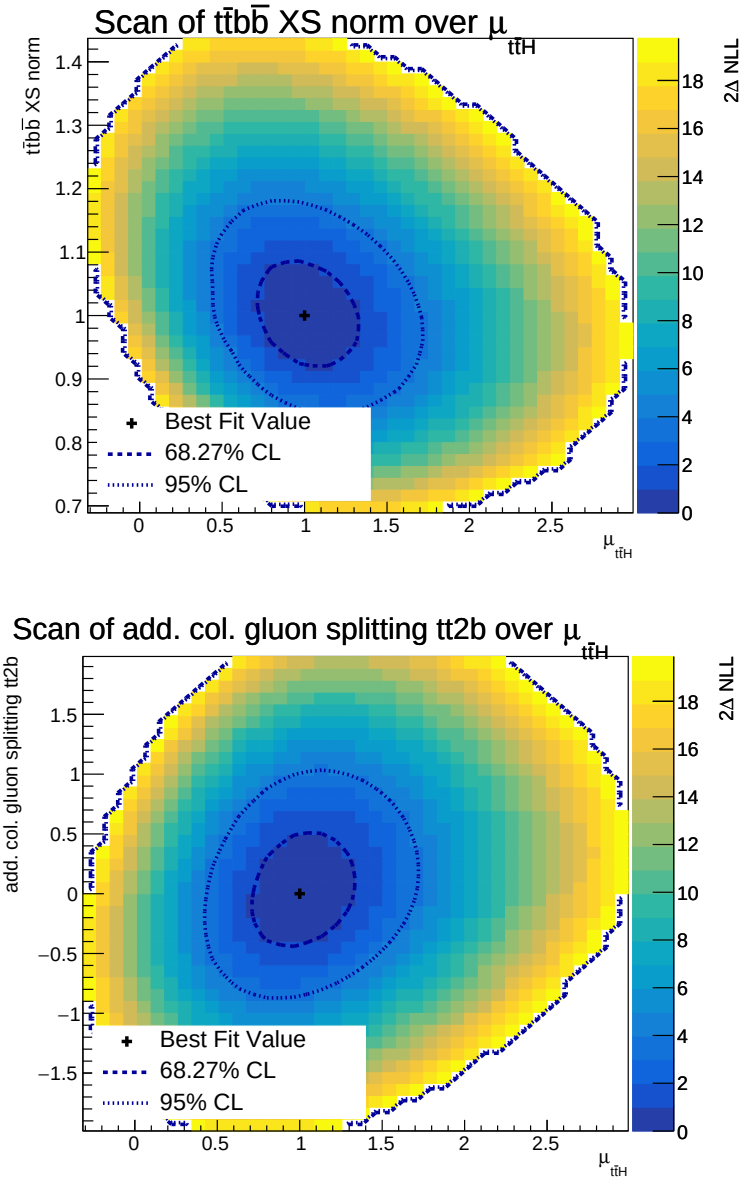


Figure 8.15: **Likelihood scans of $\mu_{\mathrm{t\bar{t}H}}$ and the $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ normalization uncertainties.** Shown are the likelihood scans of the freely-floating $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ normalizations (top) and the additional 50 % uncertainty on collinear gluon splitting (bottom) as a function of the signal-strength modifier $\mu_{\mathrm{t\bar{t}H}}$, respectively. A pseudo data set construction from the sum of the signal and background processes is used for the parameter estimation. Additionally, the best fit value (cross) as well as the $1\,\sigma$ and $2\,\sigma$ confidence intervals (dotted lines) are shown.

Table 8.3: **Results for the estimation of signal-strength modifier in SL Channel**. Shown are the results in the measurements in the individual years and their combination. The two middle columns show the estimated central value for μ with the total uncertainty in the measurement (second column) and the uncertainty due to the statistical inference only (third column). The significance is reported in units of standard deviations. For comparison, the results from previous measurements are shown [21, 133].

	Stat+Syst	Stat-Only	Significance obs	(exp)
HIG-17-026[†]	+0.72 +0.45/-0.45	+0.72 +0.24/-0.24	1.6	(2.2)
Combined SL (2016)	+1.00 +0.57/-0.55	+1.00 +0.41/-0.40		(1.7)
HIG-18-030 (SL channel)	+1.84 +0.62/-0.56	+1.84 +0.26/-0.26	3.3	(1.9)
Combined SL (2017)	+1.00 +0.57/-0.53	+1.00 +0.38/-0.37		(1.9)
Combined SL (2018)	+1.00 +0.52/-0.46	+1.00 +0.33/-0.32		(2.2)
Combined SL (Full Run-II)	+1.00 +0.34/-0.30	+1.00 +0.21/-0.21		(3.3)

[†]: Quoted result obtained in fit of DL and SL channels

The full Run-II combination in the SL channel yields an expected significance of $3.3\,\sigma$ with respect to the background-only hypothesis. Consequently, this channel alone is expected to be sensitive enough to claim evidence of the $t\bar{t}H(b\bar{b})$ process. Table 8.4 shows the five parameters with the largest correlations to the signal-strength modifier. The order of magnitude of the correlations in the fit to the observed data set with respect to the expected correlations is similar. This again indicates the good quality of the description of the observed data by the simulated events used in this analysis. Figure 8.15 shows the likelihood scans of the two parameters with the largest correlations to the signal-strength modifier as a function of $\mu_{t\bar{t}H}$. There is no unexpected behavior visible in the likelihood scans. The correlations shown in Tab. 8.4 are indicated by the tilt in the contours of the confidence intervals.

As expected, the parameters with the largest correlation to $\mu_{t\bar{t}H}$ in the fit to the observed data are related to the $t\bar{t} + b\bar{b}$ process. The anti-correlation between the $t\bar{t} + b\bar{b}$ cross section normalization and the signal-strength modifier is understandable. Since the distributions for the $t\bar{t} + b\bar{b}$ and the signal process are much alike, a change in the background process yield causes an anti-proportional response in the estimation of the signal-strength modifier.

The positive correlation between the signal-strength modifier and the additional 50 % gluon splitting uncertainty and the ISR $t\bar{t} + b\bar{b}$ uncertainty is surprising at first. Since these uncertainties modify the $t\bar{t} + b\bar{b}$ background process, the naive expectation is an anti-correlation to $\mu_{t\bar{t}H}$. To further investigate this, the fits to pseudo data constructed from the signal and background predictions are re-performed. In these fits, the two uncertainties under scrutinization are the only floating parameters apart from the signal-strength modifier, respectively. In these tests, the correlation between the additional gluon splitting uncertainty and $\mu_{t\bar{t}H}$ is $-59.8\,\%$. The correlation between ISR $t\bar{t} + b\bar{b}$ uncertainty and the signal-strength modifier is $-71.3\,\%$. These values match the naive expectation. The positive correlation in the full statistical model must therefore arise from the ensemble of correlations among all nuisance parameters.

Table 8.4: **Largest correlation of the nuisance parameters and the signal-strength modifier $\mu_{t\bar{t}H}$ in the SL channel**. Shown are the five parameters with the largest absolute correlations to the signal-strength modifier in the analysis of the full Run-II data set. For comparison, the correlations are shown in the fit to observed data (Unblinded) and in the fit to pseudo data generated from SM prediction (Blinded).

Parameter	Correlation to $\mu_{t\bar{t}H}$ in %	
	Unblinded	Blinded
$t\bar{t} + b\bar{b}$ XS normalization	−40.1	−24.1
add. col. gluon splitting $t\bar{t} + 2b$ (50 %)	32.7	22.6
PS: ISR ($t\bar{t} + b\bar{b}$)	24.5	23.9
MC Stats: $\geq$ 4j, $\geq$ 4t, ratio obs. bin 15 (2018)	−20.2	−11.5
MC Stats: $\geq$ 4j, $\geq$ 4t, ratio obs. bin 15 (2017)	−15.7	−11.8

Additionally, two uncertainties due to the limited size of the simulated samples have large correlations with the signal-strength modifier. These parameters affect the content of the bins with the highest signal purity in the SL channel. Therefore, the large anti-correlation is reasonable. Whenever the bin content is increased due to the systematic uncertainty, the signal-strength modifier is used to mitigate the ensuing differences between simulation and observed data.

After the discussion of the results in the SL channel, the contribution of this channel to the full Run-II measurement of the full analysis phase space is reviewed in the following chapter.

9 Contribution to the combined $t\bar{t}H(b\bar{b})$ analysis

The analysis in the semileptonic (SL) channel described in Chap. 8 is one of the input channels for the full $t\bar{t}H(b\bar{b})$ analysis described in Ref. [137]. In this analysis, the SL channel is combined with the analyses of the other $t\bar{t}$ decay channels introduced in Chap. 4, specifically the fully-hadronic (FH) and the dileptonic (DL) channels. As already mention in Sec. 8.3, the review of the $t\bar{t}H(b\bar{b})$ analysis is still ongoing at the time of writing this thesis. Therefore, this chapter will briefly review the current preliminary results of the full $t\bar{t}H(b\bar{b})$ combination and compare them to the observations discussed in Chap. 8. A more thorough discussion of the full $t\bar{t}H(b\bar{b})$ analysis can be found in Ref. [137].

9.1 Distributions and Goodness-of-Fit Test

First, the post-fit distributions introduced in Sec. 8.3 are compared to the distributions obtained when fitting all channels together. The comparison is shown in Fig. 9.1. Overall, the post-fit distributions for the different processes are very similar between the SL-only statistical model and the full $t\bar{t}H(b\bar{b})$ combination. The agreement between the observed data and the simulated predictions of the different processes is equally good in both models. This indicates that the parameter estimation in the SL channel already yields values that are very close to the best possible estimation. Moreover, the uncertainty bands are compatible within a few percent. This suggests that the main sensitivity for the systematic uncertainties in the fully-combined $t\bar{t}H(b\bar{b})$ analysis originates from the SL channel. The contribution of the different channels to the sensitivity of the full $t\bar{t}H(b\bar{b})$ analysis is further analyzed in Sec. 9.3.

The agreement between the observed data and the simulation is further quantified with the Goodness-of-Fit (GoF) test suit introduced in Sec. 8.2. The calculated p-values for the different input channels are shown in Tab. 9.1. The values indicate that the observed data is well-modeled in all channels. The smaller value for the full combination stems from tensions between the SL and the FH channels. The analysis of the latter introduces a sophisticated data-driven background estimation in order to control the QCD background, which is the most dominant process in the FH analysis. As part of this technique, additional parameters are introduced, which have been shown to impact the modeling parameters for the $t\bar{t}$ + jets background processes. When combined with the other $t\bar{t}H(b\bar{b})$ channels, an

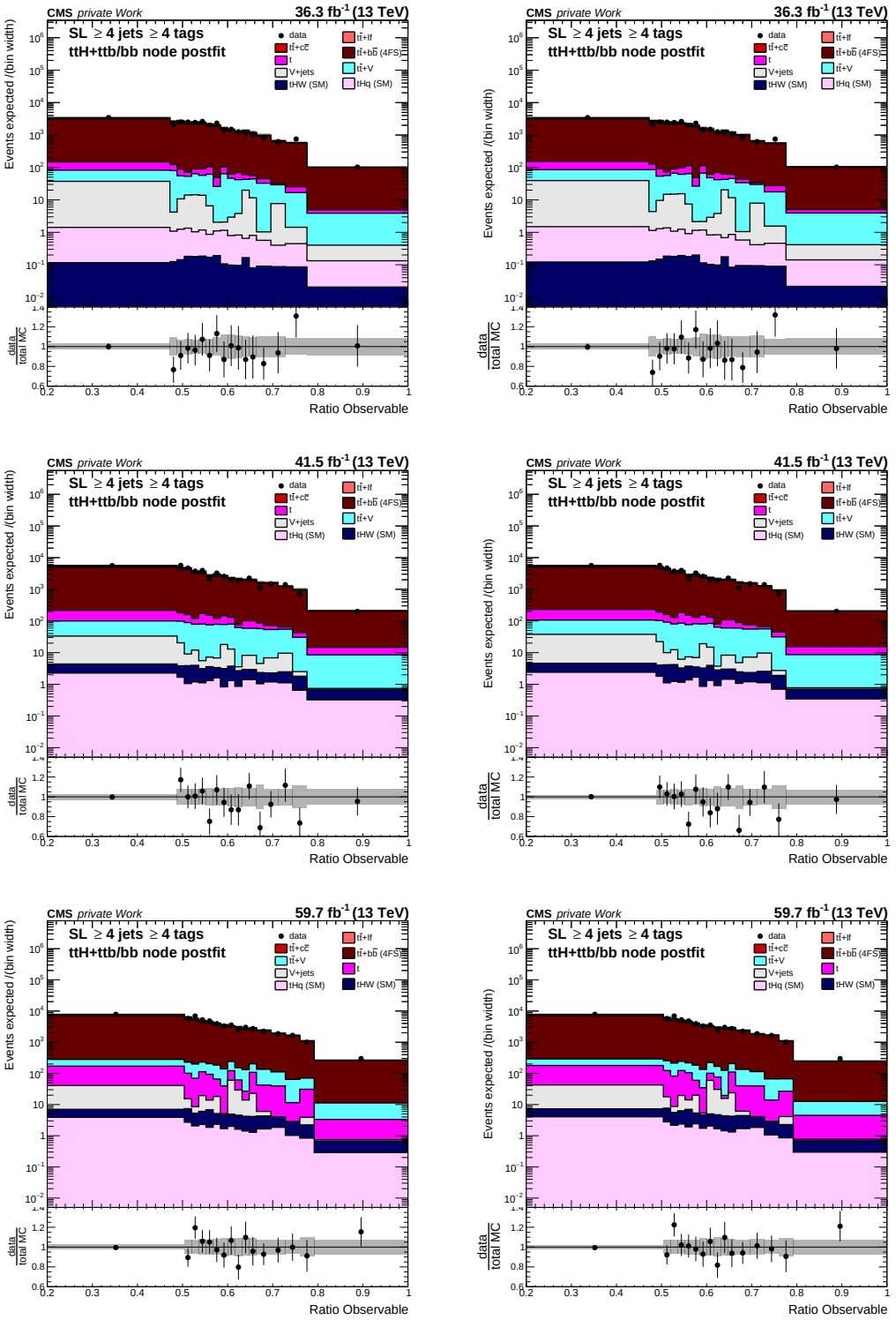


Figure 9.1: **Comparison of the ratio observable discriminant in the SL channel in the $\geq$ 4j, $\geq$ 4t category**. Shown are the post-fit distributions obtained from the analysis in the SL channel (left) and when fitting all analysis channels together (right) for 2016 (top), 2017 (middle), and 2018 (bottom). The parameter estimation is based on the full Run-II data set. The black points represent the observed data. The stacked distributions correspond to the simulated predictions for the relevant processes. The uncertainty band corresponds to the set of uncertainties in the corresponding statistical model. The distributions of the full $t\bar{t}H(b\bar{b})$ analysis are taken from Ref. [137].

Table 9.1: **Goodness-of-Fit test for the full $t\bar{t}H(b\bar{b})$ combination**. Shown are the p-values calculated in the scope of the Goodness-of-Fit test for the individual channels of the $t\bar{t}H(b\bar{b})$ analysis and combinations. Each statistical model considers the full Run-II data set. Adapted from Ref. [137].

	p value
Combined DL	1.00
SL $\geq$ 4j, $\geq$ 4t	0.99
Combined DL + FH	0.97
Combined SL	0.88
Combined DL + SL	0.79
SL $\geq$ 4j, 3t	0.70
Combined FH	0.63
Full $t\bar{t}H(b\bar{b})$ Combination	0.45
Combined SL + FH	0.42

interplay between the channels can be observed, which introduces tensions in the statistical model. These tensions were studied in the scope of the $t\bar{t}H(b\bar{b})$ analysis, where it was shown that the description of the observed data is still valid. This is also indicated by the p-value, which implies compatibility of the statistical model and the observed data within a $1\,\sigma$ confidence interval. Such fluctuations are expected within a statistical data analysis. The observed tensions in the full $t\bar{t}H(b\bar{b})$ combination are further discussed in Sec. 9.2.

Overall, both the distributions and the GoF indicate a good description of the observed data within the different statistical models. In the following chapter, the estimation of the nuisance parameters is analyzed in more detail.

9.2 Nuisance pulls and impacts

As shown in Sec. 8.3.2, a thorough study of the pulls of the nuisance parameters and their impacts on the estimation of the value of the signal strength give more information about the dominating effects that impact the signal strength inference. They are crucial to validate the statistical model and to the interpretation of the result as a whole, which is why they are further discussed in the following. Figure 9.2 compares the pulls and impacts for the 20 nuisance parameters with the largest impacts on the signal-strength estimation in the fit to observed data between the fit in the SL channel and the fully-combined fit across all $t\bar{t}H(b\bar{b})$ channels. The most important parameters in both models are related to the normalization of the $t\bar{t} + b\bar{b}$ background process. The pulls and the constraints are well-compatible with each other. The comparison also shows that the combined $t\bar{t}H(b\bar{b})$ analysis can further restrict the uncertainties on the nuisance parameters compared to capabilities of the estimation in the SL channel. For example, the relative improvement in the uncertainty on the $t\bar{t} + b\bar{b}$ normalization estimation is approximately 10 %. While this shows that the main sensitivity on the most important systematic uncertainties mainly originates from the SL channel, it also clearly illustrates that the estimation benefits from considering the entire available phase space.

This is also demonstrated in the ranking of the nuisance parameters. In the SL channel, it is dominated by experimental and theory uncertainties related to the signal and background processes. The ranking in the full $t\bar{t}H(b\bar{b})$ measurement is mostly dominated by statistical uncertainties, such as parameters that directly estimate the normalization of different background processes and uncertainties due to the limited size of the simulated samples.

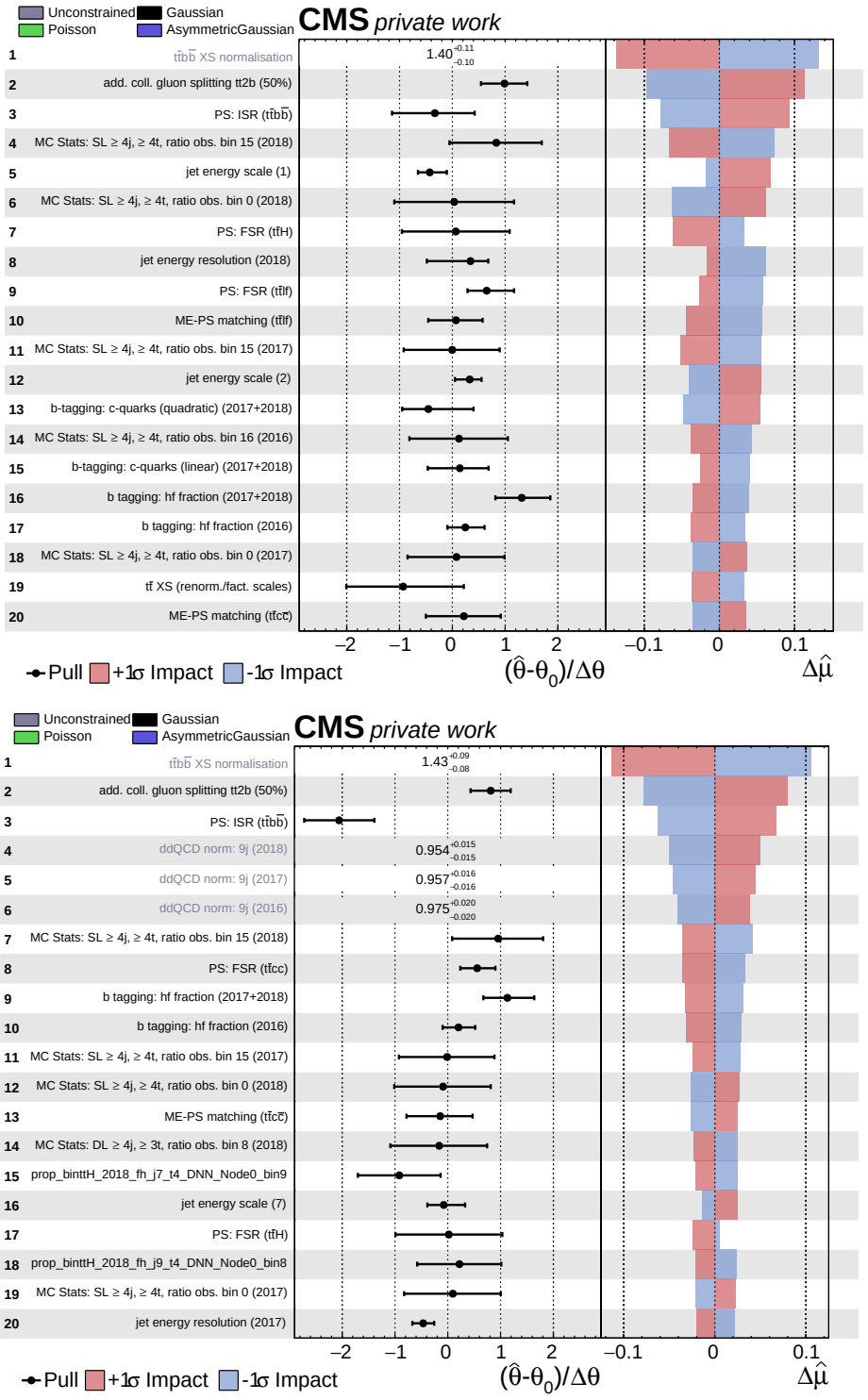


Figure 9.2: **Comparison of the largest impacts for the full Run-II analysis**. Shown are the 20 nuisance parameters with the largest impact on the signal inference in the fit to the observed data. The impacts, pulls and constraints are obtained during the parameter estimation with the full Run-II statistical model for the combined SL channel (top) and the combination of all $t\bar{t}H(b\bar{b})$ channels (bottom). Bottom figure adapted from Ref. [137].

Additionally, the impact of some parameters decreased in the full $t\bar{t}H(b\bar{b})$ combination. One prominent example is the final-state radiation (FSR) uncertainty on the $t\bar{t}H$ process, which is on rank number seven in the SL ranking and has an impact of up to 0.06 on the signal-strength estimation. In the full $t\bar{t}H(b\bar{b})$ measurement, the impact is reduced to 0.03, placing this uncertainty on rank number 16. The reduced impact of the theory and experimental uncertainties points to a better control over these nuisance parameters in the signal-strength inference in the estimation that considers the full available phase space.

Overall, the pulls and constraints on the different nuisance parameters are very similar between the estimation in the SL channel and the full $t\bar{t}H(b\bar{b})$ combination. This indicates that the SL channel has a significant contribution to the overall sensitivity in the full $t\bar{t}H(b\bar{b})$ fit. However, there are also some prominent differences. New parameters that originate in the other $t\bar{t}H(b\bar{b})$ channels are introduced, thus changing the ensemble of parameters and therefore the impact ranking of the parameters. This manifests in two effects. On the one hand, the new parameters can have a large impact on the signal inference in the full $t\bar{t}H(b\bar{b})$ combination, such as the freely-floating parameters that are part of the data-driven background estimation in the FH channel. On the other hand, the estimation of the same parameters can be different between the statistical models in the SL channel and the full $t\bar{t}H(b\bar{b})$ combination. One such difference is the estimation of the initial-state radiation (ISR) uncertainty on the $t\bar{t} + b\bar{b}$ process. The post-fit value in the SL estimation is well compatible with the initial value, while it is pulled in the full $t\bar{t}H(b\bar{b})$ fit. These differences originate from a change in the correlations between the parameters, which is induced by the changes in the statistical model itself. In this particular example, it has been shown that the large pull in the ISR uncertainty originates from an interplay between the data-driven background estimation in the FH channel and the $t\bar{t}$ + jets modeling in the SL channel [262]. This interplay causes the tension that was briefly mentioned in the previous section during the discussion of the Goodness-of-Fit test results. In the scope of the fully-combined $t\bar{t}H(b\bar{b})$ analysis, it has been shown that this large pull does not impair the description of the observed data by the statistical model and has only a small effect on the $t\bar{t} + b\bar{b}$ estimation itself.

Consequently, the estimation of the nuisance parameters gives reasonable results and is compatible with the results obtained in the SL analysis. In the following section, the estimation of the signal-strength modifier $\mu_{t\bar{t}H}$ is analyzed in more detail.

9.3 Inclusive $t\bar{t}H$ signal strength estimation

As already stated in Sec. 8.3, the step-wise procedure to fully understand the statistical inference using observed data is still ongoing at the time of writing this thesis. Therefore, the following discussion focuses on the preliminary expected sensitivity of the full Run-II analysis, which is obtained using pseudo data corresponding to the sum of the SM predictions for the signal and background processes. Figure 9.3 shows the estimation of $\mu_{t\bar{t}H}$ in the different channels of the $t\bar{t}H(b\bar{b})$ analysis. The full combination yields the most sensitive result. This is expected since this statistical model considers the full phase space available and can thus exploit the largest amount of information. Consequently, the contributions of each channel to the sensitivity are able to reduce both the statistical and systematic uncertainties as much as possible.

The figure also gives a clear indication about the sensitivity the different channels contribute to the analysis. The fully-hadronic channel has the highest amount of data available since the branching ratio of the W bosons to quarks is the largest. Therefore, the statistical uncertainties are small compared to the other channels. The high statistical precision of this channel allows for the aforementioned data-driven background estimation technique,

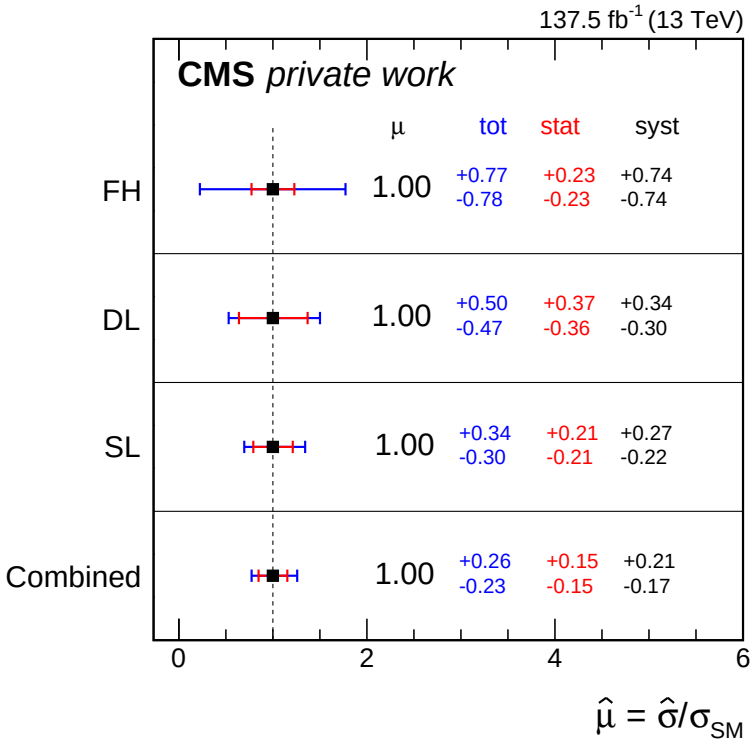


Figure 9.3: **Signal-strength modifier estimation in the different $t\bar{t}H(b\bar{b})$ channels.** Shown are the (blinded) results for the signal-strength estimation in the $t\bar{t}H(b\bar{b})$ analysis. The three columns correspond to the total (blue), the statistical (red) and the systematic uncertainty (black) in each measurement, respectively. Figure taken from Ref. [137]

which infers more information directly from the data without prior constraint terms as introduced in Chap 6. As a consequence, the statistical precision degrades and is slightly worse with respect to the precision seen in the SL channel. Since the collisions at the LHC are performed with protons, there are many hadronic objects in the events. Therefore, it is very difficult to reconstruct the $t\bar{t}H(b\bar{b})$ system in the FH channel. Hence, the systematic uncertainty in the FH channel is largest.

The reconstruction efficiency of the DL channel is very high due to the two leptons in the expected final state. Consequently, the systematic uncertainty is much smaller compared to the uncertainty in the FH channel. However, the production of such events has the lowest probability among the three $t\bar{t}H(b\bar{b})$ channels. Therefore, the statistical uncertainty is the largest among the channels, which is the limiting factor for the contribution to the total sensitivity of the DL channel.

The SL channel offers a compromise between the reconstruction efficiency and the production probability. Therefore, it yields the smallest uncertainties and contributes the most to the overall sensitivity. Consequently, it is the driving factor in the $t\bar{t}H(b\bar{b})$ measurement and crucial for the parameter estimation.

Table 9.2 shows the comparison of the previous $t\bar{t}H(b\bar{b})$ measurements performed with the 2016 and 2017 data sets and the latest full Run-II measurement that additionally considers the 2018 data set. The statistical uncertainties in the previous iterations are smaller than in the corresponding measurements in the scope of the full Run-II analysis. This is due to the parameterization of the normalization of the $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ processes. In the previous iterations of the $t\bar{t}H(b\bar{b})$ measurement, the normalization of these processes was modeled

Table 9.2: **Comparison of the signal-strength estimation to previous measurements in the $t\bar{t}H(b\bar{b})$ channel**. Shown are the results from the three iterations of the $t\bar{t}H(b\bar{b})$ analysis HIG-17-026 [133], HIG-18-030 [21] and HIG-19-011 [137] on the same data sets. The values are sorted according to the integrated luminosity and the input channels considered in the respective analysis. The two middle columns display the central estimations for the signal-strength modifier as well as the total uncertainty (second column) and the statistical contribution to the uncertainty (third column). The last column displays the observed (expected) significance of the measurement in units of standard deviations.

Parameter	Stat+Syst	Stat-Only	Significance	
HIG-17-026	+0.72 +0.45/-0.45	+0.72 +0.24/-0.24	1.6	(2.2)
HIG-19-011 DL+SL, 2016	+1.00 +0.48/-0.45	+1.00 +0.36/-0.35		(2.2)
HIG-18-030	+1.15 +0.32/-0.29	+1.15 +0.15/-0.15	3.9	(3.5)
HIG-19-011 Full Combination, 2016 + 2017	+1.00 +0.32/-0.29	+1.00 +0.20/-0.20		(3.5)
HIG-19-011 Full Combination, Full Run-II	+1.00 +0.27/-0.23	+1.00 +0.15/-0.15		(4.4)

with an a priori uncertainty of 50 %. In the full Run-II analysis, the normalizations are estimated in situ directly from the data. Therefore, the normalization parameters contribute to the statistical uncertainty in the full Run-II measurement. Since these parameters are among the leading uncertainties in the $t\bar{t}H(b\bar{b})$ measurement, the statistical uncertainty is expected to increase significantly. The statistical component of the different measurements is further decomposed in Tab. 9.3. The contribution to the uncertainty that originates from the analyzed amount of data and the pure separation power of the discriminators is smaller than it was in previous iterations of the $t\bar{t}H(b\bar{b})$ measurements. This is due to refinements in the analysis strategy and the construction of the final observables, which increased the statistical sensitivity of the full Run-II analysis.

Nevertheless, the overall sensitivity of the full Run-II analysis is comparable with the sensitivity fo the previous iterations of the $t\bar{t}H(b\bar{b})$ measurement. As shown in Ref. [137], this is mainly due to the new $t\bar{t} + b\bar{b}$ simulation used for the full Run-II analysis. While this simulation is expected to provide an improved description of the kinematic properties of this background process, it also introduces larger uncertainties to the statistical model. These effects are mitigated by the improvements in the analysis strategy, which allows for a better control over the most relevant background process without loosing sensitivity.

Table 9.3: **Comparison of the statistical uncertainties in the signal-strength estimation**. Shown are the contributions to the statistical uncertainty in the different iterations of the $t\bar{t}H(b\bar{b})$ measurement. The values in the last column are obtained by subtracting the contribution of the normalization parameters from the total statistical uncertainty in quadrature.

	Total Stat	Normalization	Other
HIG-17-026	0.24	–	0.24
HIG-19-011 DL+SL, 2016	0.35	0.27	0.22
HIG-18-030	0.15	–	0.15
HIG-19-011 Full Combination, 2016 + 2017	0.20	0.15	0.13
HIG-19-011 Full Combination, Full Run-II	0.15	0.12	0.10

Part III

Further Interpretation and Prospects

10 Towards Differential Measurements in $t\bar{t}H(b\bar{b})$

After the description of the inclusive $t\bar{t}H(b\bar{b})$ analysis, a more involved measurement that is possible with the available full Run-II statistics is reviewed in this chapter. The analysis presented in Ref. [137] includes the first interpretation towards a differential measurement in this channel within the Simplified Template Cross Section (STXS) framework. This approach enables the direct comparison of cross sections in given bins of observables, as is further discussed in Sec. 10.1, and allows for a direct comparison of predictions from other models beyond the Standard Model (BSM). It is therefore a powerful tool that can be used to formulate and test new theories.

This chapter discusses the STXS approach in the semileptonic (SL) $t\bar{t}H(b\bar{b})$ channel. First, the general concept of the STXS measurements is introduced. Afterwards, the strategy used in the SL channel is discussed in more detail. As previously stated in Chaps. 8 and 9, the $t\bar{t}H(b\bar{b})$ analysis is still under final review at the time of writing this thesis. Therefore, the final part of this chapter reviews the preliminary expected results and compares this expected sensitivity with predictions obtained from an extension of the SM.

10.1 Introduction to Simplified Template Cross Sections

This section will give a short introduction to the STXS interpretation in $t\bar{t}H$-related analyses. First, the general concept of the technique is briefly introduced. Afterwards, the definition that is used in the $t\bar{t}H(b\bar{b})$ analysis is shown. For a more complete overview of STXS measurements and the studies required to formulate the final strategy, please refer to Refs. [19, 29, 263].

The core concept of the STXS interpretation is to perform cross section measurements in regions of certain observables, which are referred to as STXS bins in the following. The definitions of these STXS bins are synchronized between the experimental measurement and the theoretical calculation. This and the direct measurement of cross sections that are independent of uncertainties related to theory calculations facilitate a direct comparison of the observed values and predictions from different theories, which is illustrated in Sec. 10.4.

The implementation of the STXS interpretation is very similar to a likelihood-based unfolding procedure. First, a property of interest is defined based on the model that is to

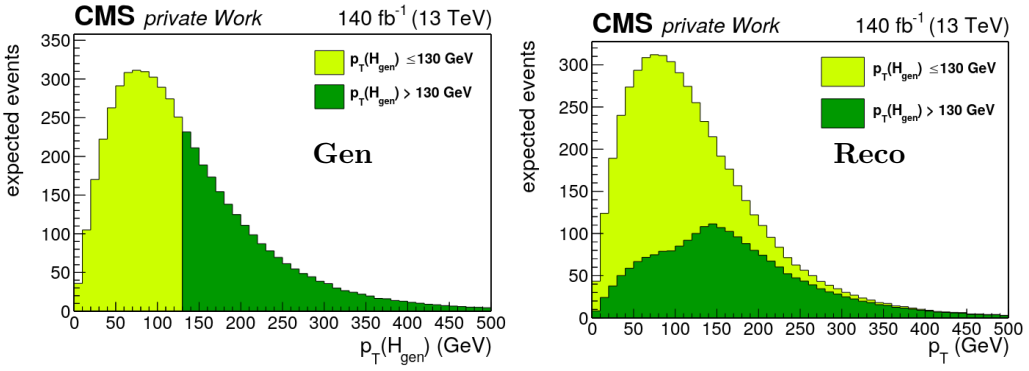


Figure 10.1: **Exemplary Definition of STXS Bins**. The figure illustrates the definition of two exemplary regions of interest (bins) at truth-level (left) and after the reconstruction (right). Figures taken from Ref. [263].

be tested. As already shown in Sec. 2.2, the transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$ of the Higgs boson is expected to be sensitive to contributions arising from operators beyond the SM. Therefore, it is the observable of interest in the context of the $t\bar{t}H$ production channel. The following explanation is structured accordingly.

Figure 10.1 illustrates a simple example of such a definition. Here, the predicted spectrum of $p_{\mathrm{T}}^{\mathrm{H}}$ is split into two STXS bins, one in the low-p_{T} and one in the high-p_{T} region. The definition of the STXS bins is based on quantities of physics processes prior to the detector simulation, which is the point that is closest to the theoretical calculation in the simulation chain. This point is commonly referred to as the truth-level definition.

As shown in Fig. 10.1, the well-defined separation of the two regions at truth level disappears after the detector simulation and they overlap. This behavior arises due to detector acceptance effects and inefficiencies. The inversion of these smearing effects is an ill-posed mathematical problem, thus rendering the direct extraction of the truth-level information impossible. The STXS approach aims to revert these effects by propagating the truth-level definition to the reconstruction after the detector simulation and finally measuring the contribution of each STXS bin simultaneously. In essence, this enables the inference of the truth-level properties while accounting for detector effects.

The STXS measurement in $t\bar{t}H$ analyses follows the same principles. The bins were defined in a common effort from experts in different Higgs-boson-related topics in both experiment and theory in the scope of the LHC Higgs Cross Section Working Group [29]. The definition for the $t\bar{t}H$ channel is a further development of this scheme. It is designed to be sensitive to current BSM theory predictions for the $t\bar{t}H$ process in a common framework with other Higgs-boson analyses. This facilitates easy combination of different measurements with common definitions in the experimental strategy and the theoretical predictions, which allows to extract the maximum amount of information.

The bins are defined in two stages. First, the region of interest is limited to Higgs bosons with a rapidity $|y^{\mathrm{H}}| < 2.5$. Afterwards, the $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum is divided into five STXS bins, which are shown in Fig. 10.2. These STXS bins are the basis for the strategy in the SL channel.

10.2 STXS Strategy in $t\bar{t}H(b\bar{b})$

Detectors in particle physics are only able to detect stable particles that are able to traverse to the sensitive regions used for the measurement. Consequently, information about highly

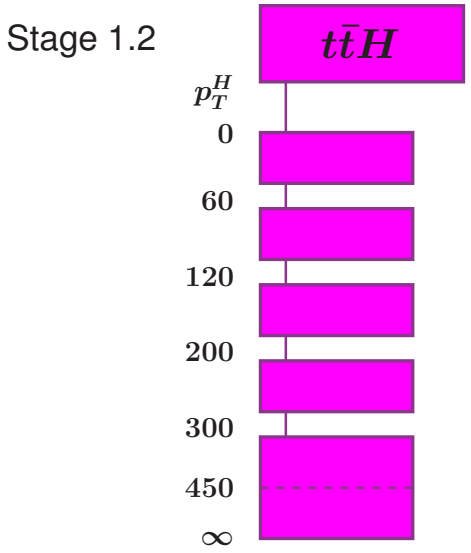


Figure 10.2: **Definition of STXS Bins in $t\bar{t}H$-related Analyses**. The figure shows the five STXS bins used in $t\bar{t}H$-related measurements. Additionally to the categorization shown here, the rapidity of the Higgs boson is restricted to $|y^{\mathrm{H}}| < 2.5$. Figure taken from Ref. [30].

unstable, heavy particles like the Higgs boson or top quark, can only be inferred from their decay products. The correct reconstruction of these heavy particles based on the measured objects in the final state is one of the main challenges for experimental physicists. Therefore, the correct reconstruction of the transverse momentum of the Higgs boson $p_{\mathrm{T}}^{\mathrm{H}}$ is one of the most important aspects for the definition of the analysis strategy for the STXS interpretation.

In the SL channel, events are assigned to one of the STXS bins based on the output of dedicated Artificial Neural Networks (ANNs), which is schematically shown in Fig. 10.3. The reconstruction is performed in two stages. First, the event classification used in the inclusive measurement and described in Sec. 8.2 is performed. In each analysis region, only events that are classified as one of the $t\bar{t}H$-related classes are considered for the STXS classification. In particular, this means that events in the $\geq$ 4j, $\geq$ 4t region introduced in Sec. 8.2 which are classified as $t\bar{t}H$ or $t\bar{t}b\bar{b}$/b, and events in the $\geq$ 4j, 3t region classified as $t\bar{t}H$ are used. These events are then further classified by additional ANNs (STXS ANNs) in each analysis region.

Much like the ANNs used for the classification of the process class, the STXS ANNs consider different properties of the individual objects and the event as a whole as input features. Additionally, the outputs of multivariate analysis techniques are used, such as the aforementioned Matrix Element Method (MEM) [257] and the event reconstruction based on the jet assignment boosted decision trees (JABDT) [137]. The full list of input features can be found in Tab. D.10. The quality of the description of these observables in observed data is evaluated with the same method as described in Sec. 8.2. In particular, the Goodness-of-Fit (GoF) test is performed for each observable and for the correlation between all pairs of input features for each analysis region. The results of these tests are summarized in App. F.1. Only observables with a p-value of above 5 % are considered for the training of the STXS ANNs.

Both STXS ANNs aim to classify the events according to the aforementioned STXS bins and assign a probability-like number for each of the classes for each event. Figure 10.4

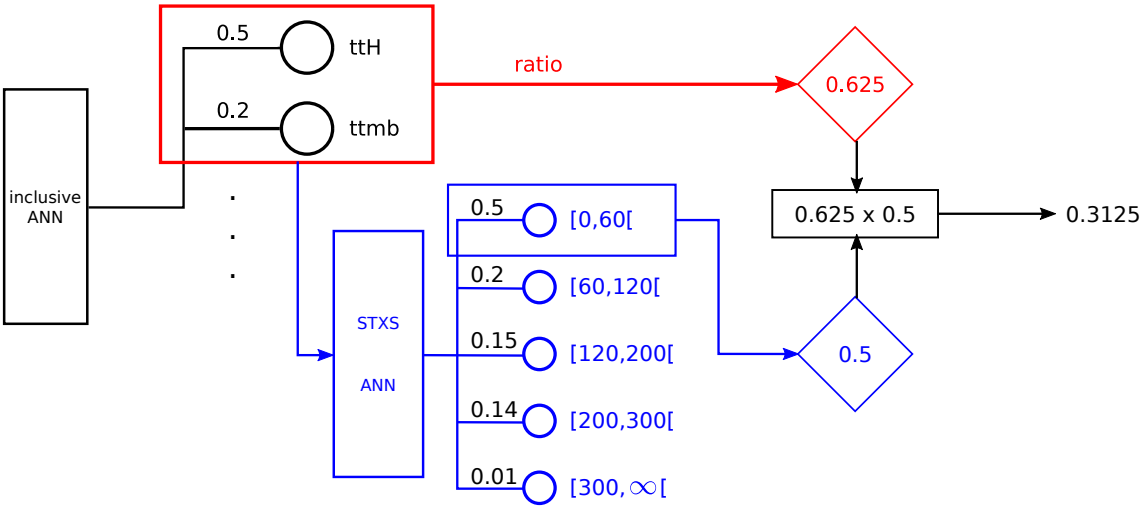


Figure 10.3: **Classification for STXS Interpretation in the $\geq$ 4j, $\geq$ 4t Regions.** Events are first assigned to a process class with the ANN used for the inclusive measurement (red). Signal-like events are then further categorized according to the reconstructed transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$ (blue). The final observable is constructed the product of the output values of both ANNs.

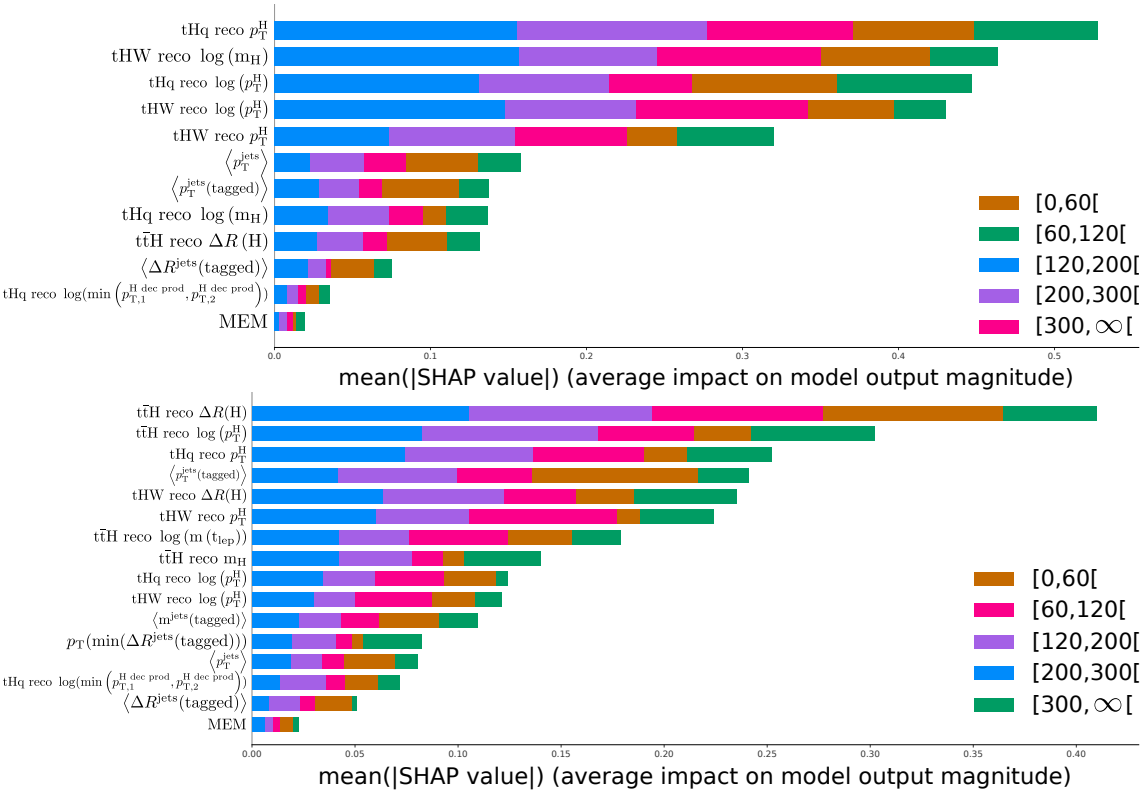


Figure 10.4: **Ranking of Input Features in STXS ANNs.** Shown are the rankings based on the Shapley metric for the ANNs in the $\geq$ 4j, 3t (top) and the $\geq$ 4j, $\geq$ 4t regions (bottom), respectively.

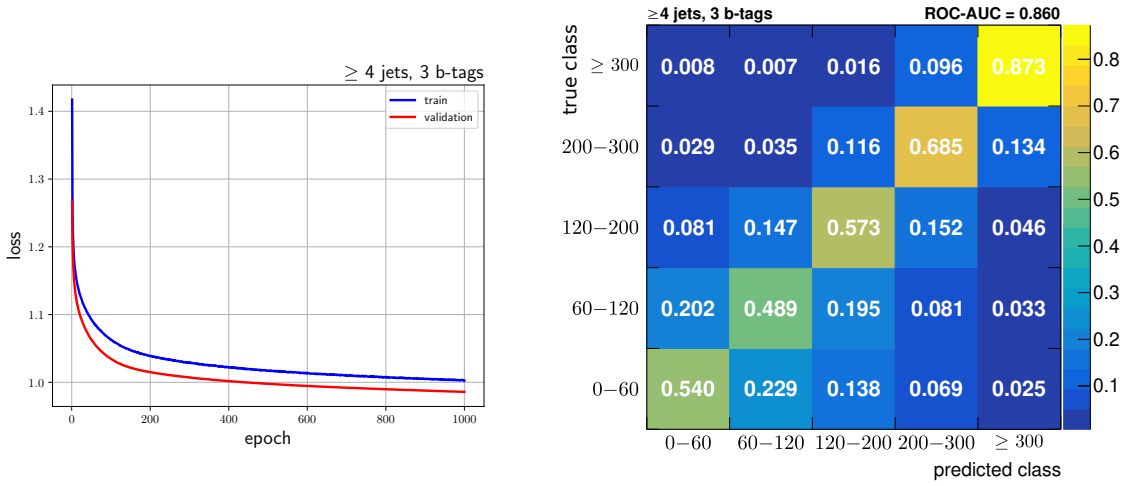


Figure 10.5: **Training Metrics of the STXS ANN in the ≥ 4j, 3t Region**. Shown are the loss function as a function of training epochs (left) and the confusion matrix of the final ANN (right).

displays the ranking of the observables used for training based on the Shapley metric [259] as introduced in Sec. 8.2 for the STXS ANNs. In both cases, the transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$ obtained from the JABDTs related to the reconstruction of Higgs-boson-related topologies are among the most important variables. Since the task of the STXS ANNs is the classification according to $p_{\mathrm{T}}^{\mathrm{H}}$, this ranking is reasonable. The prior reconstruction of the events under the tHq, tHW and $t\bar{t}H$ hypotheses combines information of multiple observables, which is expected to yield good estimates for $p_{\mathrm{T}}^{\mathrm{H}}$. A high ranking of these estimates is therefore to be expected. The other observables used for the training of the ANNs are related to other properties of the event, such as the distance between two (b-tagged) jets, the reconstructed mass of the Higgs boson and the sum of the transverse momenta of all jets. These properties yield further information about the event and especially about the reconstructed Higgs boson candidate. Consequently, they are expected to be important in the classification of the correct STXS bin, which makes the ranking of the input features plausible.

Figures 10.5 and 10.6 show the training metrics for the STXS ANNs in the two SL analysis regions, respectively. In both cases, no overtraining is observed. The reconstruction efficiency implied by the confusion matrices is equally well in both categories. Both networks indicate a large amount of confusion in the STXS bins below $p_{\mathrm{T}}^{\mathrm{H}} \leq 200\,\mathrm{GeV}$ as well as between neighboring classes. In light of the list of input features, this behavior is plausible. The observables used as inputs for the STXS ANNs are expected to be similar for similar, i.e. neighboring, $p_{\mathrm{T}}^{\mathrm{H}}$ regions, and become more distinctive for regions that are further apart. For example, the distance between two b-tagged jets is expected to decrease with increasing $p_{\mathrm{T}}^{\mathrm{H}}$ since the Higgs boson decay products become more collimated. Another good example to understand the confusion between neighboring STXS bins is the average transverse momentum of the b-tagged jets. Since the Higgs boson decays into bottom quarks, the emerging jets are naturally expected to have a larger p_{T} on average with increasing $p_{\mathrm{T}}^{\mathrm{H}}$. Consequently, the values of these observables are expected to be more distinct between the very low- and very high-$p_{\mathrm{T}}^{\mathrm{H}}$, whereas the difference between the neighboring STXS classes are less pronounced. Therefore, the observed classification power is reasonable.

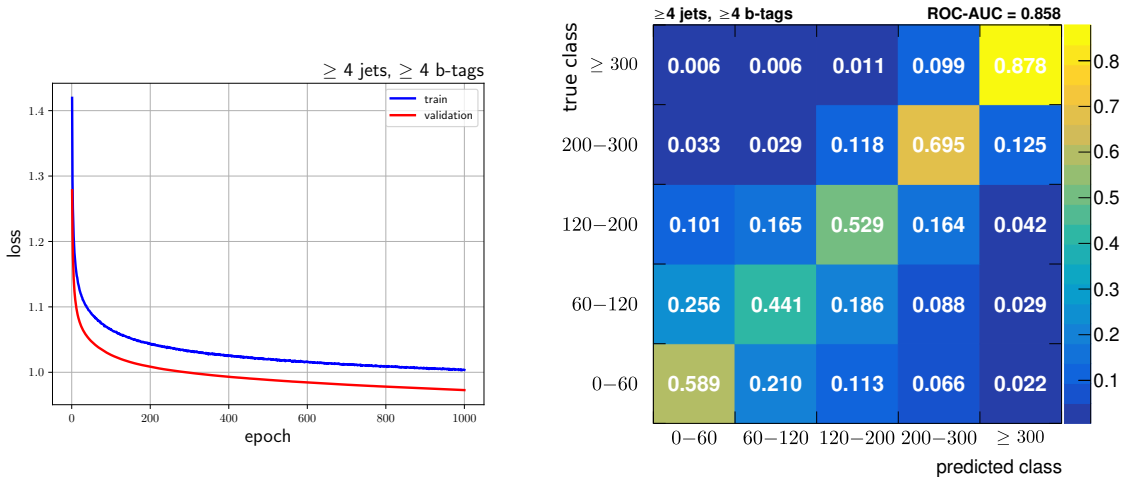


Figure 10.6: **Training Metrics of the STXS ANN in the ≥ 4j, ≥ 4t Region**. Shown are the loss function as a function of training epochs (left) and the confusion matrix of the final ANN (right).

The final observable entering the parameter estimation is the distribution of the STXS ANN output values weighted by the corresponding network used for the initial process classification. Consequently, this metric contains information about the quality of both the initial process classification and the final assignment to a given STXS bin. The treatment of the output values in the background-related nodes is the same as in the inclusive analysis. This yields powerful discriminators suitable for the simultaneous cross section measurement of the $p_{\mathrm{T}}^{\mathrm{H}}$ regions with good control over the background processes contributing to the analysis phase space. These categories are the basis for the final evaluation, which is discussed in the subsequent section.

10.3 Results

This chapter summarizes the results of the final parameter estimation of the STXS measurement. As mentioned at the end of the previous section, the statistical inference considers the five STXS-related ANN categories and the background categories. Consequently, the inference for the cross section measurement evaluates eleven categories in the ≥ 4j, 3t and ten categories in the ≥ 4j, ≥ 4t region. In principle, the analysis strategy is the same as the strategy for the inclusive measurement, which was described in Sec. 8.1. The expected yield λ_i in histogram bin i introduced in Sec. 6 is modified to enable the simultaneous measurement of the different STXS bins:

$$\lambda_i = \sum_j \sigma_j \cdot s_{i,j}^{\mathrm{t\bar{t}H}}(\vec{\theta}) + b_i(\vec{\theta}), \tag{10.1}$$

where the signal contribution in bin i for each STXS bin j is multiplied with a corresponding parameter σ_j. This new definition of the expected process yields is used in the construction of the test statistic $\tilde{q}_\mu$. The construction itself as well as the underlying principles of the parameter estimation are the same as discussed previously.

The measurements presented in Secs. 8.3 and 9.3 interpret the results in terms of the signal strength modifier μ. While this method is advantageous for the direct comparison of parameters, it is likely to suffer from theoretical calculations. Consequently, whenever there is an update of the calculation of the $\mathrm{t\bar{t}H(b\bar{b})}$ process and the corresponding uncertainties,

the measurement of the signal strength modifier has to be repeated. Additionally, two measurements are not comparable if they do not employ the same underlying theory calculations.

Therefore, the STXS measurement presented in this thesis is performed as a direct measurement of the cross section. To this end, uncertainties related to the theoretical calculation of the $t\bar{t}H$ process are removed from the statistical model, specifically

- the previously-introduced rate and shape parameters modeling effects arising from the renormalization and factorization scales as well as the PDF set and
- the shape uncertainties modeling the effects arising from additional initial (ISR) and final state radiation (FSR).

Additionally, the measures to improve the stability of the statistical model introduced in Sec. 8.1 are also applied to the STXS classifiers.

This measurement of the cross section enables a direct comparison of the predictions from different theory calculations, which is illustrated in Sec. 10.4. The following sections first review the details of the cross section measurement in the scope of the STXS framework. In order to illustrate the different challenges in this measurement, the discussion focuses on the first STXS category with $p_{\mathrm{T}}^{\mathrm{H}} \in [0, 60[\,\mathrm{GeV}$ and the last STXS bin with $p_{\mathrm{T}}^{\mathrm{H}} \geq 300\,\mathrm{GeV}$.

At the time of writing this thesis, the process of analyzing the observed data is still ongoing. The results presented in this chapter reflect the current status and the discussion follows the same structure used in Chap. 8. Specifically, the comparison of observed data and the simulated samples except for the signal processes and the estimation of the nuisance parameters in the fit to observed data are discussed. In these comparisons, the parameter estimation is performed with the full statistical model, i.e. considering the signal contributions. The estimation of the cross sections in the different STXS bins is still under scrutiny. Consequently, the results expected in the scope of the SM are discussed in the latter part of this section.

10.3.1 Distributions and Goodness-of-Fit

The complete overview of the distributions used in the cross section measurement can be found in App. H.2. Figure 10.7 shows the final discriminant for the first STXS bin in the $\geq 4\mathrm{j}$, $\geq 4\mathrm{t}$ region. Since the STXS classification is performed after the initial process assignment by the neural network described in Sec. 8.2, the separation between the STXS bins and the background processes is on a similar level as in the inclusive $t\bar{t}H(b\bar{b})$ analysis. Moreover, the pre-fit distributions in particular show the separation between the different STXS classes. The distribution for the first STXS bin is enriched at large discriminant values. This is expected since this metric is specifically designed to be sensitive to this particular STXS bin. The discrimination to the other STXS classes increases gradually with rising $p_{\mathrm{T}}^{\mathrm{H}}$, i.e. the distribution for the second STXS bin is also enriched at large values, whereas the distribution for the last STXS bin peaks at smaller values of the discriminant. This behavior is also reflected in the number of events categorized into this ANN output. Events that belong to the first STXS bin on truth level have the largest contribution to this observable, whereas events originating from the last STXS category have the smallest yield. This is very similar to the observations made in the discussion of the STXS ANN confusion matrices in Sec. 10.2. The separation power between neighboring STXS bins is small and increases with large differences in $p_{\mathrm{T}}^{\mathrm{H}}$. Therefore, the distributions are fully in line with the expected characteristics.

Figure 10.8 displays the distributions for the category with the highest sensitivity to the last STXS bin. In principle, the features discussed for the discriminant for the first STXS

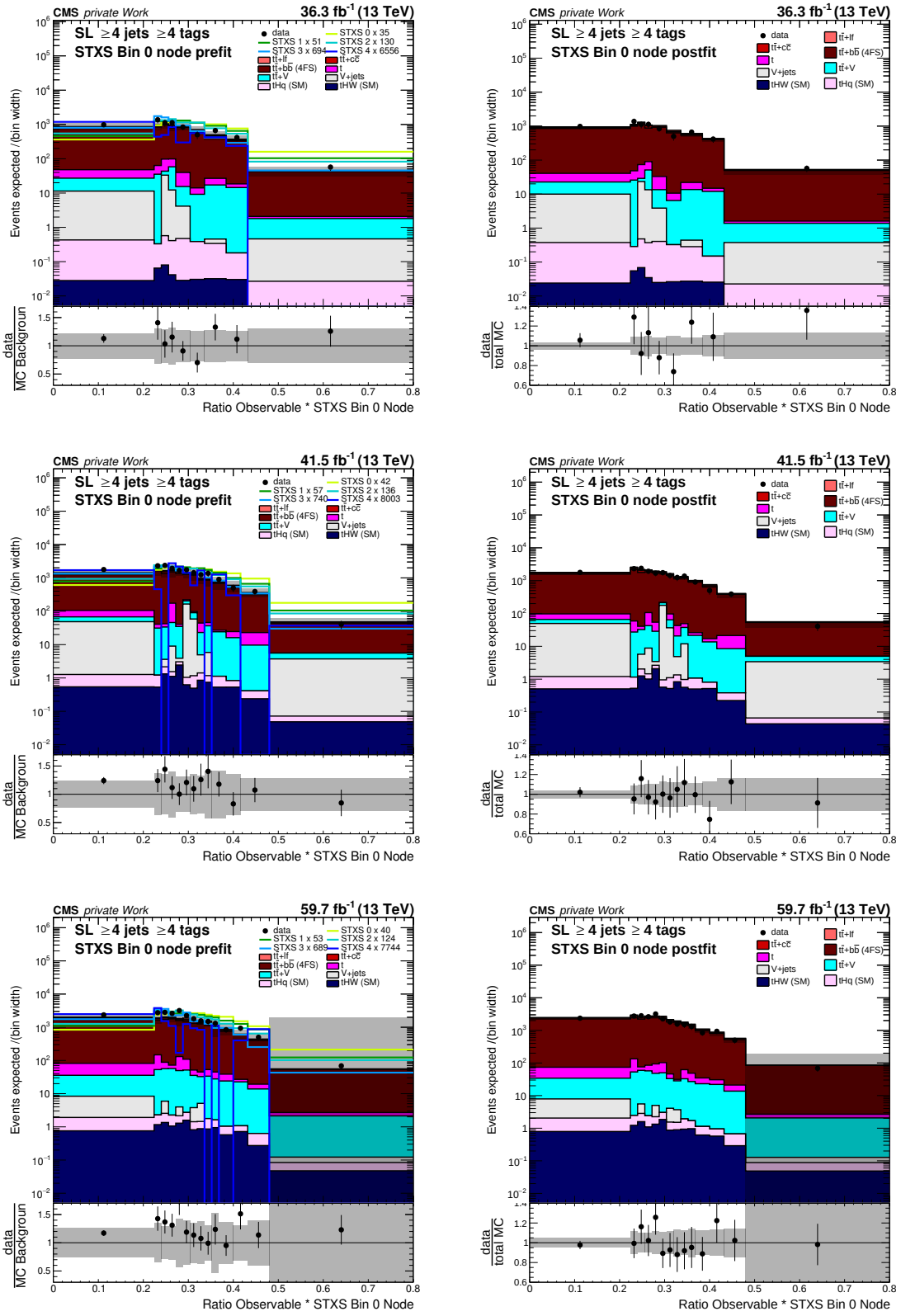


Figure 10.7: **Discriminants in the SL channel in the ratio observable times STXS ANN output for the first STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

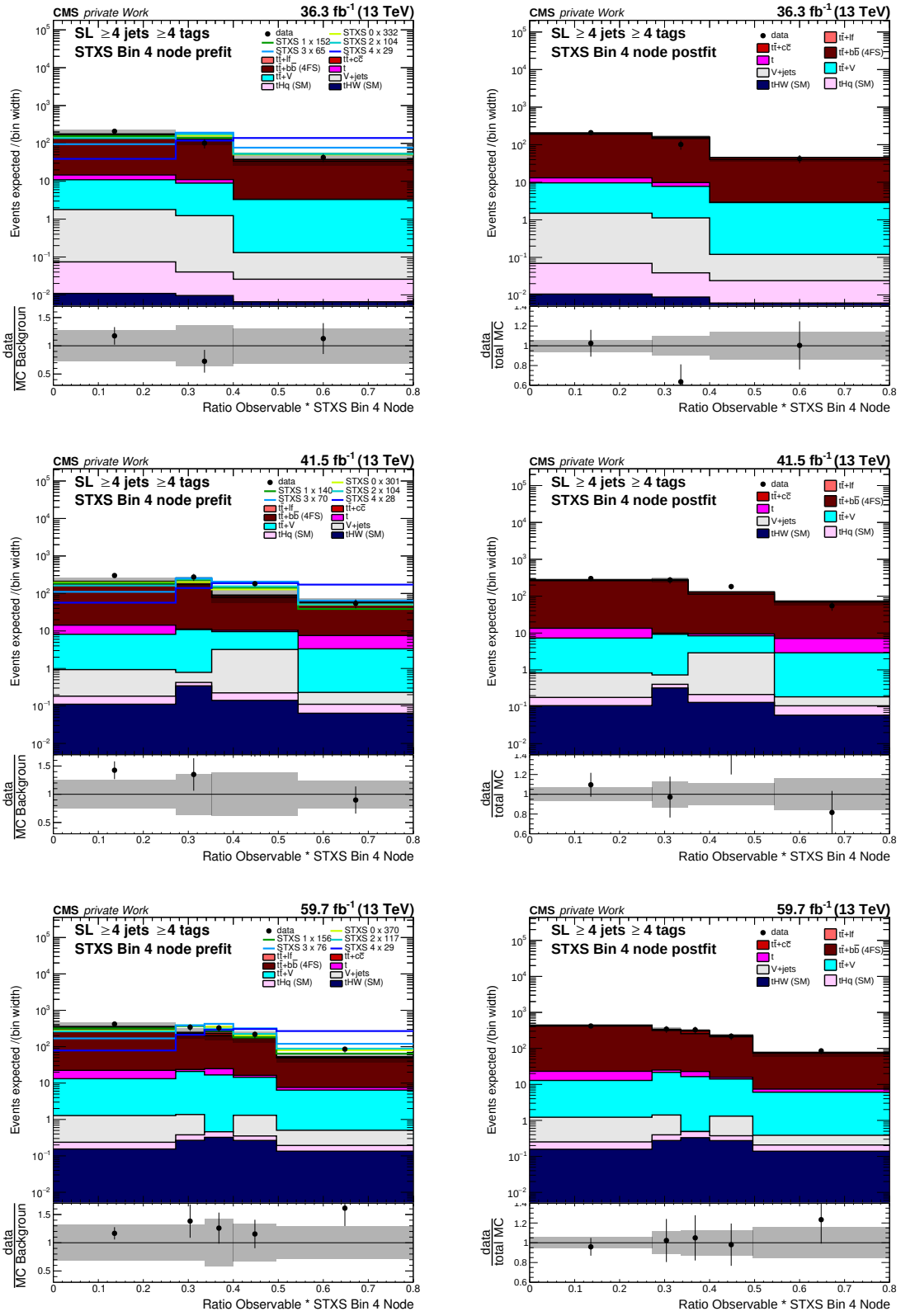


Figure 10.8: **Discriminants in the SL channel in the ratio observable times STXS ANN output for the last STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

Table 10.1: **Goodness-of-Fit p-values for the statistical model used for the cross section measurement**. Shown are the p-values for the individual SL analysis categories and their combination. Each block corresponds to either one of the years of data taking or the full Run-II luminosity. The values are sorted according to the p-values in each block, respectively.

Combination	p Values
$\geq$ 4j, $\geq$ 4t (2016)	0.99
Combined (2016)	0.96
$\geq$ 4j, 3t (2016)	0.87
Combined (2017)	0.98
$\geq$ 4j, $\geq$ 4t (2017)	0.97
$\geq$ 4j, 3t (2017)	0.88
$\geq$ 4j, 3t (2018)	1.00
Combined (2018)	0.98
$\geq$ 4j, $\geq$ 4t (2018)	0.87
$\geq$ 4j, $\geq$ 4t (full Run-II)	1.00
Combined (full Run-II)	0.94
$\geq$ 4j, 3t (full Run-II)	0.85

bin are also valid here. The observable shows a good separation of the STXS classes and the background processes. Moreover, the previously discussed separation between the individual STXS bins follows the behavior that is highlighted in the confusion matrices. Additionally, the coarse binning of the distributions indicates that the number of events in the category is much smaller than for the discriminant in the first STXS bin. Given the spectrum of the $p_{\mathrm{T}}^{\mathrm{H}}$ discussed in Chap. 2 and Sec. 5.3, this is in line with the expected behavior. Overall, all STXS categories display a reasonable compatibility of the simulated samples and the observed data within the uncertainties.

This compatibility is further quantified with the Goodness-of-Fit (GoF) method as introduced in Sec. 8.2. The resulting p-values are listed in Tab. 10.1 and further indicate excellent compatibility between the simulated samples and the observed data within the statistical model used for the cross section measurement.

10.3.2 Nuisance parameter pulls and impacts

The impact of the nuisance parameters on the cross section inference in the respective STXS bins is calculated as described in Sec. 8.3.2. Figure 10.9 shows the 20 nuisance parameters with the largest impact on the inference of the cross section for the first STXS bin. The expected ranking is defined by performing the parameter estimation with a pseudo data set which corresponds to the SM expectation, i.e. the sum of the signal and background simulations. The signal inference for the first STXS bin is expected to be dominated by modeling uncertainties on the $\mathrm{t\bar{t}}$ + jets backgrounds as well as experimental uncertainties and uncertainties due to the limited number of simulated events. These parameter groups are discussed in more detail in the following.

Since the $\mathrm{t\bar{t}}$ + jets processes are the most important background processes in the $\mathrm{t\bar{t}H(b\bar{b})}$ measurement, a large impact of the corresponding modeling parameters is reasonable. In particular, most of the high-ranking parameters are related to the modeling of the $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ process, such as the uncertainties on the renormalization and factorization scales, the

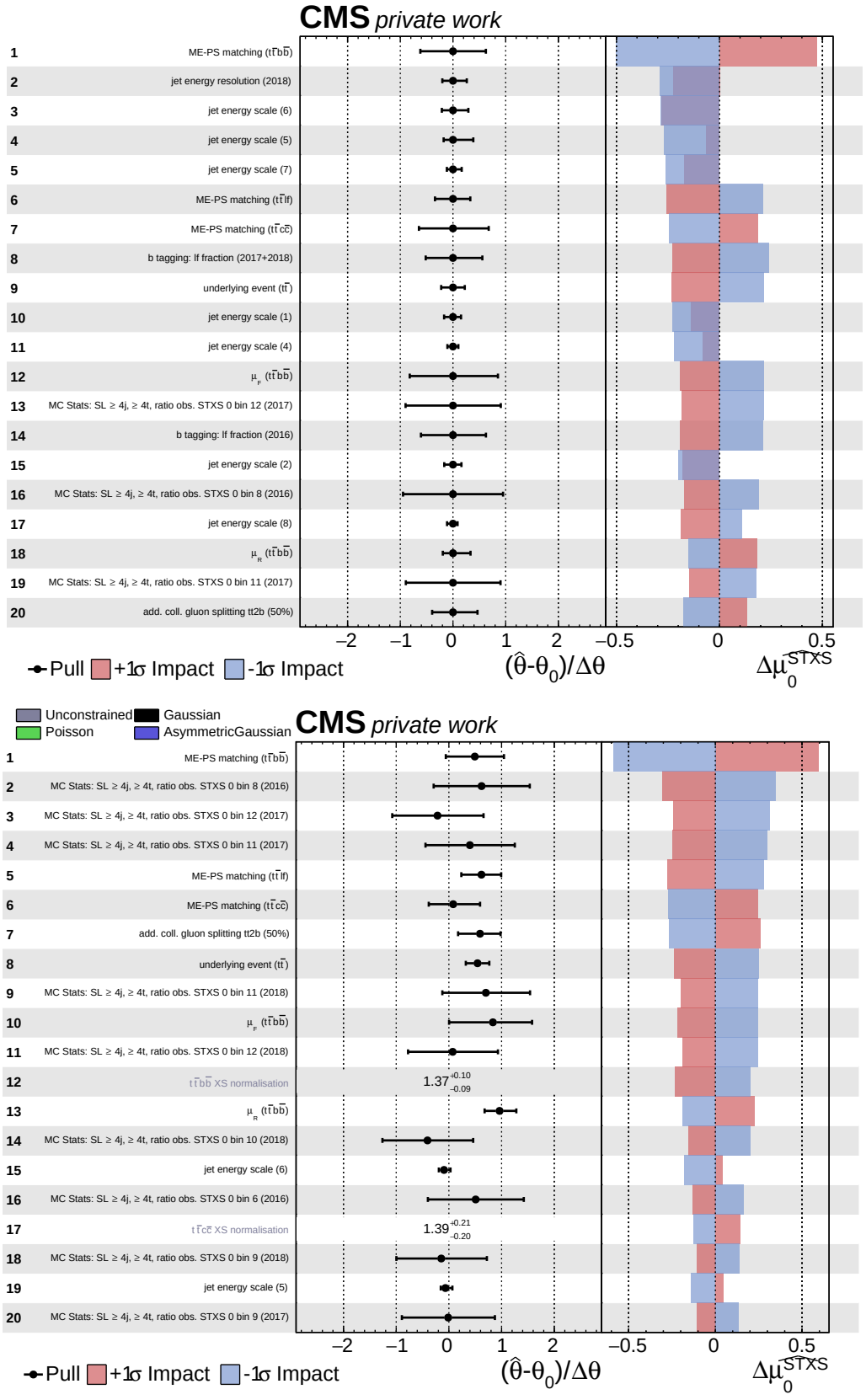


Figure 10.9: **Impacts in the first STXS bin for the full Run-II SL combination**. Shown are 20 nuisance parameters with the largest impact on the signal inference in the fit to a pseudo data set corresponding to the SM (top) and to the observed data (bottom). The post-fit pulls and constraints are shown in the middle panel. The right-hand side panel indicates the magnitude of the impact on the signal inference of the corresponding parameter.

ME-PS matching, the underlying event simulation and additional gluon splitting effects. This behavior is to be expected since this is an irreducible background in the $t\bar{t}H(b\bar{b})$ analysis.

Among the experimental uncertainties, different sources of the Jet Energy Scale (JES) and Jet Energy Resolution (JER) have the largest impact on the cross section inference in the first STXS bin. Since these uncertainties model variations in the energy measurement of the reconstructed jets, they can cause migrations of events in and out of the analysis phase space. These migration effects are expected to be most pronounced for jets with transverse momenta of the order of the lower limit in the jet selection as introduced in Sec. 4.3, which is 30 GeV. Furthermore, many of the JES and JER uncertainties are large for small values of the transverse momentum of reconstructed jets, as shown in Ref. [198]. Due to its definition in the low-$p_{\mathrm{T}}^{\mathrm{H}}$ region, the first STXS bin is the category that is the most likely to contain such lower-energy jets with respect to the other STXS regions. Therefore, it is plausible that the signal inference in the first STXS bin is susceptible to effects by the JES and JER. Additionally, it is notable that most of these uncertainties are strongly constrained. This is due to the limited number of simulated events in the corresponding varied distributions, an effect that has been observed and documented in the $t\bar{t}H(b\bar{b})$ analyses presented in Refs. [21, 133, 137]. Additionally, the b-tagging uncertainty modeling the effects of the $t\bar{t}+\mathrm{lf}$ contamination in the heavy-flavor control region used for deriving the b-tagging scale factors for heavy-flavor jets is expected to have a large impact. This contamination increases in the low-energy regime since additional heavy-flavor jets are less likely in this region. Of all STXS bins, the first is expected to be the most sensitive to this kinematic range. Therefore, the large impacts of the parameters modeling the $t\bar{t}+\mathrm{lf}$ contamination in the derivation of the b-tagging scale factors is reasonable.

Finally, the last group of uncertainties that are among the 20 parameters with the largest impact on the signal inference for the first STXS bin are related to the limited size of the simulated samples. These parameters vary the content of the bins to large discriminant values of the 2016 and 2017 distributions for the category dedicated to the first STXS bin in the $\geq 4\mathrm{j}$, $\geq 4\mathrm{t}$ region. As already discussed in the previous section, these bins are expected to be especially enriched in events attributed to the first STXS bin. Consequently, it is reasonable that parameters that cause a variation in these particular high-sensitivity bins have a large impact on the inference of the cross section in the first STXS bin.

Although the magnitude of the impacts of the different nuisance parameters is different, the list of uncertainties is generally similar to the list shown for the inclusive $t\bar{t}H(b\bar{b})$ analysis. Given the distribution of $p_{\mathrm{T}}^{\mathrm{H}}$ introduced in Secs. 2 and 5.3, this is expected since most of the $t\bar{t}H(b\bar{b})$ events should be located in the first three STXS bins. Consequently, the inclusive $t\bar{t}H(b\bar{b})$ analysis is dominated by the effects in the low-$p_{\mathrm{T}}^{\mathrm{H}}$ regime, which is why the general ranking of the impacts is comparable.

The bottom illustration of Fig. 10.9 shows the impact ranking in the fit to observed data. Overall, the ranking and the order of magnitude of the impacts agree well with the expectation, which indicates the good compatibility of the simulated samples and the observed data already observed in the previous section. The uncertainties on the modeling on the $t\bar{t}+\mathrm{jets}$ processes that were expected to have a large impact on the signal inference for the cross section in the first STXS bin show the same behavior in the fit to observed data. Additionally, the freely-floating normalization parameters for the $t\bar{t}+b\bar{b}$ and the $t\bar{t}+c\bar{c}$ processes are now among the 20 parameters with the largest impact. This behavior arises from the reduced constraints on these parameters with respect to the expected values, which is approximately 25 % for the $t\bar{t}+b\bar{b}$ and 10 % for the $t\bar{t}+c\bar{c}$ normalization. The increased parameter range for these uncertainties leads to a corresponding larger impact

on the signal inference, which is shown in the figure. As discussed in Sec. 8.3.2, a large impact of the freely-floating normalization parameter is to be expected when considering the final states of these processes.

The ranking in the fit to observed data includes fewer experimental uncertainties than expected. This is due to the fact that the parameter estimation yields tighter constraints on these parameters with respect to the expectation. Smaller constraints impede the impact of these uncertainties on the signal inference, which is why the magnitude of the impacts is smaller. Consequently, they are not among the 20 parameters with the largest impact.

Finally, the number of uncertainties related to the limited size of the simulated samples among the 20 parameters with the largest impact on the signal inference is larger than expected. This feature arises from the diminished impacts of the experimental uncertainties. While the magnitude of the impacts for these parameters decreases in the fit to observed data with respect to the expected values, the impact of the uncertainties on the limited size of the simulated samples is comparable with their expected impact on the signal inference. As a result, these uncertainties are among the 20 parameters with the largest impact on the cross section measurement in the first STXS bin. Similar to the discussion for the expected impacts of the uncertainties, the parameters that are now part of the top-ranking uncertainties modify the yield in the high-sensitivity bins of the category most-enriched in events attributed to the first STXS bin. Consequently, the large impact of these parameters on the signal inference is reasonable.

Generally, the size of the impacts and the post-fit constraints on the respective parameters are compatible with the expected values. Additionally, the post-fit estimations for the different parameters are compatible with the expected values within approximately two standard deviations, which indicates a reasonable degree of compatibility. This further validates the simulated samples and the statistical model used for the parameter estimation.

The upper part of Fig. 10.10 shows the expected impact ranking for the last STXS bin. Similar to the first STXS bin, uncertainties related to the modeling of the $\mathrm{t\bar{t}}$ + jets processes are expected to have a large impact on the cross section measurement in the last STXS bin. Additionally, the final-state radiation (FSR) uncertainties on the $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ and $\mathrm{t\bar{t}} + \mathrm{c\bar{c}}$ processes are among the 20 parameters with the largest impact. The probability for emissions in the final state is proportional to the energy of the particles. Therefore, it is reasonable that the FSR uncertainties for the most dominant background processes have a large impact on the measurement in the last STXS bin, which is expected to be sensitive to effects at high energies by construction. Furthermore, the PDF shape uncertainty for the $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ process has an impact of approximately 10 % on the signal inference. This is plausible since this parameter adds a degree of freedom to the modeling of the most signal-like background process, which is reflected in a corresponding correlation of approximately 9 %.

Compared to the measurement of the first STXS bin, there are fewer JES and JER uncertainties among the parameters with the largest impact on the inference in the last STXS bin. As discussed before, many of the JES and JER uncertainty sources decrease in size as a function of the reconstructed jet p_{T} [198]. Consequently, it is not surprising that fewer of these parameters have a large impact on the signal inference in the last STXS bin. Indeed, many of these uncertainties that are expected to be among the 20 parameters with the largest impact on the measurement are specifically constructed for reconstructed jets with large $|\eta|$ values. As shown in Fig. 10.4, some of the most important observables for the classification of events into the last STXS bin are related to the tHq hypothesis, which employs information of such jets. Therefore, the ranking of these parameters shown in the upper part of Fig. 10.10 is reasonable.

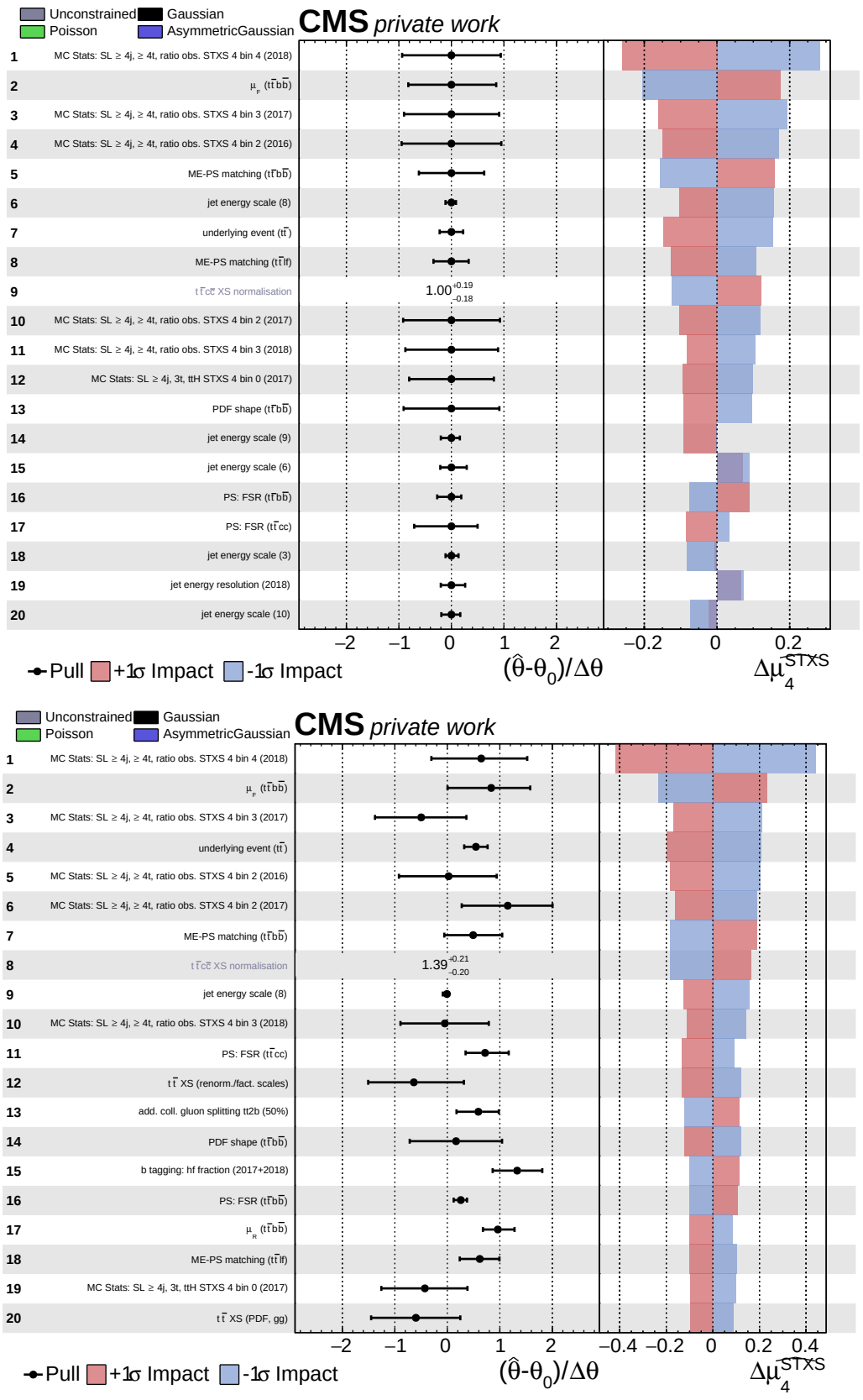


Figure 10.10: **Impacts in the last STXS bin for the full Run-II SL combination**. Shown are 20 nuisance parameters with the largest impact on the signal inference in the fit to a pseudo data set corresponding to the SM (top) and to the observed data (bottom). The post-fit pulls and constraints are shown in the middle panel. The right-hand side panel indicates the magnitude of the impact on the signal inference of the corresponding parameter.

Finally, uncertainties arising from the limited size of the simulated samples are expected to have a large impact on the measurement in the final STXS bin. Compared to the measurement in the first STXS bin, the fraction of such parameters among the 20 parameters with the greatest impact is larger. This is plausible considering the distributions in the categories dedicated to the last STXS bin as shown in Fig. 10.8. The binning of these distributions is much coarser, reflecting the smaller number of events expected for the high-$p_{\mathrm{T}}^{\mathrm{H}}$ region. As a result, the statistical inference in the last STXS bin is more susceptible to parameters that model the statistical fluctuations in this region, thus explaining their large impact.

The lower part of Fig. 10.10 displays the 20 parameters with the largest impact on the measurement in the last STXS bin in the fit to observed data. Compared to the expected ranking of the uncertainties, the fraction of JES- and JER-related parameters decreased and more modeling uncertainties for the $\mathrm{t\bar{t}}$ + jets background processes are ranked high. This can be explained by the stringent constraints placed on the experimental uncertainties, as previously discussed in the review of the impacts on the measurement in the first STXS bin. The parameters that migrated into the list of the 20 uncertainties with the largest impact on the measurement are related to the dominant $\mathrm{t\bar{t}}$ + jets background processes, which is reasonable as discussed previously. In general, the constraints and the size of the impacts for the different parameters agree well with the expected values. The post-fit estimations of the uncertainties agree with the expectation within approximately two standard deviations, which again indicates good agreement of the simulated samples and the observed data.

As shown in Chap. 9, the consideration of the dileptonic and fully-hadronic $\mathrm{t\bar{t}}$ decay modes increases the sensitivity of the inclusive $\mathrm{t\bar{t}H(b\bar{b})}$ measurement. The same is true for the cross section measurement in the scope of the STXS framework, which is indicated in the comparisons of the impacts shown in Figs. 10.11 and 10.12. For this comparison, the STXS measurement presented in Ref. [137] is modified to enable the interpretation in terms of a direct measurement of the cross section as explained in Sec. 10.1. The post-fit estimations for the different parameter values in the SL channel are compatible with the estimation in the full $\mathrm{t\bar{t}H(b\bar{b})}$ measurement within one standard deviation, which further validates the previously discussed measurements. Furthermore, the overall ranking of the parameters is similar to the previously presented results in the SL channel. Therefore, the previous explanations are also applicable to the fully-combined $\mathrm{t\bar{t}H(b\bar{b})}$ measurement.

The order of magnitude of the impacts in the fully-combined $\mathrm{t\bar{t}H(b\bar{b})}$ measurement is smaller than in the SL fits. Furthermore, the fraction of uncertainties related to the limited size of the simulated samples is larger with respect to the fit using the combined statistical model in the SL channel. This indicates that the additional information provided by the other decay channels improves the control over the theoretical and experimental uncertainties in the fully-combined $\mathrm{t\bar{t}H(b\bar{b})}$ measurement, thus making the parameters related to the limited number of events available for the parameter estimation more important. This effect is especially evident for the measurement in the last STXS bin, which is completely dominated by these uncertainties.

Overall, the impact of the nuisance parameters on the measurement as well as their post-fit estimations and constraints are reasonable, thus indicating that the statistical models in the SL channel and the fully-combined $\mathrm{t\bar{t}H(b\bar{b})}$ measurement are capable of describing the observed data. The following section will focus on the expected sensitivity of the cross section measurement.

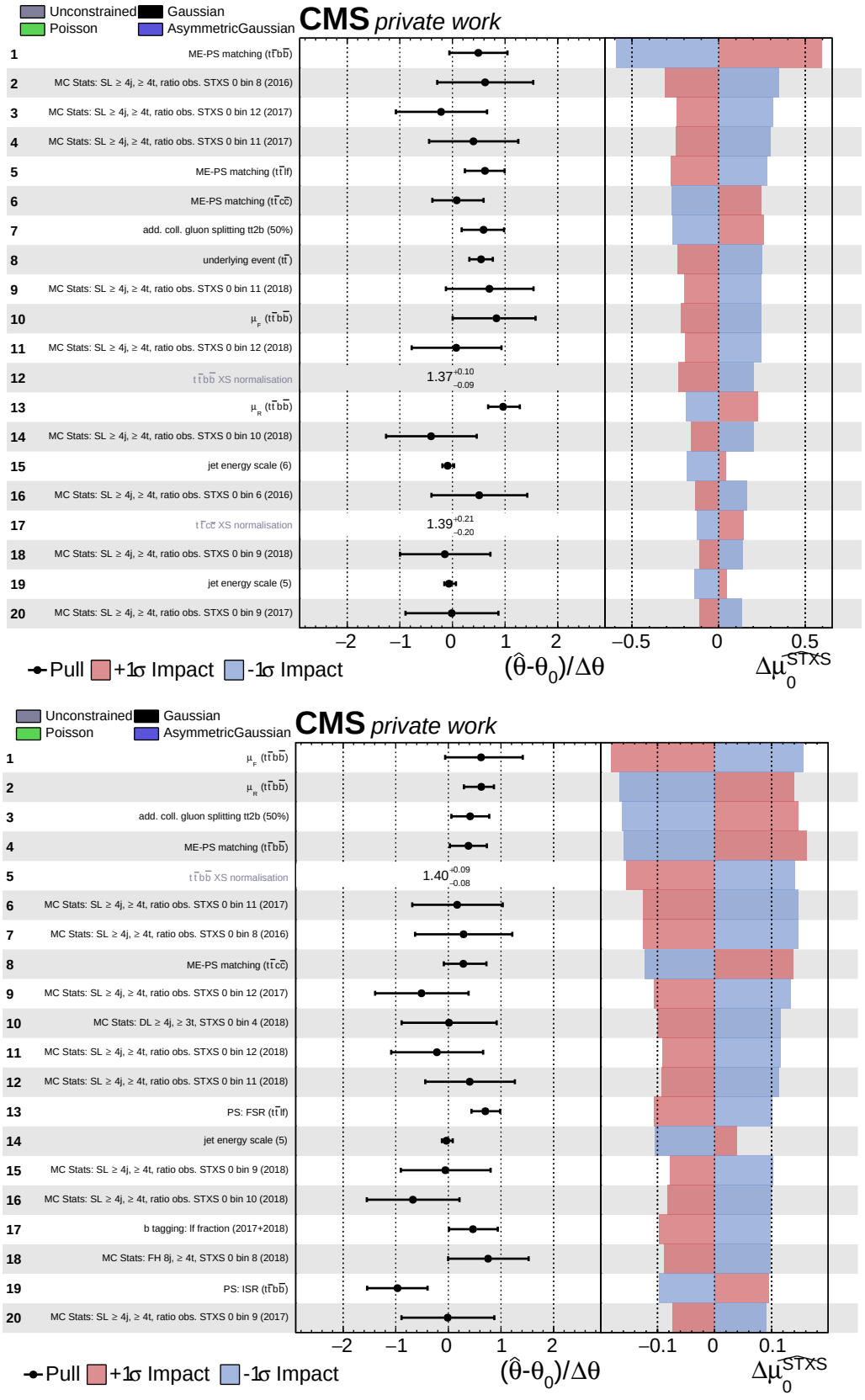


Figure 10.11: **Comparison of the impacts computed for observed data in the first STXS bin.** Shown are 20 nuisance parameters with the largest impact on the signal inference in the fit to the observed data in the statistical model for the full Run-II SL (top) and the full $t\bar{t}H(b\bar{b})$ combination (bottom). The post-fit pulls and constraints are shown in the middle panel. The right-hand side panel indicates the magnitude of the impact on the signal inference of the corresponding parameter.

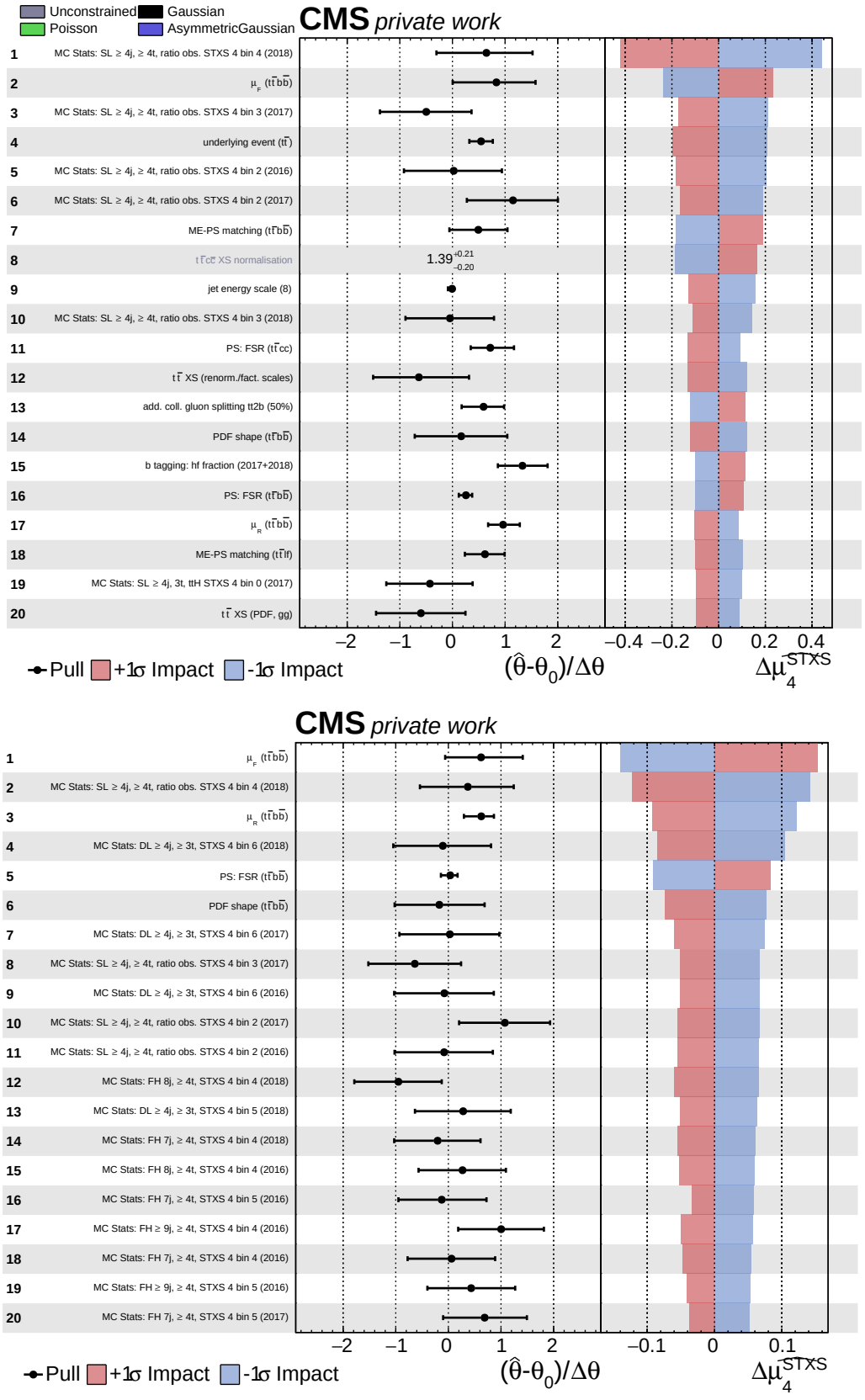


Figure 10.12: **Comparison of the impacts computed for observed data in the last STXS bin**. Shown are 20 nuisance parameters with the largest impact on the signal inference in the fit to the observed data in the statistical model for the full Run-II SL (top) and the full $t\bar{t}H(b\bar{b})$ combination (bottom). The post-fit pulls and constraints are shown in the middle panel. The right-hand side panel indicates the magnitude of the impact on the signal inference of the corresponding parameter.

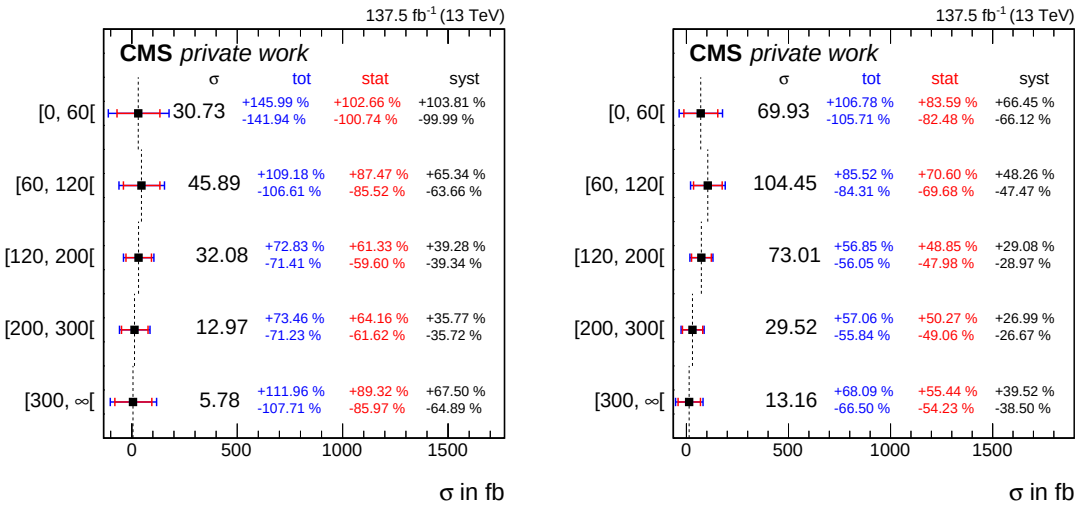


Figure 10.13: **Expected sensitivity of the cross section measurement**. The uncertainties are extracted in a fit to pseudo data corresponding to the full Run-II SL combination (left) and the full $t\bar{t}H(b\bar{b})$ measurement (right).

10.3.3 Cross section measurement

As already mentioned in the beginning of this section, the following discussion focuses on the preliminary expected sensitivity obtained in a fit to pseudo data corresponding to the SM predictions of the signal and background processes. Figure 10.13 shows the expected cross sections in the respective STXS bins and the corresponding uncertainties in the parameter estimation. All measurements are dominated by the statistical uncertainty, which is expected given the small $t\bar{t}H$ production cross sections in the individual STXS bins. The measurement of the first STXS bin in the SL channel shows a large contribution of the systematic uncertainty. This is in line with the behavior observed for the impacts of the nuisance parameters, which showed a large influence of both theoretical and experimental uncertainties on the inference in this STXS bin. The contributions of the systematic uncertainties decreases in the following STXS bins, before it increases again in the measurement in the last bin. The cross section inference in the third and forth STXS bin are expected to be the most precise measurements in the SL channel.

The right-hand side of Fig. 10.13 shows the measurement using the full $t\bar{t}H(b\bar{b})$ combination. The additional information provided by the DL and FH decay channels is expected to greatly improve the sensitivity in all STXS bins. This improvement is most pronounced in the measurement of the first and last bins, where the contributions of the systematic uncertainties decrease by up to 40 %. This is due to the same behavior as previously discussed in the comparison of the impacts of the nuisance parameters on the signal inference. The additional information provided by the other decay channels improves the control over the systematic uncertainties, thus improving the measurement. The statistical uncertainties improve by up to 30 % in the STXS bins with the smallest production probability with respect to the results obtained in the SL channel. This is the expected behavior as the full $t\bar{t}H(b\bar{b})$ phase space yields more data for the analysis, thus improving the statistical precision. A comparison with Fig. H.48 in App. H.2 indicates that the largest improvement in the full combination is expected to originate from the DL channel. Following the discussion in Chap. 9, this behavior is reasonable due to the high signal purity in this channel. The addition of the FH channel decreases the contributions of both the statistical and the systematic uncertainties by approximately 10 %. The expected

sensitivity of this measurement is comparable with preliminary results published by the ATLAS collaboration presented in Ref. [264].

Currently, the only available STXS measurement for the $t\bar{t}H$ production channel within the CMS collaboration analyzed events where the Higgs boson decays into two photons, which is presented in Ref. [265]. The precision of this measurement is between 52 % and 126 %, depending on the exact STXS bin. In particular, the uncertainties in the high-$p_{\mathrm{T}}^{\mathrm{H}}$ bins are especially large compared to the expected results presented in this thesis. Consequently, a combination of the analyses presented in Ref. [265] and in this thesis could potentially increase the sensitivity of future measurements significantly.

This STXS measurement is the first statistical analysis that provides details about $p_{\mathrm{T}}^{\mathrm{H}}$ in the $t\bar{t}H(b\bar{b})$ channel in the CMS collaboration. Apart from the tests of the SM discussed so far in this chapter, this also allows for powerful tests of BSM theories. The subsequent section will discuss the potential contribution to such tests that the $t\bar{t}H(b\bar{b})$ analysis can provide.

10.4 Comparison with the Standard Model Effective Field Theory

As already discussed in Chap. 2, the $t\bar{t}H$ production channel is potentially sensitive to BSM contributions. To this end, the predictions of the Standard Model Effective Field Theory (SMEFT) are compared to the expected cross section measurements in the scope of the STXS framework reviewed in the previous section. As already discussed, the full $t\bar{t}H(b\bar{b})$ combination yields the highest precision for the cross section measurements. Therefore, the results obtained in full combination are used to gauge the potential sensitivity of the $t\bar{t}H(b\bar{b})$ channel to the different SMEFT operators at next-to-leading-order QCD perturbation theory: the chromomagnetic dipole moment $\mathcal{O}_{tG}$, the effective gluon-Higgs coupling $\mathcal{O}_{\varphi G}$ and the shift in the top-Higgs Yukawa coupling $\mathcal{O}_{t\varphi}$. The comparison is performed using the simulated samples described in Sec. 5.3 in two schemes.

In the individualized scheme, the three operators of interest introduced in Chap. 2 are considered individually, while the contributions of the other operators are neglected. The results are shown in Fig. 10.14. The constraints on the contributions of the three SMEFT operators in this scheme from other measurements are already strong, which is indicated by the uncertainty bands in the corresponding colors. The simulated predictions for the individual SMEFT operators are compatible with the corresponding SM prediction. As discussed in Ref. [69], this is driven by the high precision of other Higgs-related analyses, such as the gluon fusion production channel. While the $t\bar{t}H(b\bar{b})$ channel is indeed sensitive to the BSM contributions, the current expected sensitivity of the measurement in the STXS framework is too small to constrain the contributions in this scheme.

In the marginalized scheme, the contributions of all SMEFT operators are considered by using the marginalized likelihood for the parameter estimation as described in Ref. [69]. The comparison of the predictions in this approach to the cross section measurement in the $t\bar{t}H(b\bar{b})$ analysis are shown in Fig. 10.15. Since the posterior correlations of contributions of the different operators are not known for the event simulation, the other operators are fixed to their central values for the calculation of the uncertainties for the individual contributions. Consequently, the parameter estimation considers the contribution of all SMEFT operators simultaneously, resulting in larger confidence interval compared to the individualized scheme. Additionally, the central values in this approach are larger, thus increasing the contribution of the individual SMEFT operators. This is especially apparent in the high-$p_{\mathrm{T}}^{\mathrm{H}}$ bin, where the deviation from the corresponding SM prediction is most

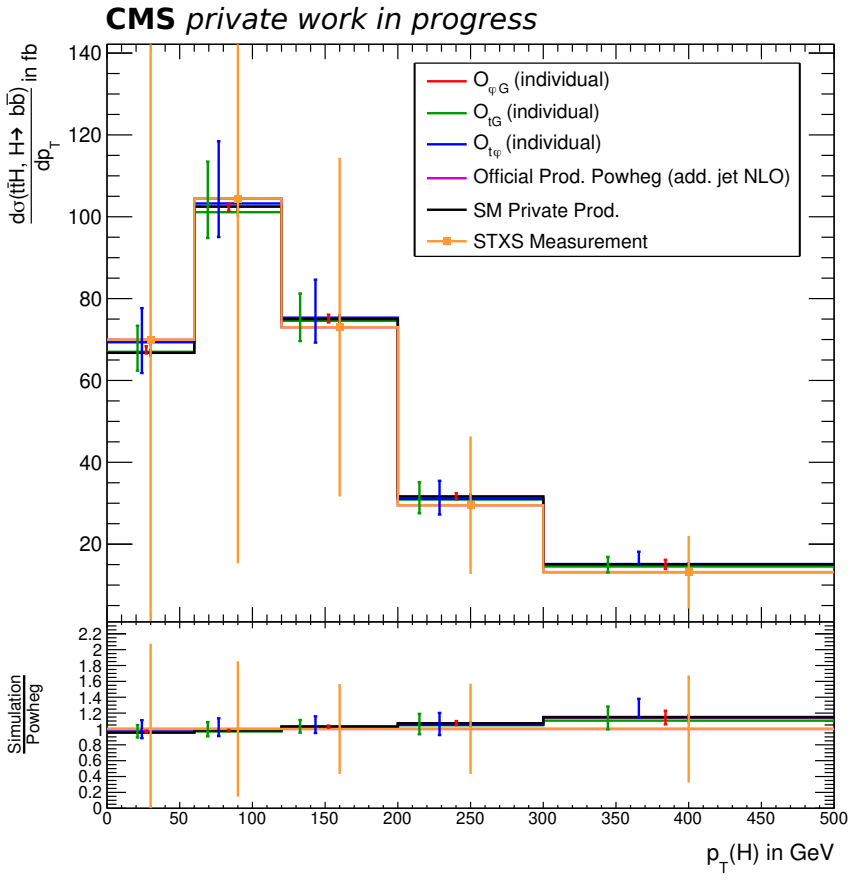


Figure 10.14: **Comparison to the SMEFT predictions using the individual scheme**. The STXS measurement corresponds to the expected sensitivity obtained in the full $t\bar{t}H(b\bar{b})$ combination. The values used to generate the SMEFT predictions and the uncertainties indicated by the vertical bars are listed in Tab. 5.7.

pronounced. The p_T^H spectrum indicates especially large variations due to the shift in the top-Higgs Yukawa coupling and the monochromatic dipole moment. This is in good agreement with the expected behavior as presented in Ref [17].

Overall, the expected precision of the cross section measurement in the $t\bar{t}H(b\bar{b})$ analysis is comparable with the current sensitivity of the parameter estimation presented in Ref. [69]. In particular, the measurement of the high-p_T^H region is expected to further contribute to the sensitivity to these SMEFT operators. This is in line with the recently-published discussion presented in Ref. [18], where the authors show that especially analyses targeting the $H \to b\bar{b}$ decay channel in the boosted regime are sensitive to contributions originating from the chromomagnetic dipole and the modified top-Higgs Yukawa coupling. The results shown in Fig. 10.15 confirm this claim. Consequently, the $t\bar{t}H(b\bar{b})$ channel can contribute to the sensitivity in future iterations of analyses aiming to measure the SMEFT contributions.

After this preliminary test of the potential constraining power of the $t\bar{t}H(b\bar{b})$ analysis, future dedicated measurements in this channel are of special interest in the search for BSM physics. Contrary to the results at truth-level shown here, this requires the full simulation chain common in high-energy physics. Moreover, a statistical model with a proper parameterization of the SMEFT effects needs to be constructed, which will enable dedicated parameter estimations similar to the results presented in this thesis. Such

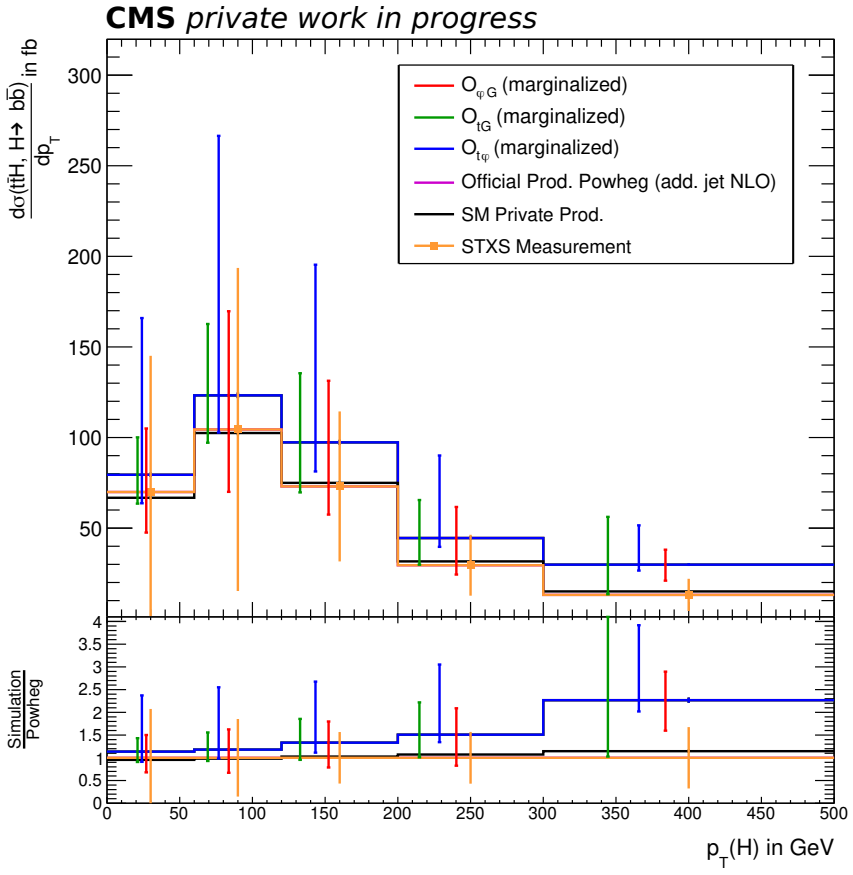


Figure 10.15: **Comparison to the SMEFT predictions using the marginalized scheme**. The STXS measurement corresponds to the expected sensitivity obtained in the full $t\bar{t}H(b\bar{b})$ combination. The central values for the three SMEFT operators in this scheme and thus their nominal predictions is the same in this scheme. The values used to generate the SMEFT predictions and the uncertainties indicated by the vertical bars are listed in Tab. 5.7.

measurements pose challenges that are yet to be addressed and are beyond the scope of this thesis.

11 Projected expected sensitivity of the $t\bar{t}H(b\bar{b})$ measurement at the HL-LHC

After discussing the current state of the analysis in Part II and the first steps towards differential measurements in Chap. 10, it is intuitive to investigate the future of the $t\bar{t}H(b\bar{b})$ analysis. In the scope of this thesis, the expected sensitivity of the $t\bar{t}H(b\bar{b})$ analysis was projected to future data sets, in particular at the HL-LHC. The future scenarios considered in this study are listed in Tab. 11.1.

The initial values for the interpolation are based on the $t\bar{t}H(b\bar{b})$ analysis using the data set recorded in 2016 with $35.9\,\text{fb}^{-1}$ described in Ref. [133]. The studies contributed to the formulation of the European strategy for particle physics and are summarized in Refs. [33, 266]. The following sections will briefly discuss these results. First, the assumptions and scenarios of the projection studies are discussed. Afterwards, the results of the study are discussed and compared with the current sensitivity of the $t\bar{t}H(b\bar{b})$ analysis achieved at the end of Run-II of the LHC.

11.1 Projection scenarios and assumptions

As already mentioned, the starting point for the projection is the $t\bar{t}H(b\bar{b})$ analysis with the dataset recorded in 2016 with the CMS detector. This implies that no improvements

Table 11.1: **Future data sets considered for the sensitivity projection**. The projections are performed by scaling the expected event yields and uncertainty constraints to the respective integrated luminosity L_{int}.

Future Dataset	Integrated Luminosity L_{int} in fb^{-1}
Initial dataset (2016)	35.9
Full LHC Run-II†	100
Full LHC Run-III	300
First runs of HL-LHC	1000
Full HL-LHC	3000

†: Value corresponds to luminosity expected prior to LHC Run-II, does not correspond to final luminosity achieved after Run-II

Table 11.2: **Uncertainty scaling in the projection scenarios**. Scaling of the systematic uncertainties relative to the 2016 data analysis under the scenarios S1 and S2. The value R_{int} is the ratio of the assumed L_{int} with respect to $35.9\,\text{fb}^{-1}$. The scenarios are further defined in the text. Taken from [266].

Source	S1	S2
b tag HF/LF (stat)	same	$1/\sqrt{R_{\text{int}}}$
b tag HF/LF (others)	same	same
lepton ID/iso	same	$1/\sqrt{R_{\text{int}}}$ (floor 0.5)
lepton trigger	same	$1/\sqrt{R_{\text{int}}}$
JER; JES, flavor components; JES, resolution components	same	$1/\sqrt{R_{\text{int}}}$ (floor 0.5)
JES, pileup components	same	same
JES, absolute components	same	$1/\sqrt{R_{\text{int}}}$ (floor 0.3)
JES, relative components	same	$1/\sqrt{R_{\text{int}}}$ (floor 0.2)
JES, time components; JES, method components	same	$1/\sqrt{R_{\text{int}}}$
luminosity	same	1%
theory uncertainties	same	halved
incl. $t\bar{t}$ cross-section (renorm./fact. scales, PDF)		
incl. $t\bar{t}H$ cross-section (renorm./fact. scales, PDF)		
μ_R/μ_F scales (ME generator)		
PS ISR/FSR/matching		
underlying event		
pileup		
PDF (shape contribution)		
$t\bar{t}$ + HF background norm (add. 50% cross-section uncertainty)	same	halved
MC stat.	none	none

in the analysis strategy since then are covered within the study. The projection is later compared to the full Run-II result of the $t\bar{t}H(b\bar{b})$ analysis discussed in Part II to validate this assumption.

In the scope of the projection studies, experts in both experimental and theoretical physics formulated two scenarios to estimate the physics potential of the CMS detector:

Run-2 systematic uncertainties (S1): The performance of the CMS detector and the accuracy of the theoretical predictions are assumed to stay constant. This scenario is used as a conservative baseline.

YR systematic uncertainties (S2): Theoretical uncertainties are halved to account for more accurate calculations and techniques. The experimental uncertainties are scaled as a function of the square root of the integrated luminosity L_{int} until the uncertainties reach a defined minimum value. This accounts for the improved statistical sensitivity of experimental measurements in e.g. dedicated control regions. The floor values are chosen according to the expected performance of the physics objects with the upgraded CMS detector at the HL-LHC summarized in Sec. 3.4 and presented in more detail in Ref. [80].

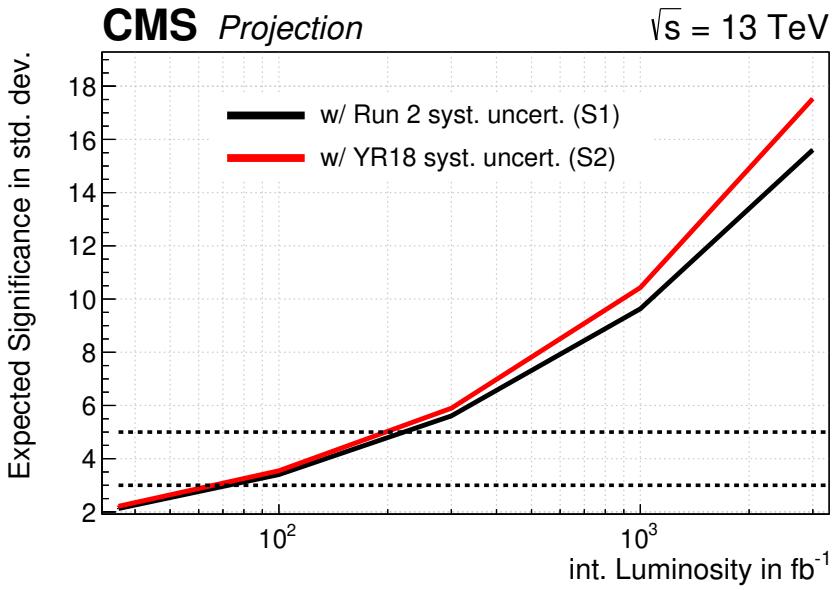


Figure 11.1: **Expected significance on $\mu_{\mathrm{t\bar{t}H}}$ as a function of the integrated luminosity**. The expected significance of the $\mathrm{t\bar{t}H(b\bar{b})}$ measurement is shown in units of Gaussian standard deviations. The solid lines show the results for the two scenarios considered in this projection study. The dashed lines indicate the significance required to claim evidence ($3\,\sigma$) and an observation ($5\,\sigma$) of $\mathrm{t\bar{t}H(b\bar{b})}$.

In both scenarios, the uncertainty due to the limited number of simulated events is neglected, assuming sufficiently large samples in future iterations of the analysis. The assumptions for both scenarios are summarized in Tab. 11.2. Technically, the prior uncertainty that enters the test statistic used for the statistical inference is scaled as depicted in the table. More details can be found in Ref. [266].

The following section discusses the expected sensitivity for these scenarios.

11.2 Projected sensitivity

The expected sensitivity is studied in terms of expected significance and the corresponding uncertainty on the $\mathrm{t\bar{t}H(b\bar{b})}$ signal strength modifier $\mu_{\mathrm{t\bar{t}H}}$ as defined in Sec. 8.1. Figure 11.1 shows the expected significance for $\mu_{\mathrm{t\bar{t}H}}$ as a function of the integrated luminosity. The significance grows continuously with increased luminosity, signifying the rising sensitivity due to the larger amount of data. Both scenarios agree well up to approximately $300\,\mathrm{fb}^{-1}$. Beyond this point, the optimistic scenario S2 projects a higher sensitivity. This is expected due to the reduction of the prior systematic uncertainties in this scheme that becomes dominant with increased luminosity. Both scenarios predict an observation of the $\mathrm{t\bar{t}H(b\bar{b})}$ process with approximately $200\,\mathrm{fb}^{-1}$, which corresponds to slightly more than the data recorded at the CMS experiment during Run-II of the LHC. Both scenarios indicate a high sensitivity to the $\mathrm{t\bar{t}H(b\bar{b})}$ process at the time of the HL-LHC. This will enable future analyses to perform more sophisticated measurements of e.g. differential distributions, which will in turn yield more insight and sensitivity to predictions of the Standard Model and beyond.

To gain a deeper understanding of the evolution of the sensitivity as a function of the integrated luminosity, the uncertainty in the measurement of $\mu_{\mathrm{t\bar{t}H}}$ is discussed next. The total uncertainty is split into groups according to the following sources:

Table 11.3: **Projected contributions of the different uncertainty groups to the total uncertainty in the estimation of $\mu_{t\bar{t}H}$.** All values are given in percent relative to $\mu_{t\bar{t}H} = 1$. Table taken from Ref. [266]

	S1			S2		
Source	$35.9\,\text{fb}^{-1}$	$300\,\text{fb}^{-1}$	$3000\,\text{fb}^{-1}$	$35.9\,\text{fb}^{-1}$	$300\,\text{fb}^{-1}$	$3000\,\text{fb}^{-1}$
Total	48.7	20.4	11.1	46.1	17.6	7.3
Stat	26.7	9.3	2.9	26.7	9.3	2.9
Sig. Theory	10.8	9.3	8.7	5.0	4.5	4.4
Bkg. Theory	28.6	10.3	4.1	25.6	9.6	3.5
Add. $t\bar{t}$ + HF XS	14.6	2.6	0.8	16.5	3.1	0.7
Exp	17.4	8.7	4.2	16.6	6.7	2.6
Luminosity	1.6	1.8	1.8	0.5	0.7	0.8
B tagging	12.0	6.1	2.8	10.8	4.4	1.6
JES	10.9	4.5	1.6	11.3	4.4	1.6

Stat: the statistical uncertainty of the fit;

Sig. Theory: all $t\bar{t}H(b\bar{b})$-related theory uncertainties, specifically on the inclusive production cross section;

Bkg. Theory: all theory uncertainties related to the $t\bar{t}$ background processes, namely: parton shower (PS) initial-state radiation (ISR)/final-state radiation (FSR)/ME-PS matching, underlying event, pileup, shape contributions of the μ_R/μ_F scales in the matrix element (ME) calculation and the PDF set, effects on the inclusive cross section (renormalization/factorization scales, PDF) and $t\bar{t}$ + HF background normalization (additional 50 % cross section uncertainty);

- **add $t\bar{t}$ + HF XS:** the subgroup of the background theory uncertainty that contains the additional 50 % $t\bar{t}$ + HF background normalization uncertainty parameters, i.e. for the $t\bar{t}$ + $b\bar{b}$, $t\bar{t}$ + 2b, $t\bar{t}$ + b and $t\bar{t}$ + $c\bar{c}$ normalizations, respectively;

Exp: uncertainties related to the experimental setup. This group is further divided into the following subgroups:

- **Luminosity:** the luminosity uncertainty;
- **b-tagging:** all sources of uncertainties related to the b-tagging of heavy-flavor, light-flavor and charm-flavor jets;
- **JES:** all Jet Energy Scale (JES) uncertainties;

The projected uncertainties are shown in Fig. 11.2. Table 11.3 lists the values for a subset of these projections. Since the discriminants are the same in both scenarios, the contributions of the statistical uncertainties to the sensitivity are the same. They approximately fall with $1/\sqrt{L_{\text{int}}}$ which is also the expected behavior for Poisson statistics.

The theoretical uncertainties on the background prediction (Bkg. Theory) has the largest contribution to the total expected uncertainty at $35.9\,\text{fb}^{-1}$. This behavior is mainly driven by the $t\bar{t}$ + HF normalization parameters, which are constructed with a large prior uncertainty. The $t\bar{t}$ + HF background consists of the irreducible background processes for the $t\bar{t}H(b\bar{b})$ measurement. Since the distinction between the processes is very difficult, large uncertainties on the background processes are expected to have a sizable impact on the $t\bar{t}H(b\bar{b})$ measurement. Therefore, the large contribution to the total uncertainty is to be expected. In both scenarios, the contributions of the background theory uncertainties are

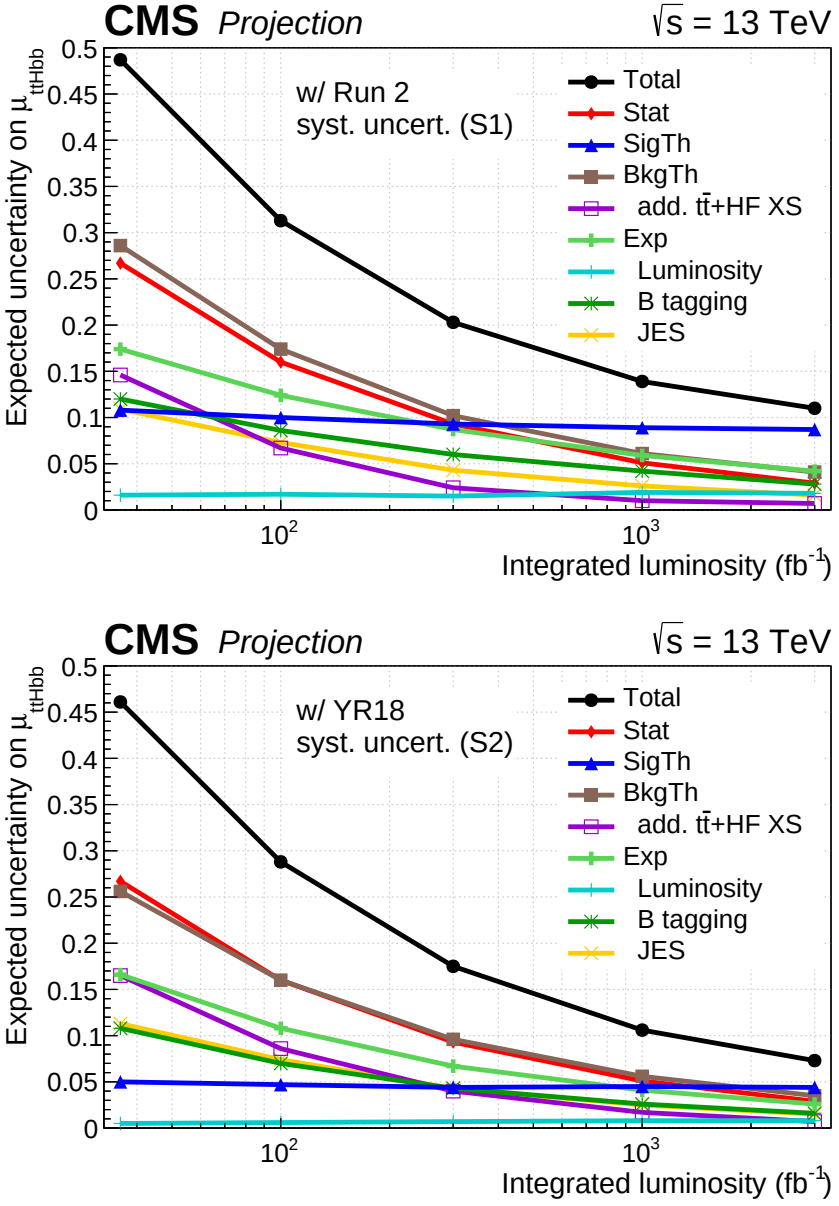


Figure 11.2: **Expected uncertainties on the $t\bar{t}H(b\bar{b})$ signal-strength modifier as a function of the integrated luminosity**. Shown are the total expected uncertainty as well as the contributions of the respective groups of uncertainty sources for the conservative Run-2 systematic uncertainties (S1) (top) and the optimistic YR systematic uncertainties (S2) (bottom) scenarios. Figures taken from Ref. [33].

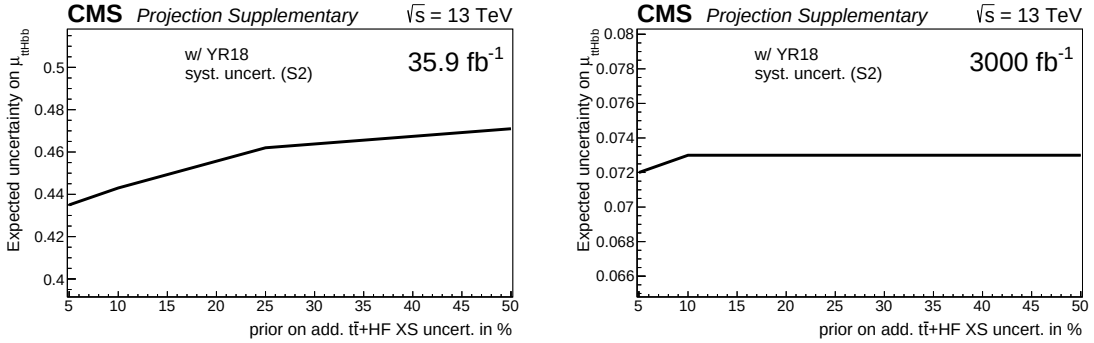


Figure 11.3: **Expected uncertainty on $\mu_{\mathbf{t\bar{t}H}}$ as a function of the $\mathbf{t\bar{t}+HF}$ normalization uncertainties**. Shown is the total expected uncertainty on the $\mu_{t\bar{t}H}$ estimation as a function of the prior uncertainty assigned to the $t\bar{t}+\text{HF}$ normalization in the optimistic scenario S2 at $35.9\,\text{fb}^{-1}$ (left) and $3000\,\text{fb}^{-1}$ (right). Figures taken from supplementary material of Ref. [266].

strongly decreasing with an increasing amount of luminosity. This is reasonable since the increased statistical precision allows for a more powerful measurement of the background, thus constraining the uncertainties and limiting the influence on the signal inference.

It is noticeable that the contribution of the background theory uncertainty group is very similar between the scenarios, despite the reduction of the a priori uncertainties in scenario S2 by a factor of two. To further understand this effect, the projection was performed as a function of the prior uncertainty on the $t\bar{t}+\text{HF}$ normalization. The results are summarized in Fig. 11.3. The study showed that already at $35.9\,\text{fb}^{-1}$, a reduction of the a priori uncertainty by a factor of ten leads to a relative improvement in the total post-fit uncertainty on $\mu_{t\bar{t}H}$ of 8 %. The projection for $3000\,\text{fb}^{-1}$ indicates a relative improvement of 3 %. Consequently, a mere reduction of the a priori uncertainty on the modeling of the irreducible background improves the sensitivity on the $t\bar{t}H(b\bar{b})$ measurement only by a little, which necessitates improvements in the modeling of these processes. This was one of the main inspirations for the changes in the analysis strategy which are implemented in the full Run-II analysis described in Part II.

The experimental uncertainties have a sizable contribution to the total uncertainty of the $\mu_{t\bar{t}H}$ measurement. At $35.9\,\text{fb}^{-1}$, this group is dominated by the b-tagging and JES uncertainties. This is very reasonable since the event selection and signal discrimination depends heavily on the b-tagging. In both scenarios, these uncertainties decrease with increased statistical precision and thus growing constraining power of the different processes involved in the analysis. The contribution of the b-tagging uncertainties decreases faster in S2, indicating that the improved performance expected for the CMS detector has the potential to improve the sensitivity of the analysis. The contribution of the JES uncertainties is very similar between the scenarios. This suggests that the enhanced performance of the detector will not be able to improve the sensitivity with the strategy for the JES uncertainties used in these projection studies. This necessitates a more refined parameterization of these sources of uncertainty, which is yet to be defined. The projection also clearly shows the increasing relative importance of the luminosity uncertainty. While it only contributes little at $35.9\,\text{fb}^{-1}$, it is one of the dominant experimental uncertainties at the full HL-LHC statistical precision. Since this source of uncertainty cannot be constrained in situ in any cross section measurement by construction, it is essential to improve the precision by dedicated measurements. This is indicated by S2, which shows that the

contribution of the luminosity uncertainty can be reduced by further reducing the a priori uncertainty entering the $t\bar{t}H(b\bar{b})$ analysis.

Finally, the theory uncertainties on the signal prediction contribute to the total expected uncertainty on the $t\bar{t}H(b\bar{b})$ measurement. Much like the luminosity uncertainty, these uncertainties have a direct impact on the signal inference and can therefore not be constrained in situ. Consequently, the signal theory uncertainties gain relative importance as the statistical power of the analysis increases. At the end of the HL-LHC, this group of parameters is expected to be the dominant source of uncertainty in the $t\bar{t}H(b\bar{b})$ measurement. This clearly illustrates the importance of more precise theory calculations and simulations of the $t\bar{t}H(b\bar{b})$ process.

The impact of individual sources of uncertainty on the $t\bar{t}H(b\bar{b})$ measurement is shown in Figs. 11.4 and 11.5 at $35.9\,\mathrm{fb}^{-1}$ and $3000\,\mathrm{fb}^{-1}$, respectively. This quantity is calculated as described in Sec. 8.3.3. Each nuisance parameter is set to its $\pm1\,\sigma$ post-fit confidence interval value. At each of these points, the signal strength is re-estimated. Finally, the impact is the difference between the global best-fit estimation of $\mu_{t\bar{t}H}$ and the re-estimated values. The ranking of the nuisance parameters generally reflects the effects discussed above for the different uncertainty groups. At $35.9\,\mathrm{fb}^{-1}$, the uncertainty with the largest impact on the signal inference by far is the additional 50 % cross section uncertainty on the $t\bar{t} + b\bar{b}$ process. Since this process is the irreducible background to the signal process, it is reasonable that a large uncertainty that changes the $t\bar{t} + b\bar{b}$ yield causes some ambiguity when measuring the yield of a similar processes such as the $t\bar{t}H(b\bar{b})$ process. Whenever the background yield is modified, the signal yield has to be updated as well to mitigate the ensuing differences.

The impact ranking changes at $3000\,\mathrm{fb}^{-1}$. The most important change with respect to the projected impacts at $35.9\,\mathrm{fb}^{-1}$ is the reduction of the impact of the additional cross section uncertainties on the $t\bar{t} + \mathrm{HF}$ background processes in both scenarios. This is due to the increased amount of data, which allows a precise measurement of the yield of these processes. Consequently, the constraints on the additional cross section uncertainties are strong, limiting their impact on the signal inference. Moreover, the parameters with the highest impacts on $\mu_{t\bar{t}H}$ are related to the calculation of the inclusive cross section of $t\bar{t}H$. As already mentioned in the discussion of the contributions to the total expected uncertainty on $\mu_{t\bar{t}H}$ by the respective uncertainty groups, these uncertainties cannot be constrained in situ during the fit. On the other hand, the other parameters are constrained, which reduces their impact on the measurement. Therefore, the relative importance of the uncertainties on the inclusive $t\bar{t}H$ cross section increases. This effect becomes more pronounced with higher luminosities and is most prominent at $3000\,\mathrm{fb}^{-1}$. As previously discussed, this shows the importance of future improvements of the calculations for the $t\bar{t}H$ process in order to gain more sensitivity in future $t\bar{t}H(b\bar{b})$ measurements.

In scenario S1, most of the parameters with the highest expected impact are related to the b-tagging procedure and the modeling of the $t\bar{t} + \mathrm{jets}$ background processes. This is very close to the behavior at $35.9\,\mathrm{fb}^{-1}$, which is expected since the uncertainty model in S1 stays the same. While some of the constraints on the nuisance parameters certainly benefit from the improved statistical precision, the underlying prior uncertainties do not change. Therefore, it is reasonable that uncertainty sources that are not benefiting from increased number of events, such as the b-tagging or background model parameters, have a high impact at both luminosities.

Contrary to this behavior, only the group of these nuisance parameters with the largest prior uncertainties are ranked high in S2. This is due to the reduction of the a priori uncertainties in this scenario, which becomes stronger with increasing luminosity especially

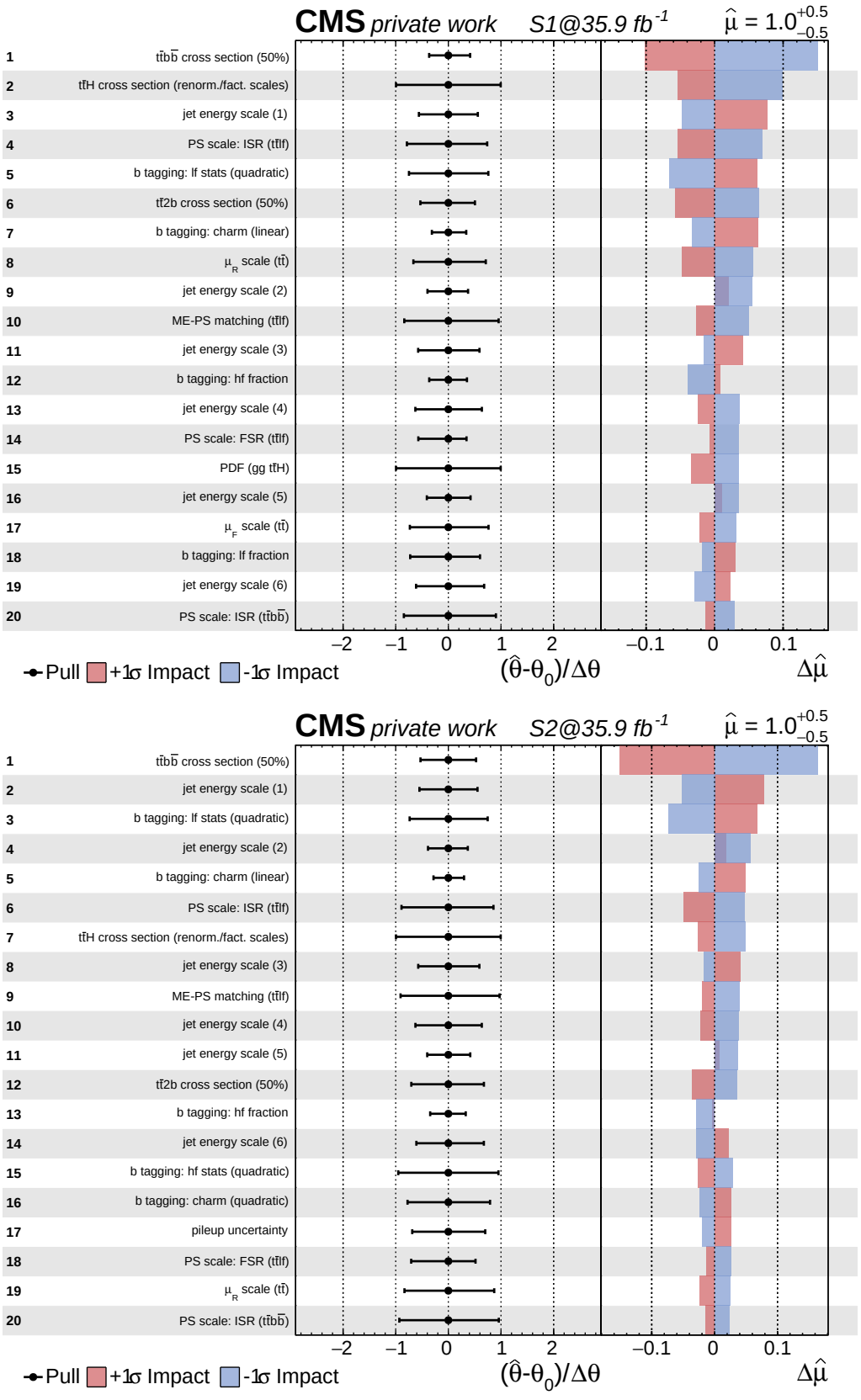


Figure 11.4: **Top 20 parameters with the largest projected impact on the $\mu_{t\bar{t}H}$ estimation at 35.9 fb^{-1}.** Shown are the 20 parameters with the largest projected impact on $\mu_{t\bar{t}H}$ at 35.9 fb^{-1} in scenario S1 (top) and S2 (bottom), respectively. All impacts were calculated with pseudo data generated from the sum of the signal and background predictions.

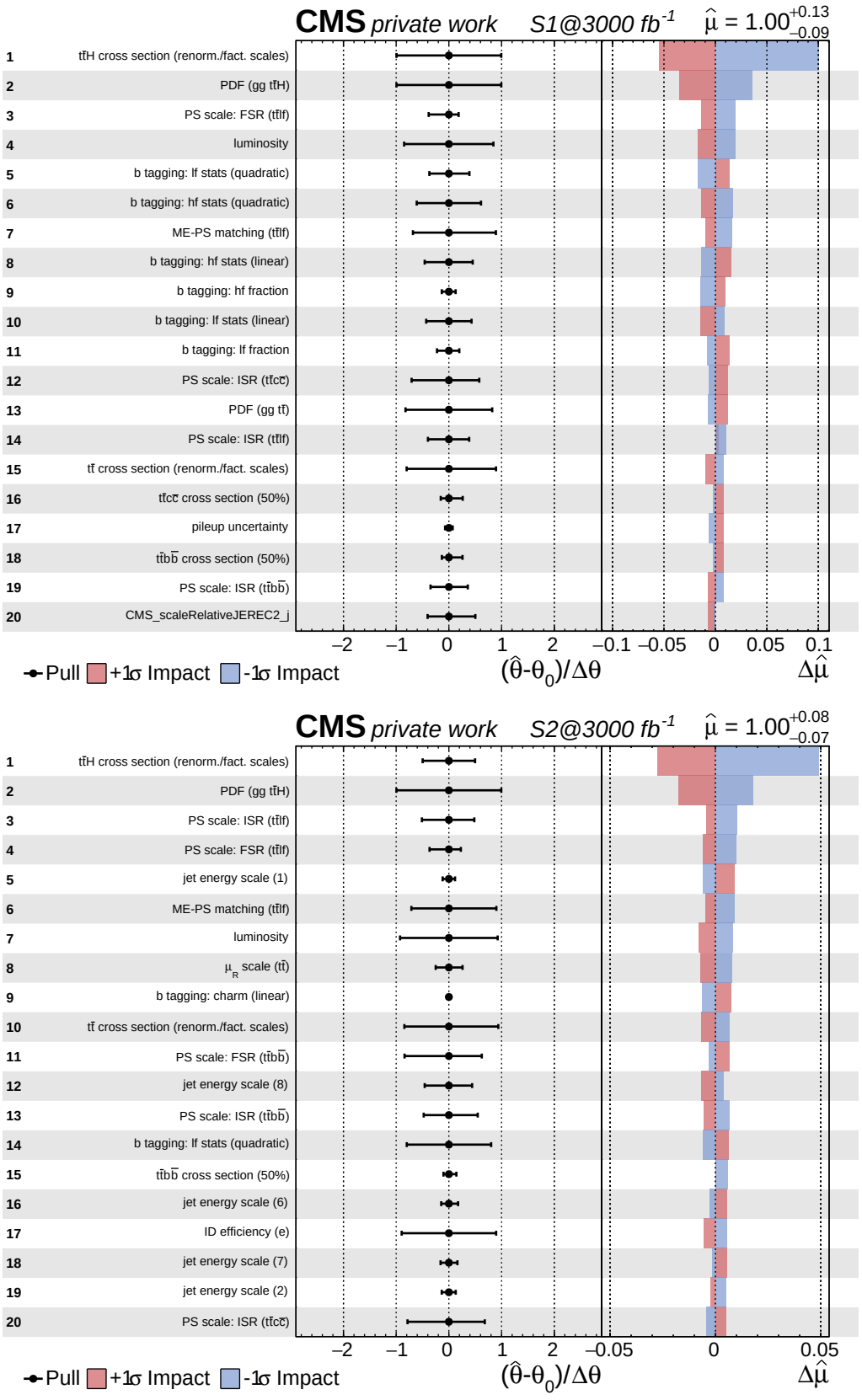


Figure 11.5: **Top 20 parameters with the largest projected impact on the $\mu_{\mathrm{t\bar{t}H}}$ estimation at 3000 fb^{-1}.** Shown are the 20 parameters with the largest projected impact on $\mu_{\mathrm{t\bar{t}H}}$ at 3000 fb^{-1} in scenario S1 (top) and S2 (bottom), respectively. All impacts were calculated with pseudo data generated from the sum of the signal and background predictions.

for the b-tagging parameters. This enables the fit to place stronger constraints on these uncertainties, which in turn limits the impact on the signal inference. This impact ranking clearly indicates which sources of uncertainty are in need of improvements beyond the estimation used in this projection. On the theory side, the uncertainties on the inclusive $t\bar{t}H$ cross section at NLO and the next-to-next-to-leading-logarithmic (NNLL) parton shower uncertainties for the $t\bar{t}$ + jets prediction are among the most important parameters. Apart from calculations at higher orders, especially the latter uncertainties will benefit from more sophisticated simulation techniques, such as more precise calculations of the inclusive $t\bar{t}$ pair production or dedicated predictions for the $t\bar{t}$ + jets processes. Such improvements are expected to yield a better description of the shape in the final observable for these processes, which cannot be fully accounted for in the scope of this projections. On the experimental side, the uncertainties on the b-tagging procedure for c-jets and the JES are expected to have the largest impact on the $\mu_{t\bar{t}H}$ measurement. While these uncertainties will certainly benefit from the improved detector performance in the future, it is crucial to develop new techniques to reduce these uncertainties.

Overall, all studies point to a high sensitivity of the future $t\bar{t}H(b\bar{b})$ analysis. This will open new avenues for more sophisticated measurements, such as differential measurements. Moreover, the increased statistical precision will allow the simultaneous measurement of the signal and important background processes. This was already started in the full Run-II $t\bar{t}H(b\bar{b})$ analysis described in Part II and will give valuable input for future calculations of these processes. All of this will allow for more stringent tests of theory predictions from the Standard Model or future models that have yet to be defined.

As mentioned at the beginning of the chapter, the basic assumption of these projection scenarios is that the analysis strategy of the measurement does not change. However, as new techniques and technologies are developed over time, the analysis strategy is continuously updated to exploit the resulting new possibilities. In fact, many of the improvements in the $t\bar{t}H(b\bar{b})$ analysis strategy have been developed in the scope of this thesis, as was already discussed in Part II. The following section briefly summarizes the improvements between the analysis of the data recorded in 2016 and the full Run-II analysis and compares the current sensitivity with the sensitivity projected in the course of this study.

11.3 Comparison to the full Run-II analysis

The main differences between the setup used in the analysis of the data recorded in 2016 and the full Run-II analysis setup are listed in Tab. 11.4. First, the definition of the analysis phase space changed. The $t\bar{t}H(b\bar{b})$ analysis of the 2016 data set described in Ref. [133], which is the base for the projection studies, only considered events where the $t\bar{t}$ system decays in at least one lepton. In the full Run-II analysis, all $t\bar{t}$ decay channels are used. The consideration of all $t\bar{t}$ decay modes increases the sensitivity of the full Run-II analysis by approximately 10 %.

The differences in the baseline selections mainly arise from the improved b-tagging algorithm. After the projection studies were finished, the algorithm was updated twice. During each update, the change in the selection efficiency was quantified using the ratio of signal events over the square root of background events ($S/\sqrt{B}$), which quantifies the signal yield in terms of the expected statistical fluctuations of the background processes and can thus be used as a proxy for the sensitivity of the final categorization. The projection baseline analysis employed the combined secondary vertex (CSVv2) algorithm [267], which was based on a neural-network approach. This technique was superseded by its successor DeepCSV [268], which used more information about the tracks of charged particles and included more layers in its architecture than its predecessor. This improved the $S/\sqrt{B}$

Table 11.4: **Differences between the projection baseline and the full Run-II analysis**. The table summarizes the main differences between the initial analysis for the projection and the full Run-II analysis that have a potential impact on the sensitivity to the $t\bar{t}H(b\bar{b})$ signal.

Topic	Baseline for Projections	Full Run-II Analysis
Input Channels for the Analysis	Dileptonic and Semileptonic Channels	Dileptonic, Semileptonic and Full-Hadronic Channels
Categorization	4 jets, 5 jets, $\geq$ 6 jets	($\geq$ 4 jets, 3 b-tags), ($\geq$ 4 jets, $\geq$ 4 b-tags)
b-tagging Algorithm	Combined Secondary Vertex (BDT) [267]	DeepJet (DNN) [120, 202]
Simulated Events	Old underlying event tune	Updated underlying event tune
$t\bar{t} + b\bar{b}$ Simulation	Part of inclusive $t\bar{t}$ NLO simulation (first add. b-jet at ME)	Dedicated NLO $t\bar{t} + b\bar{b}$ simulation (both add. b-jets at ME)
Network Training Sample	Dedicated 2016 simulation	Combination of simulations for 2016, 2017 and 2018
Network Architecture per Category	Up to four Hidden Layers with 100 nodes each	Four Hidden Layer with up to 2048 nodes each
Network Classification	Classification into $t\bar{t}H(b\bar{b})$, $t\bar{t} + 2b$, $t\bar{t} + b$, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$, $t\bar{t} + lf$	Classification into $t\bar{t}H(b\bar{b})$, $t\bar{t} + 2b$, $t\bar{t}b\bar{b}$/b, $t\bar{t} + c\bar{c}$, $t\bar{t} + lf$, tHq, tHW
Uncertainty Definition	Latest uncertainty definition in 2017	Latest uncertainty definition in 2020
Fitted Observables	Distributions of ANN outputs	Distributions of ANN outputs, ratio observable
Treatment of Background Normalization uncertainty	a-priori uncertainty of 50 % for $t\bar{t} + b\bar{b}$, $t\bar{t} + 2b$, $t\bar{t} + b$ and $t\bar{t} + c\bar{c}$, respectively	Freely-floating parameters for $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$, additional 50 % uncertainty for $t\bar{t} + 2b$

in the signal-enriched category by approximately 18 % [21]. Afterwards, the b-tagging algorithm was updated to the latest current standard in the CMS experiment, which is the DeepJet tagger [120]. This improved the $S/\sqrt{B}$ again by approximately 30 % in the signal-enriched categories [137]. This improvement in sensitivity allowed for a different categorization, which is based on the number of b-tagged jets instead of the number of jets. This increased the number of events per category without significantly losing sensitivity to the $t\bar{t}H(b\bar{b})$ signal process, which has proven to stabilize the uncertainty estimation.

Furthermore, the accuracy of the $t\bar{t} + b\bar{b}$ background process simulation was improved. In the 2016 analysis, the $t\bar{t} + b\bar{b}$ description was based on an inclusive NLO $t\bar{t}$ simulation in the five-flavor scheme introduced in Sec. 2.3. While the $t\bar{t}$ pair production is modeled very precisely in this sample, only the first additional b-jet is part of the matrix element (ME) calculation. All other additional jets originate from the parton shower (PS), which means that the corresponding observables like the p_T of the jets had only NNLL accuracy. As already introduced in Sec. 5.1.2, the full Run-II analysis relies on a dedicated NLO $t\bar{t} + b\bar{b}$ sample in the four-flavor scheme, where the additional b-jets are both generated directly at matrix element (ME) level. This simulation is expected to describe the kinematic properties

better than the NNLL simulation. This change in the kinematic properties leads to a higher $t\bar{t} + b\bar{b}$ contribution in the signal-enriched categories, thus degrading the purity of the final observables. This effect leads to a loss in sensitivity of approximately 10 % [137].

Additionally, the neural network architecture and training paradigm was updated. As described in Part II, the full Run-II analysis considers dedicated simulated samples for each year, respectively. As previously mentioned, the total number of events available for training and the stability of the training are improved by using events from the simulated samples for all three years simultaneously. The increased number of training events also allowed for a more complicated network architecture. In the end, the signal efficiency in the signal-enriched category increases by approximately 3 %. Furthermore, the selection efficiency of the background processes is improved by up to 130 %. This in turn helps to constrain the uncertainties on the backgrounds in the analysis, which helps improving the final signal inference.

Finally, the statistical model itself was updated for the full Run-II analysis. In general, both analyses used the latest recommendations from the dedicated expert groups within the CMS collaboration at the time, which are updated regularly. The main differences relevant for the $t\bar{t}H(b\bar{b})$ measurement impact the following groups of uncertainties:

Sig. Theory: The full Run-II analysis additionally considers ISR, FSR, μ_R and μ_F uncertainties for the $t\bar{t}H$ processes, which were previously not part of the statistical model.

JES Uncertainties: The recommendation was updated by the corresponding group of experts, reducing the number of uncertainty sources from 26 to eleven.

$t\bar{t}$ + heavy flavor Modeling: The parameterization of the $t\bar{t}$ + heavy flavor processes was changed in order to fully exploit the improved $t\bar{t} + b\bar{b}$ simulation. Most importantly, the a priori uncertainty on the additional $t\bar{t}$ + heavy flavor cross sections was removed and the normalizations are inferred directly from the data. This means that the additional $t\bar{t}$ + heavy flavor cross section parameters are part of the statistical uncertainty in the full Run-II fit model, which was already discussed in Sec. 9.3

Additionally, the full Run-II analysis exploits a dedicated likelihood ratio observable, which further increases the sensitivity of the analysis by approximately 30 % [137, 260] and was introduced in Sec. 8.2.

Figure 11.6 compares the significance projected by the two scenarios S1 and S2 with the values obtained in the course of the full Run-II analysis. For every luminosity value, the (expected) full Run-II significance exceeds both the conservative S1 and the optimistic S2 scenario. Consequently, the improvements in the analysis strategy that were discussed previously had an impact greater than what was conceivable at the time of the analysis of the 2016 data set. This is especially true when considering the difficulties that were introduced, such as the higher contribution of the irreducible background in the analysis phase space or the simultaneous fit of the signal and the largest backgrounds. These changes were only possible by the improvements in the discriminators and uncertainty model, which were rewarded with a much better sensitivity.

Table 11.5 compares the contributions of the uncertainty sources in the optimistic S2 scenario with the full Run-II analysis. In the following discussion, it is important to note that the projected luminosity for the end of Run-II is lower than the true luminosity recorded at the CMS experiment. Therefore, the improvement visible in Tab. 11.5 stems in part from the difference in the number of events. However, there are also features that are independent of the improved statistical precision.

The statistical uncertainty in the full Run-II analysis is slightly smaller than the projected value at $100\,\mathrm{fb}^{-1}$. However, when considering the increased statistical precision using

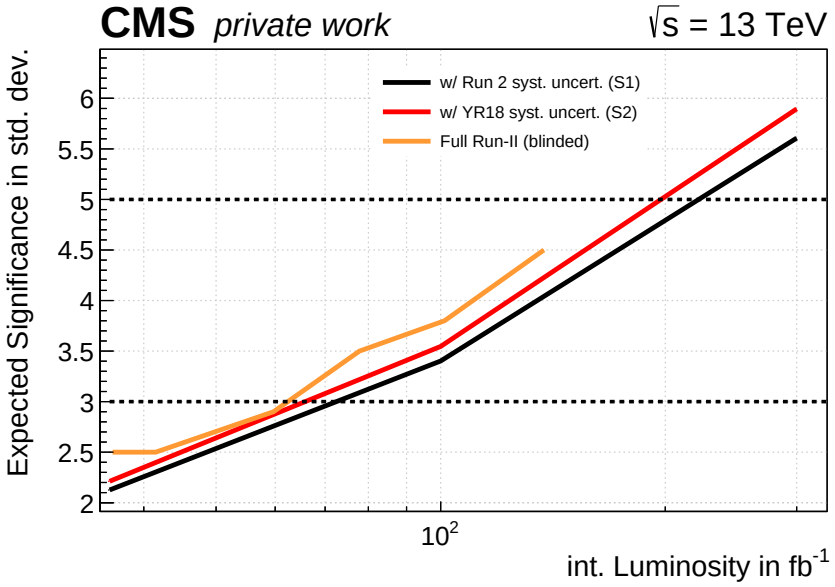


Figure 11.6: **Comparison of the projected significance in the $t\bar{t}H(b\bar{b})$ measurement and the sensitivity of the full Run-II analysis**. Shown are the projected significances for scenarios S1 and S2 as well as the values obtained in the scope of the Full Run-II analysis, respectively. The dashed lines indicate the thresholds for evidence ($3\,\sigma$) and observation ($5\,\sigma$).

Table 11.5: **Comparison of the uncertainty contributions to the total precision in the $\mu_{t\bar{t}H}$ estimation**. Shown are the contributions of the individual sources of uncertainty to the sensitivity. All values are given in percent. The table compares projected values in scenario S2 at $100\,\text{fb}^{-1}$ with the values obtained in the full Run-II analysis. For better comparability, the projected values were scaled by $\sqrt{100\,\text{fb}^{-1}/137.5\,\text{fb}^{-1}}$, which is the naively expected gain in sensitivity due to improved statistical precision without consideration of correlations. The values were obtained from pseudo data containing the expected signal and background contributions.

Source	S2 @ $100\,\text{fb}^{-1}$	Scaled S2 @ $137.5\,\text{fb}^{-1}$	Full Run-II Analysis (blinded)
Total	28.8	24.6	24.0
Stat	16.0	13.6	15.3
Sig. Theory	4.7	4.0	8.9
add. $t\bar{t}$ + heavy flavor XS†	8.6	7.3	11.5
Bkg. Theory	16.0	13.6	9.3
Exp	10.8	9.2	6.7
JES	7.4	6.3	3.9
Luminosity	1.0	0.9	1.6
B tagging	7.0	6.0	4.8
MC uncertainty	–	–	8.9

†: Parameterization in fits different

the naive scaling of the uncertainty, the statistical uncertainty in the full Run-II analysis seemingly deteriorated with respect to the value expected within the projection scenario S2. This is expected due to the change in the parameterization of the analysis. Since the $t\bar{t}$ + heavy flavor normalization parameters are part of the statistical uncertainty in the full Run-II analysis, the contribution of this source of uncertainty to the total uncertainty increased. In fact, given the large contribution of the $t\bar{t}$ + heavy flavor normalization uncertainties, it is remarkable that the difference is not more pronounced. This indicates that the discrimination power of the final observables has increased, which is the expected outcome from the improved training of the Artificial Neural Network (ANN) as well as the newly-introduced ratio observable used in the final fit.

The contribution of the signal theory uncertainties is larger than the expected value in the projection scenario. This is due to the assumption in the S2 scenario that new calculations for the $t\bar{t}H$ process will half the uncertainties. However, there have been no improvements in the accuracy of the predictions for the $t\bar{t}H$ process since the projection study was published. Therefore, it is expected that the contribution of the signal theory uncertainties in the full Run-II analysis are approximately twice as large as the values predicted by scenario S2, which is indeed the case.

The contribution of the additional $t\bar{t}$ + heavy flavor normalization uncertainties is slightly increased in the full Run-II analysis with respect to the projected value. Much like with the signal theory uncertainties, the a priori uncertainties on these parameters were halved in scenario S2. However, the precision of the cross section calculations stayed the same. Therefore, one would naively expect a factor of approximately two between the projected value and the contribution in the full Run-II uncertainty. Additionally, the a priori uncertainties on these parameters were removed in the full Run-II analysis. As shown in Sec. 11.2, the exact a priori uncertainty on the $t\bar{t}$ + heavy flavor normalization parameters has only a small effect on the total expected uncertainty on $\mu_{t\bar{t}H}$. Consequently, one would expect to see a small additional effect originating from the removal of the a priori uncertainty. However, this is not the case. The contribution of the $t\bar{t}$ + heavy flavor normalization parameters is approximately 30 % larger than what is projected in the S2 scenario, which is well below the expectation. This again indicates the improved discrimination of signal and background processes, which yields a more powerful background estimation.

The contribution of the background theory uncertainties is larger in the S2 projection scenario than in the full Run-II analysis. This is in part due to the fact that the additional $t\bar{t}$ + heavy flavor cross section uncertainties are part of the background theory group in the projection study while they are moved to the statistical uncertainties in the full Run-II measurement. However, even after subtracting the contribution of the $t\bar{t}$ + heavy flavor parameters, the contribution of the background uncertainty group in the projection study is larger than what is seen in the full Run-II analysis. Additionally, the a priori uncertainties in this group of uncertainties were halved in the S2 scenario, which means that one would naively expect that the projected contribution is much smaller than the value measured in the full Run-II analysis. This improvement must therefore originate from the improved sensitivity of the discriminators and the changes in the uncertainty parameterization, which result in tighter constraints on the uncertainties and thus limit their impact on the signal inference.

The contribution of the experimental uncertainties is much smaller in the full Run-II analysis than in the projection scenario. This effect originates mainly from the JES and b-tagging uncertainties. The full Run-II analysis deploys an improved paramterization of the JES uncertainty sources that gain sensitivity by combining the different years. The b-tagging uncertainties profit from the improved detector performance due to upgrades

such as the new pixel detector that was implemented in 2017 explained in Sec. 3.2.2. Additionally, the b-tagging efficiency was improved by deploying powerful algorithms such as the DeepJet tagger. These improvements exceed the originally estimated upgrades and thus reduce the contribution of these uncertainty sources to the total uncertainty on $\mu_{t\bar{t}H}$. The contribution of the luminosity uncertainty to the total precision is increased in the full Run-II analysis. The parameterization is based on new measurements of the luminosity recorded at the CMS experiment and yields the most realistic uncertainties to date. However, they are still far from the estimated precision of the luminosity at the HL-LHC, which is used in the projection scenario S2. As discussed in Sec. 11.2, there is still room for future improvements here, from which the $t\bar{t}H(b\bar{b})$ analysis will benefit.

Finally, the relative contribution of the uncertainties due to the limited number of simulated events to the total uncertainty on $\mu_{t\bar{t}H}$ is approximately 40 %. Consequently, this group of uncertainties has one of the largest contributions to the total uncertainty. Therefore, the assumption that the uncertainties on the size of the simulated samples will be negligible in future analyses is not fulfilled in the full Run-II analysis, which is a shortcoming of the projection studies.

Before finally evaluating the validity of the projection studies, it is important to compare the impacts of the individual parameters on the $t\bar{t}H(b\bar{b})$ measurement. Figure 11.7 shows the 20 parameters that are expected to have the largest impacts on the $\mu_{t\bar{t}H}$ estimation for the projection scenario S2 and the full Run-II analysis, respectively. Overall, the underlying sources of uncertainties are similar. The cross section normalization uncertainties for $t\bar{t} + b\bar{b}$ and $t\bar{t} + 2b$ as well as the cross section uncertainty on $t\bar{t}H$ due to a variation of the strong coupling α_s in the μ_R and μ_F scales and the PDF set are ranked high in both the projection scenario and the full Run-II analysis. Additionally, the full Run-II impact lists many parameters related to the background estimation in the fully-hadronic input channel. Since this channel was not considered in the projection, these parameters cannot be compared.

In both rankings, theory uncertainties due to the renormalization and factorization scales, as well as initial state and final state radiation are among the parameters with the highest impacts on the $t\bar{t}H(b\bar{b})$ measurement. However, the exact sources of uncertainties are slightly different. While the projected impact ranking is mainly dominated by $t\bar{t}$ + lf-related parameters, the full Run-II ranking shows many uncertainties related to the $t\bar{t} + b\bar{b}$ process. This is due to the changes in the $t\bar{t} + b\bar{b}$ simulation, which is expected to model this background process more accurately. However, the uncertainties of this simulation are larger, particularly the shape uncertainty arising from the renormalization scale μ_R in the calculation of the matrix element. Since the kinematic properties of the $t\bar{t} + b\bar{b}$ background process, and thus the shape in the final discriminator, is very similar to the signal process, an increase in the a priori uncertainty can have a larger effect on the signal strength estimation. Additionally, as shown before, the discriminators in the full Run-II analysis are more sensitive to both the $t\bar{t}H(b\bar{b})$ signal and the $t\bar{t} + b\bar{b}$ background. Consequently, these parameters are ranked higher in the full Run-II analysis than in the projection.

The ranking of the projected impacts shows several experimental uncertainties, mostly related to the b-tagging and the JES parameterization. This is not the case in the full Run-II impact ranking, which shows two b-tagging parameters and three JES uncertainties. This indicates again the improvements in the parametrization of the experimental uncertainties and the improved b-tagging efficiency. These improvements are beyond the underlying expectations of the projection studies, which is why the ranking between the cases is different.

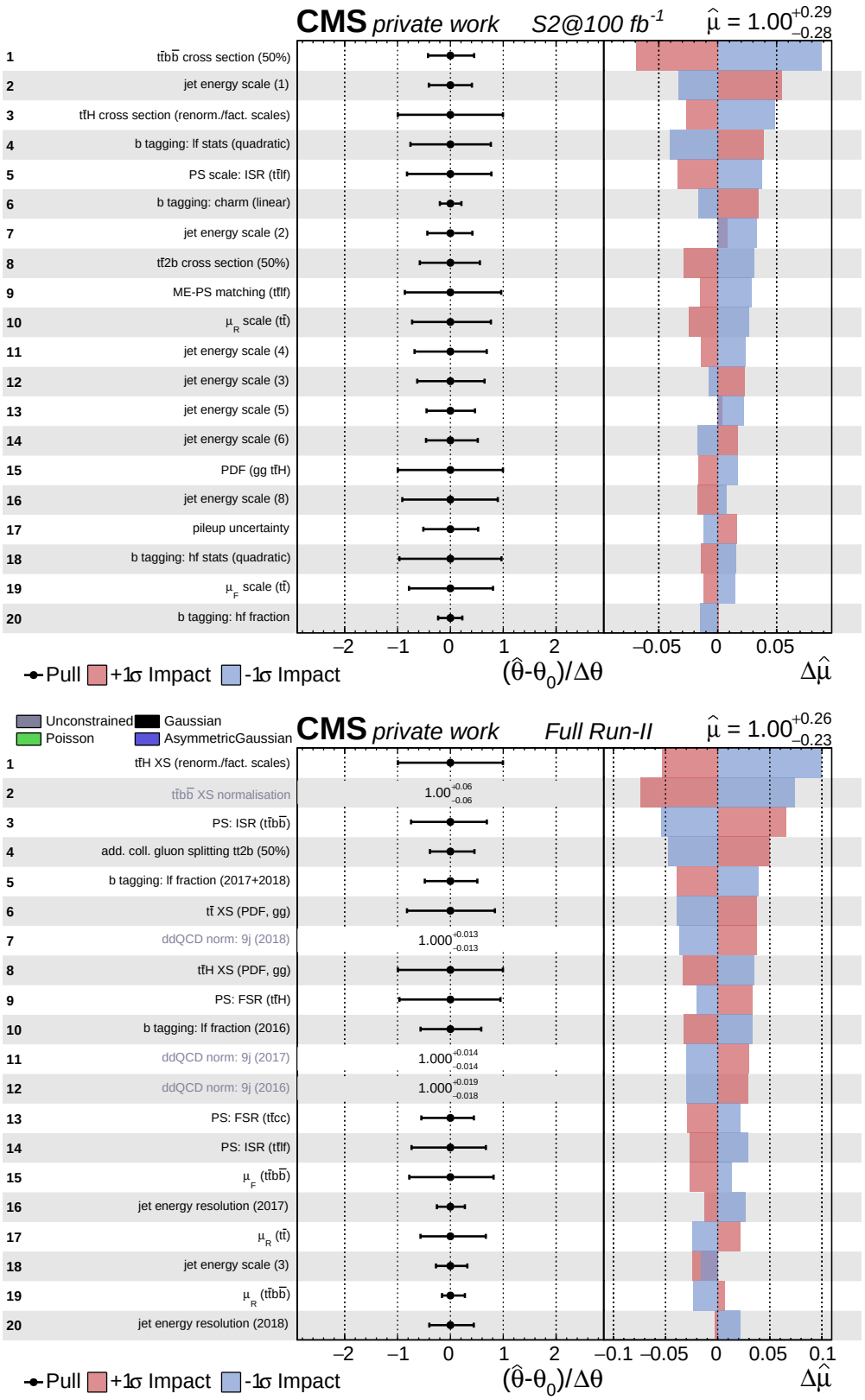


Figure 11.7: **Comparison of the parameter impacts between the projection study and the full Run-II analysis**. Shown are the 20 parameters that are expected to have the largest impact on the signal-strength estimation for the YR systematic uncertainties (S2) scenario (top) and the full Run-II analysis (bottom). The impacts are calculated using pseudo data corresponding to the Standard Model predictions for the signal and background processes corresponding to the respective integrated luminosities.

Overall, the results predicted in the optimistic S2 scenario are close to the full Run-II results. However, the improvements in the analysis strategy yield a higher sensitivity that goes beyond a mere improvement of the statistical sensitivity. The more powerful discriminants, the improved techniques and algorithms and the newest uncertainty parameterization culminate in a sensitivity that exceeds the expectations. This will lay the foundation for future improvements, which will surely go beyond what is conceivable today. These improvements will open the way to even more sophisticated tests of the Standard Model that will give us even more insights into the fundamental mechanisms of particle physics.

Part IV

Conclusions

12 Conclusion

In the absence of significant deviations between observed data and the predictions made by the Standard Model of particle physics (SM), the search for New Physics is still ongoing. The observation of a particle with properties compatible with the predicted Higgs boson in 2012 marks the start of a large-scale campaign for probing the Higgs sector to construct stringent tests of the SM and contributions from physics beyond (BSM).

This thesis presented the analysis of the $t\bar{t}H(b\bar{b})$ channel in the semileptonic final state using the full Run-II data set recorded at the CMS experiment. It was shown that the $t\bar{t}H(b\bar{b})$ analysis is of special interest due to its direct access to the top-Higgs Yukawa coupling and the possible sensitivity to BSM effects. Due to the large contribution of the $t\bar{t}+$ jets processes and the irreducible $t\bar{t}+b\bar{b}$ process in particular, a classification based on the output of multiclass Artificial Neural Networks (ANNs) was chosen as analysis strategy. These networks were trained with observables describing the underlying physics obtained from simulated samples and enhance the sensitivity to the $t\bar{t}H(b\bar{b})$ signal.

It was shown that the statistical model used for the inclusive $t\bar{t}H(b\bar{b})$ measurement is capable of describing the observed data well. Furthermore, the expected and observed estimations for the systematic uncertainties have been interpreted and discussed. The observed values are well-compatible with observations made in previous measurements in the $t\bar{t}H(b\bar{b})$ channel as well as other analyses. The parameters with the largest impact on the signal strength estimation are related to the modeling of the $t\bar{t}+$ jets backgrounds and particularly to the $t\bar{t}+b\bar{b}$ process, as well as uncertainties related to the b-tagging method and the limited size of the used simulated samples, which is compatible with the expectation. At the time of writing this thesis, the final estimation of the signal strength modifier is still under revision. Therefore, the expected sensitivity to the $t\bar{t}H(b\bar{b})$ signal evaluated with simulated data with a signal strength of one was discussed. The expected sensitivity in the semileptonic channel is $+0.34/-0.30$, which corresponds to an expected significance of $3.3\,\sigma$ compared to the background-only hypothesis.

Furthermore, the contribution of the semileptonic channel to the analysis presented in Ref. [137] has been discussed. It was shown that the semileptonic analysis is the driving factor in the combined $t\bar{t}H(b\bar{b})$ measurement, which also profits from the contributions from the dileptonic and the fully-hadronic channels. The expected sensitivity in this analysis is compatible with the values obtained in previous $t\bar{t}H(b\bar{b})$ measurements at the CMS experiment when the corresponding data sets recorded in 2016 and 2017 are evaluated.

The unprecedented amount of data allowed for the first $t\bar{t}H$ cross section measurement as a function of the transverse momentum of the Higgs boson p_T^H in the context of the Simplified Template Cross Section (STXS) framework in the $t\bar{t}H(b\bar{b})$ channel at the CMS experiment. Similar to the inclusive $t\bar{t}H(b\bar{b})$ measurement, the cross section estimation in the low-p_T^H region is sensitive to systematic uncertainties related to the $t\bar{t}+$ jets modeling and experimental uncertainties. Contrary to this, the uncertainties with the largest impact on the cross section measurement in the high-p_T^H region are mainly related to the limited size of the simulated samples and the modeling of the $t\bar{t}+b\bar{b}$ process. The expected sensitivities of the cross section measurement using all $t\bar{t}H(b\bar{b})$ channels range between 56–107 % and are mostly dominated by statistical uncertainties.

Additionally, the first feasibility study of the potential sensitivity to effects described in the Standard Model Effective Field Theory (SMEFT) was discussed. The study compared the expected contributions of the three dimension-six operators at NLO QCD in the scope of the currently available constraints presented in Ref. [69] to the sensitivity of the cross section measurement in the $t\bar{t}H(b\bar{b})$ channel. The study indicated that the $t\bar{t}H(b\bar{b})$ analysis with the currently available sensitivity could potentially add constraining power to estimations in the marginalized scheme. This is especially true for effects originating from the monochromatic dipole moment $\mathcal{O}_{tG}$ and the modified top-Higgs Yukawa coupling $\mathcal{O}_{t\varphi}$ in the high-p_T^H regions, where these operators yield an effect of approximately 120 %, corresponding to an excess of little over one standard deviation with respect to the SM prediction and the sensitivity for this particular STXS bin. The large sensitivity to these operators in particular is in good agreement with the results presented in Ref. [18].

Finally, the projected sensitivity used for the formulation of the European strategy for particle physics presented in Ref. [33] was discussed in detail. These projections predict a sensitivity of 7 % (11 %) in an optimistic (conservative) scenario for a data set of 3000 fb^{-1}, which corresponds to the expected amount of data at the end of the High Luminosity Large Hadron Collider (HL-LHC). Similar to the observations made in the analysis of the full Run-II data set, the systematic uncertainties that are expected to have the largest impact on the signal strength estimation are related to the modeling of the signal process and the $t\bar{t}+$ jets background processes. The sensitivity predicted in the scope of the projection studies for the then-expected Run-II data set was compared to the sensitivity obtained in the real full Run-II analysis. The expected sensitivity exceeds the projected value in the optimistic scenario by approximately 7 %, which is attributed to the improvements in the analysis strategy. This indicates that future analyses might reach sensitivities that are lower than the projected 7 %, which will enable even more sophisticated measurements in the $t\bar{t}H(b\bar{b})$ channels.

In summary, future $t\bar{t}H(b\bar{b})$ analyses will have the opportunity to perform measurements with increasingly differential information and will be able to probe the available phase space in far more detail. Here, the high-p_T^H regime in particular shows high sensitivity to interesting new physics prospects, which are already available for comparison using tools such as the SMEFT framework. However, such measurements involve challenges that are yet to be addressed, such as the full description of all contributing operators in the simulation, the arising new interference effects with background processes and the construction of suitable statistical models for the final inference of the higher-dimensional contributions. Furthermore, the studies show that the description of the $t\bar{t}+$ jets and the $t\bar{t}+b\bar{b}$ process in particular is crucial for the $t\bar{t}H(b\bar{b})$ measurement. In order to improve the sensitivity, a better control and description of these leading background processes is needed, which necessitates more dedicated studies such as the analysis presented in Ref. [25]. Overall, the physics perspective in the $t\bar{t}H(b\bar{b})$ channel will become even more interesting with an increased amount of data, and is sure to inspire new measurements.

Bibliography

[1] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Physical Review Letters* 13 (9 Aug. 31, 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[2] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Physical Review Letters* 13 (16 Oct. 19, 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[3] P. W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Physical Review* 145.4 (May 27, 1966), pp. 1156–1163. DOI: 10.1103/physrev.145.1156.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Physical Review Letters* 13 (20 Nov. 16, 2016), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[5] P. W. Anderson. "Plasmons, Gauge Invariance, and Mass". In: *Physical Review* 130 (1 Apr. 1, 1963), pp. 439–442. DOI: 10.1103/PhysRev.130.439. URL: https://link.aps.org/doi/10.1103/PhysRev.130.439.

[6] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (Jan. 29, 2013), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.

[7] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716 (1 July 31, 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214v2 [hep-ex].

[8] B. Hoeneisen. "A Study of Dark Matter with Spiral Galaxy Rotation Curves". In: *International Journal of Astronomy and Astrophysics* 09.02 (June 2019), pp. 71–96. DOI: 10.4236/ijaa.2019.92007.

[9] D. Paraficz et al. "The Bullet cluster at its best: weighing stars, gas, and dark matter". In: *Astronomy & Astrophysics* 594.A121 (Oct. 24, 2016), p. 14. DOI: 10.1051/0004-6361/201527959.

[10] A. Mazumdar, S. Mohanty, and P. Parashari. "Evidence of dark energy in different cosmological observations". In: *The European Physical Journal Special Topics* 230.9 (July 8, 2021), pp. 2055–2066. DOI: 10.1140/epjs/s11734-021-00212-y.

[11] G. Degrassi et al. "Higgs mass and vacuum stability in the Standard Model at NNLO". In: *Journal of High Energy Physics* 98 (May 29, 2012). DOI: 10.1007/JHEP08(2012)098. arXiv: 1205.6497 [hep-ph].

[12] F. Bezrukov and M. Shaposhnikov. "Why should we care about the top quark Yukawa coupling?" In: *Journal of Experimental and Theoretical Physics* 120 (Apr. 15, 2015), pp. 335–343. DOI: 10.1134/S1063776115030152. eprint: arXiv:1411.1923. URL: https://cds.cern.ch/record/1968356.

[13] S. P. Martin. "A Supersymmetry Primer". In: *Perspectives on Supersymmetry* (1998), pp. 1–98. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356 [hep-ph].

[14] R. Contino. *Tasi 2009 lectures: The Higgs as a Composite Nambu-Goldstone Boson.* Tech. rep. May 24, 2010. arXiv: 1005.4269 [hep-ph].

[15] C. Csáki and P. Tanedo. "Beyond the Standard Model". In: *2013 European School of High-Energy Physics, Paradfurdo, Hungary.* Feb. 12, 2016, pp. 169–268. DOI: 10.5170/CERN-2015-004.169. arXiv: 1602.04228 [hep-ph].

[16] G. Li et al. "Resolving the degeneracy in top quark Yukawa coupling with Higgs pair production". In: *Physics Letters B* 800 (Jan. 10, 2020), p. 135070. DOI: 10.1016/j.physletb.2019.135070.

[17] F. Maltoni, E. Vryonidou, and C. Zhang. "Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD". In: *Journal of High Energy Physics* 2016.123 (Oct. 24, 2016). DOI: 10.1007/JHEP10(2016)123. arXiv: 1607.05330 [hep-ph].

[18] M. Battaglia et al. *Sensitivity to BSM effects in the Higgs p_T spectrum within SMEFT.* Tech. rep. Sept. 7, 2021. arXiv: 2109.02987 [hep-ph].

[19] LHC Higgs Cross Section Working Group. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector.* en. 2017. DOI: 10.23731/CYRM-2017-002.

[20] K. El Morabit. "Measurement of ttH(H→bb) production in the semi- leptonic tt decay channel at the CMS Experiment". PhD thesis. Karlsruhe Institute of Technology (KIT), 2021. DOI: 10.5445/IR/1000135447.

[21] CMS Collaboration. *Measurement of* $\mathrm{t\bar{t}H}$ *production in the* $\mathrm{H} \to \mathrm{b\bar{b}}$ *decay channel in* $41.5\,\mathrm{fb}^{-1}$ *of proton-proton collision data at* $\sqrt{s} = 13\,\mathrm{TeV}$. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2675023.

[22] CMS Collaboration. "Observation of $t\bar{t}H$ Production". In: *Physical Review Letters* 120.231801 (23 June 4, 2018). DOI: 10.1103/PhysRevLett.120.231801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.231801.

[23] CMS Collaboration. "Observation of Higgs Boson Decay to Bottom Quarks". In: *Physical Review Letters* 121.121801 (12 Sept. 17, 2018). DOI: 10.1103/PhysRevLett.121.121801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.121801.

[24] CMS Collaboration. "Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Physics Letters B* 776 (2018), pp. 355–378. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2017.11.043.

[25] A. Saibel. "Phenomenology of ttH Production with Top Quark Running Mass and the Differential Cross-Section Measurement of tt+b-jets Production in the Dilepton Channel at $\sqrt{s} = 13\,\mathrm{TeV}$ with the CMS Experiment". PhD thesis. Universität Hamburg, 2021. URL: https://ediss.sub.uni-hamburg.de/handle/ediss/9155.

[26] S. Willenbrock and C. Zhang. "Effective Field Theory Beyond the Standard Model". In: *Annual Review of Nuclear and Particle Science* 64.1 (2014), pp. 83–100. DOI: 10.1146/annurev-nucl-102313-025623. eprint: https://doi.org/10.1146/annurev-nucl-102313-025623. URL: https://doi.org/10.1146/annurev-nucl-102313-025623.

[27] A. Falkowski. "Effective field theory approach to LHC Higgs data". In: *Pramana* 87.39 (Aug. 23, 2016). DOI: 10.1007/s12043-016-1251-5.

[28] I. Brivio and M. Trott. "The standard model as an effective field theory". In: *Physics Reports* 793 (2019). The standard model as an effective field theory, pp. 1–98. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2018.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0370157318303223.

[29] M. Delmastro et al. *Simplified Template Cross Sections – Stage 1.1*. Tech. rep. 14 pages, 3 figures. Geneva: CERN, Apr. 2019. arXiv: 1906.02754. URL: https://cds.cern.ch/record/2669925.

[30] LHC Higgs Cross Section Working Group. *LHCHWG Fiducial And STXS*. TWiki: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGFiducialAndSTXS r24.

[31] G. Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports: Monographs. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-004. URL: https://cds.cern.ch/record/2284929.

[32] CERN. *Project Schedule for the LHC/HL-LHC Plan*. Nov. 1, 2021. URL: https://project-hl-lhc-industry.web.cern.ch/content/project-schedule (visited on 11/01/2021).

[33] A. Dainese et al. *Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC*. Tech. rep. Geneva, Switzerland, 2019. DOI: 10.23731/CYRM-2019-007. URL: https://cds.cern.ch/record/2703572.

[34] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Dec. 2013. DOI: 10.1017/9781139540940.

[35] M. E. Peskin. *An Introduction To Quantum Field Theory*. CRC Press, May 2018. DOI: 10.1201/9780429503559.

[36] D. Griffiths. *Introduction to elementary particles*. Weinheim: Wiley-VCH, 2004. ISBN: 978-3-527-61847-7.

[37] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[38] E. Fermi. "Zur Quantelung des idealen einatomigen Gases". In: *Zeitschrift für Physik* 36.11 (Nov. 1926), pp. 902–912. ISSN: 0044-3328. DOI: 10.1007/BF01400221.

[39] B. T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". In: *The Astrophysical Journal* 496.1 (Mar. 1998), pp. 505–526. DOI: 10.1086/305343.

[40] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81 (8 Aug. 24, 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[41] K. Zuber. "The discovery of neutrino oscillations". In: *Annalen der Physik* 528.6 (2016), pp. 452–457. DOI: https://doi.org/10.1002/andp.201600097. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.201600097. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.201600097.

[42] S. Fukuda et al. "Constraints on Neutrino Oscillations Using 1258 Days of Super-Kamiokande Solar Neutrino Data". In: *Physical Review Letters* 86 (25 June 18, 2001), pp. 5656–5660. DOI: 10.1103/PhysRevLett.86.5656. URL: https://link.aps.org/doi/10.1103/PhysRevLett.86.5656.

[43] D. V. Forero, M. Tórtola, and J. W. F. Valle. "Neutrino oscillations refitted". In: *Physical Review D* 90.093006 (9 Nov. 20, 2014). DOI: 10.1103/PhysRevD.90.093006. URL: https://link.aps.org/doi/10.1103/PhysRevD.90.093006.

[44] M. Sanchez et al. *Observation of atmospheric neutrino oscillations in Soudan 2.* Tech. rep. Aug. 2003. DOI: 10.2172/815661. URL: https://www.osti.gov/biblio/815661.

[45] CMS Collaboration et al. "A measurement of the Higgs boson mass in the diphoton decay channel". In: *Physics Letters B* 805 (June 10, 2020). DOI: 10.1016/j.physletb.2020.135425.

[46] S. Bose. "Planck's law and the light quantum hypothesis". In: *Journal of Astrophysics and Astronomy* 15.1 (Mar. 1994), pp. 3–7. ISSN: 0973-7758. DOI: 10.1007/BF03010400.

[47] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Physical Review Letters* 30.26 (June 25, 1973), pp. 1343–1346. DOI: 10.1103/physrevlett.30.1343.

[48] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Physical Review Letters* 30.26 (June 25, 1973), pp. 1346–1349. DOI: 10.1103/physrevlett.30.1346.

[49] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Physical Review* 96.1 (Oct. 1, 1954), pp. 191–195. DOI: 10.1103/physrev.96.191.

[50] D. J. Gross and F. Wilczek. "Asymptotically Free Gauge Theories. I". In: *Physical Review D* 8 (10 Nov. 15, 1973), pp. 3633–3652. DOI: 10.1103/PhysRevD.8.3633. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.3633.

[51] D. J. Gross and F. Wilczek. "Asymptotically free gauge theories. II". In: *Physical Review D* 9 (4 Feb. 15, 1974), pp. 980–993. DOI: 10.1103/PhysRevD.9.980. URL: https://link.aps.org/doi/10.1103/PhysRevD.9.980.

[52] A. Salam and J. C. Ward. "Electromagnetic and Weak Interactions". In: *Selected Papers of Abdus Salam*. WORLD SCIENTIFIC, May 1994, pp. 210–213. DOI: 10.1142/9789812795915_0029.

[53] S. Weinberg. "A Model of Leptons". In: *Physical Review Letters* 19.21 (Nov. 20, 1967), pp. 1264–1266. DOI: 10.1103/physrevlett.19.1264.

[54] S. L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (Sept. 9, 1969), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[55] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Physical Review Letters* 10.12 (June 15, 1963), pp. 531–533. DOI: 10.1103/physrevlett.10.531.

[56] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1, 1973), pp. 652–657. DOI: 10.1143/ptp.49.652.

[57] J. Goldstone. "Field theories with Superconductor solutions". In: *Il Nuovo Cimento (1955-1965)* 19.1 (Sept. 8, 1960), pp. 154–164. ISSN: 1827-6121. DOI: 10.1007/BF02812722. URL: https://doi.org/10.1007/BF02812722.

[58] Y. Nambu. "Quasi-Particles and Gauge Invariance in the Theory of Superconductivity". In: *Physical Review* 117 (3 Feb. 1, 1960), pp. 648–663. DOI: 10.1103/PhysRev.117.648. URL: https://link.aps.org/doi/10.1103/PhysRev.117.648.

[59] J. Schwinger. "Quantum Electrodynamics. I. A Covariant Formulation". In: *Physical Review* 74 (10 Nov. 15, 1948), pp. 1439–1461. DOI: 10.1103/PhysRev.74.1439. URL: https://link.aps.org/doi/10.1103/PhysRev.74.1439.

[60] S. Tomonaga. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields." In: *Progress of Theoretical Physics* 1.2 (Aug. 1, 1946), pp. 27–42. DOI: 10.1143/PTP.1.27.

[61] R. P. Feynman. "The Theory of Positrons". In: *Physical Review* 76 (6 Sept. 15, 1949), pp. 749–759. DOI: 10.1103/PhysRev.76.749. URL: https://link.aps.org/doi/10.1103/PhysRev.76.749.

[62] T. Banks and A. Zaks. "Composite gauge bosons in 4-fermi theories". In: *Nuclear Physics B* 184 (Oct. 31, 1980), pp. 303–322. ISSN: 0550-3213. DOI: 10.1016/0550-3213(81)90220-0.

[63] M. Waßmer. "Search for the production of Dark Matter in hadronic mono-top signatures with the CMS experiment". PhD thesis. Karlsruher Institut für Technologie (KIT), 2021. 247 pp. DOI: 10.5445/IR/1000129768.

[64] T. Appelquist and J. Carazzone. "Infrared singularities and massive fields". In: *Physical Review D* 11 (10 May 15, 1975), pp. 2856–2861. DOI: 10.1103/PhysRevD.11.2856. URL: https://link.aps.org/doi/10.1103/PhysRevD.11.2856.

[65] C. Degrande et al. *Standard Model Effective Theory at One-Loop in QCD*. May 1, 2021. URL: http://feynrules.irmp.ucl.ac.be/wiki/SMEFTatNLO.

[66] A. Helset and A. Kobach. "Baryon number, lepton number, and operator dimension in the SMEFT with flavor symmetries". In: *Physics Letters B* 800 (Jan. 10, 2020). DOI: 10.1016/j.physletb.2019.135132.

[67] B. Grzadkowski et al. "Dimension-six terms in the Standard Model Lagrangian". In: *Journal of High Energy Physics* 2010.10 (Oct. 2010). DOI: 10.1007/jhep10(2010)085.

[68] C. Degrande et al. "Automated one-loop computations in the SMEFT". In: *Physical Review D* 103.096024 (May 24, 2021). DOI: 10.1103/PhysRevD.103.096024. arXiv: 2008.11743 [hep-ph].

[69] J. Ellis et al. "Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory". In: *Journal of High Energy Physics* 2021.279 (Apr. 29, 2021). DOI: 10.1007/JHEP04(2021)279. arXiv: 2012.02779 [hep-ph].

[70] M. D. Schwartz. "TASI Lectures on Collider Physics". In: *Anticipating the Next Discoveries in Particle Physics*. 2018. Chap. Chapter 2, pp. 65–100. DOI: 10.1142/9789813233348_0002. eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789813233348_0002. URL: https://www.worldscientific.com/doi/abs/10.1142/9789813233348_0002.

[71] J. D. Bjorken. "Asymptotic Sum Rules at Infinite Momentum". In: *Physical Review* 179 (5 Mar. 25, 1969), pp. 1547–1553. DOI: 10.1103/PhysRev.179.1547. URL: https://link.aps.org/doi/10.1103/PhysRev.179.1547.

[72] G. ’. Hooft. “Renormalizable lagrangians for massive Yang-Mills fields”. In: *Nuclear physics: B* 35.1 (1971), pp. 167–188. URL: https://dspace.library.uu.nl/handle/1874/4733.

[73] G. ’. Hooft and M. Veltman. “Regularization and renormalization of gauge fields”. In: *Nuclear Physics B* 44.1 (July 1, 1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9.

[74] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and collider physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Photography by S. Vascotto. Cambridge: Cambridge University Press, 2003. DOI: 10.1017/CBO9780511628788. URL: https://cds.cern.ch/record/318585.

[75] K. Kumerički. *Feynman Diagrams for Beginners*. Tech. rep. Feb. 8, 2016. arXiv: 1602.04182 [physics.ed-ph].

[76] F. Maltoni, G. Ridolfi, and M. Ubiali. “b-initiated processes at the LHC: a reappraisal”. In: *Journal of High Energy Physics* 2012.22 (July 3, 2012). DOI: 10.1007/jhep07(2012)022.

[77] J. C. Collins, D. E. Soper, and G. Sterman. “Factorization of Hard Processes in QCD. Perturbative QCD”. In: *Advanced Series on Directions in High Energy Physics* (May 30, 1989), pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313 [hep-ph].

[78] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.S08004 (Aug. 2008). URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[79] CMS Collaboration, G. L. Bayatian, et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CERN-LHCC-2006-001 ; CMS-TDR-8-1. Geneva, 2006.

[80] CMS Collaboration. *Expected performance of the physics objects with the upgraded CMS detector at the HL-LHC*. Tech. rep. Geneva: CERN, Dec. 2018. URL: https://cds.cern.ch/record/2650976.

[81] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: https://cds.cern.ch/record/782076.

[82] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-2. URL: http://cds.cern.ch/record/815187.

[83] M. Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: http://cds.cern.ch/record/823808.

[84] E. Mobs. *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*. Tech. rep. General Photo. July 29, 2019. URL: https://cds.cern.ch/record/2684277.

[85] T. Sakuma and T. McCauley. “Detector and Event Visualization with SketchUp at the CMS Experiment”. In: *Journal of Physics: Conference Series* 513.022032 (2 2014). Ed. by D. L. Groep and D. Bonacorsi. DOI: 10.1088/1742-6596/513/2/022032. arXiv: 1311.4942 [physics.ins-det].

[86] V. Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412.

[87] CMS Collaboration. *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: http://cdsweb.cern.ch/record/490194.

[88] CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *Journal of Instrumentation* 9.P10009 (May 26, 2014). DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569v2 [physics.ins-det].

[89] F. Hartmann. *Evolution of Silicon Sensor Technology in Particle Physics*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-64436-3.

[90] The Tracker Group of the CMS Collaboration. "The CMS Phase-1 Pixel Detector Upgrade". In: *Journal of Instrumentation* 16.P02027 (Feb. 22, 2021). DOI: 10.1088/1748-0221/16/02/P02027. arXiv: 2012.14304 [physics.ins-det].

[91] CMS Collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375.

[92] P. Bloch et al. *Changes to CMS ECAL electronics: addendum to the Technical Design Report*. Technical design report. CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/581342.

[93] A. Benaglia. "The CMS ECAL performance with examples". In: *Journal of Instrumentation* 9.C02008 (Feb. 11, 2014). URL: http://stacks.iop.org/1748-0221/9/i=02/a=C02008.

[94] CMS Collaboration. *ECAL 2016 refined calibration and Run2 summary plots*. Tech. rep. Apr. 20, 2020. URL: https://cds.cern.ch/record/2717925.

[95] CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/357153/?ln=de.

[96] CMS Collaboration. "Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data". In: *Journal of Instrumentation* 5.T03012 (Mar. 19, 2010). URL: http://stacks.iop.org/1748-0221/5/i=03/a=T03012.

[97] CMS Collaboration. *HCAL Energy Reconstruction Performance*. Tech. rep. Nov. 25, 2016. URL: http://cds.cern.ch/record/2235509.

[98] CMS Collaboration. *HCAL Calibration Status in Summer 2017*. Tech. rep. Aug. 17, 2017. URL: http://cds.cern.ch/record/2281146.

[99] CMS Collaboration. *Results related to the Phase1 HE upgrade*. Tech. rep. May 29, 2018. URL: https://cds.cern.ch/record/2320857.

[100] CMS Collaboration. *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331056?.

[101] CMS Collaboration. *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814.

[102] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of Instrumentation* 13.P06015 (June 19, 2018). DOI: 10.1088/1748-0221/13/06/p06015. URL: https://doi.org/10.1088/1748-0221/13/06/p06015.

[103] CMS Collaborations. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical design report. CMS. URL: https://cds.cern.ch/record/706847.

[104] CMS Collaboration. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/578006.

[105] CMS Collaboration. "CMS trigger performance". In: *EPJ Web of Conferences* 182.02037 (Aug. 3, 2018). Ed. by Y. Aharonov, L. Bravina, and S. Kabana, p. 14. DOI: 10.1051/epjconf/201818202037.

[106] CMS Collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 15.P10017 (Oct. 19, 2020). DOI: 10.1088/1748-0221/15/10/p10017.

[107] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (Dec. 15, 1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: http://www.sciencedirect.com/science/article/pii/0168900287908874.

[108] G. Cerati et al. "Parallelized and Vectorized Tracking Using Kalman Filters with CMS Detector Geometry and Events". In: *EPJ Web of Conferences* 214.02002 (Sept. 17, 2019), p. 8. DOI: 10.1051/epjconf/201921402002. arXiv: 1811.04141 [physics.comp-ph].

[109] CMS Collaboration. "Alignment of the CMS tracker with LHC and cosmic ray data". In: *Journal of Instrumentation* 9.P06009 (Mar. 10, 2014). DOI: 10.1088/1748-0221/9/06/P06009. arXiv: 1403.2286v2 [physics.ins-det].

[110] CMS Collaboration. "Alignment of the CMS Tracker: Latest Results from LHC Run-II". In: *Journal of Physics: Conference Series* 898.042014 (Oct. 2017). DOI: 10.1088/1742-6596/898/4/042014. URL: https://doi.org/10.1088/1742-6596/898/4/042014.

[111] CMS Collaboration. *Utilities for Accessing Pileup Information for Data: Recommended cross section*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData#Recommended_cross_section r36.

[112] CMS Collaboration. *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*. Tech. rep. CMS-PAS-PFT-09-001. Geneva: CERN, Apr. 2009. URL: https://cds.cern.ch/record/1194487.

[113] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *Journal of Instrumentation* 12.P10003 (June 15, 2017). DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965v2 [physics.ins-det].

[114] M. Dordevic. "The CMS Particle Flow Algorithm". In: *EPJ Web of Conferences* 191.02016 (Oct. 31, 2018), p. 7. DOI: 10.1051/epjconf/201819102016. URL: http://cds.cern.ch/record/2678077.

[115] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 13.P06015 (Apr. 12, 2018). DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].

[116] CMS Collaboration. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 16.P05014 (May 16, 2021). DOI: 10.1088/1748-0221/16/05/P05014. arXiv: 2012.06888 [hep-ex].

[117] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.63 (4 Apr. 16, 2008). DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189v2 [hep-ph].

[118] CMS Collaboration. *Jet algorithms performance in 13 TeV data*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2256875.

[119] CMS Collaboration. "Pileup mitigation at CMS in 13 TeV data". In: *Journal of Instrumentation* 15.P09018 (Sept. 15, 2020). DOI: 10.1088/1748-0221/15/09/p09018. URL: https://doi.org/10.1088/1748-0221/15/09/p09018.

[120] E. Bols et al. "Jet flavour classification using DeepJet". In: *Journal of Instrumentation* 15.12 (Dec. 2020), P12012–P12012. DOI: 10.1088/1748-0221/15/12/p12012.

[121] CMS Collaboration. *Heavy flavour tagging for 13 TeV 2016 Legacy data*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation2016Legacy r16. Feb. 16, 2021.

[122] CMS Collaboration. *Heavy flavour tagging for 13 TeV data in 2017 and 94X MC*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation94X r40. Feb. 16, 2021.

[123] CMS Collaboration. *Heavy flavour tagging for 13 TeV data in 2018 and 10_2_X MC*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation102X r22. Feb. 16, 2021.

[124] CMS Collaboration. *Reweighting recipe to emulate Level 1 ECAL prefiring*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/L1ECALPrefiringWeightRecipe r11.

[125] CMS Collaboration. *JetMET Hypernews posting*. https://hypernews.cern.ch/HyperNews/CMS/get/JetMET/2000.html.

[126] CMS Collaboration. *JEC Status Report HEM issue Studies*. https://indico.cern.ch/event/846608/contributions/3556096/attachments/1904250/3144533/JERCMeeting_9_9_2019_V2.pdf.

[127] CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, June 2017. URL: https://cds.cern.ch/record/2272264.

[128] CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. This is the final version, approved by the LHCC. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2283187.

[129] CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. Geneva: CERN, Nov. 2017. URL: https://cds.cern.ch/record/2293646.

[130] CMS Collaboration. *The Phase-2 Upgrade of the CMS Muon Detectors*. Tech. rep. This is the final version, approved by the LHCC. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2283189.

[131] CMS Collaboration. *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade*. Tech. rep. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2296612.

[132] CMS Collaboration. *The Phase-2 Upgrade of the CMS Level-1 Trigger*. Tech. rep. Final version. Geneva: CERN, Apr. 2020. URL: http://cds.cern.ch/record/2714892.

[133] CMS Collaboration. "Search for $t\bar{t}H$ production in the $Hb\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2019.3 (2019), pp. 1–62. URL: https://doi.org/10.1007/JHEP03(2019)026.

[134] P. Keicher. "Model Stability Studies in a Search for $t\bar{t}H$ Production in the $H \to b\bar{b}$ Channel at CMS". MA thesis. Karlsruhe Institute of Technology (KIT), 2018. URL: https://publish.etp.kit.edu/record/21907.

[135] N. Bartosik and J. Hauk. *Hadron based origin identification of heavy flavour jets at generator level*. CMS public url: https://twiki.cern.ch/twiki/bin/view/CMSPublic/GenHFHadronMatcher. CMS Collaboration, Jan. 30, 2017. (Visited on 02/09/2018).

[136] M. M. Horzela. "Merging of $t\bar{t}$ and $t\bar{t}$+$b\bar{b}$ simulations for an improved background modeling for ttäH(bbä) measurements". MA thesis. Karlsruhe Institute of Technology (KIT), 2019. URL: https://publish.etp.kit.edu/record/21675.

[137] CMS Collaboration. *Measurement of $t\bar{t}H$ and tH in the H→ $b\bar{b}$ channel with the full Run 2 data sample*. CMS internal documentation: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2019/094. 2021.

[138] LHC Top Working Group. *NLO single-top channel cross sections ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program*. TWiki: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec r34.

[139] A. Kulesza et al. "Associated production of a top quark pair with a heavy electroweak gauge boson at NLO+NNLL accuracy". In: *The European Physical Journal C* 79.3 (Mar. 2019), p. 249. DOI: 10.1140/epjc/s10052-019-6746-z. arXiv: 1812.08622 [hep-ph].

[140] T. Gehrmann et al. "W^+W^--Production at Hadron Colliders in Next to Next to Leading Order QCD". In: *Physical Review Letters* 113.212001 (Nov. 18, 2014). DOI: 10.1103/physrevlett.113.212001.

[141] N. Metropolis and S. Ulam. "The Monte Carlo Method". In: *Journal of the American Statistical Association* 44.247 (Sept. 1949), pp. 335–341. DOI: 10.1080/01621459.1949.10483310.

[142] I. M. Sobol. *A Primer for the Monte Carlo Method*. CRC Press, Apr. 2018. DOI: 10.1201/9781315136448.

[143] A. Buckley et al. "General-purpose event generators for LHC physics". In: *Physics Reports* 504.5 (July 2011), pp. 145–233. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2011.03.005. URL: https://www.sciencedirect.com/science/article/pii/S0370157311000846.

[144] P. Nason. "A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms". In: *Journal of High Energy Physics* 2004.40 (Dec. 10, 2004). DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph].

[145] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *Journal of High Energy Physics* 11 (Sept. 13, 2007). DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092v1 [hep-ph].

[146] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *Journal of High Energy Physics* 2010.6 (June 2010). DOI: 10.1007/jhep06(2010)043. URL: https://doi.org/10.1007/jhep06(2010)043.

[147] H. B. Hartanto et al. "Higgs boson production in association with top quarks in the POWHEG BOX". In: *Physical Review D* 91.094003 (9 May 5, 2015). DOI: 10.1103/PhysRevD.91.094003. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.094003.

[148] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 07.79 (July 17, 2014). DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301v2 [hep-ph].

[149] T. Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012v1 [hep-ph].

[150] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (July 1, 2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[151] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (Mar. 27, 2006), pp. 270–278. DOI: 10.1109/tns.2006.869826.

[152] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 1, 2016), pp. 186–225. DOI: 10.1016/j.nima.2016.06.125.

[153] R. D. Ball et al. "Parton distributions from high-precision collider data". In: *The European Physical Journal C* 77.10 (Oct. 2017). DOI: 10.1140/epjc/s10052-017-5199-5.

[154] T. Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nuclear Physics B* 248.2 (Dec. 24, 1984), pp. 469–502. DOI: 10.1016/0550-3213(84)90607-2.

[155] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2-3 (July 1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[156] CMS Collaboration et al. "Event generator tunes obtained from underlying event and multiparton scattering measurements". In: *The European Physical Journal C* 76.3 (Mar. 2016). DOI: 10.1140/epjc/s10052-016-3988-x.

[157] CMS Collaboration et al. "Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements". In: *The European Physical Journal C* 80.1 (Jan. 2020). DOI: 10.1140/epjc/s10052-019-7499-4.

[158] LHC Higgs Cross Section Working Group. *SM Higgs production cross sections at $\sqrt{s}$ = 13 TeV*. TWiki: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV r23.

[159] F. Demartin et al. "Higgs production in association with a single top quark at the LHC". In: *The European Physical Journal C* 75.6 (June 2015). DOI: 10.1140/epjc/s10052-015-3475-9.

[160] F. Demartin et al. "tWH associated production at the LHC". In: *The European Physical Journal C* 77.1 (Jan. 2017). DOI: 10.1140/epjc/s10052-017-4601-7.

[161] ATLAS-CMS. *NNLO+NNLL top-quark-pair cross sections*. TWiki: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO. Sept. 14, 2015. (Visited on 02/06/2018).

[162] CMS Collaboration. *Standard Model Cross Sections for CMS at 13 TeV*. CMS internal url: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat13TeV` r26.

[163] LHC Higgs Cross Section Working Group. *SM Higgs Branching Ratios and Total Decay Widths*. TWiki: `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR` r22.

[164] LHC Higgs Cross Section Working Group. *Modelling of the single top-quark production in association with the Higgs boson at 13 TeV*. CMS internal url: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV` r15.

[165] M. Beneke et al. "Hadronic top-quark pair production with NNLL threshold resummation". In: *Nuclear Physics B* 855.3 (Feb. 2012), pp. 695–741. DOI: `10.1016/j.nuclphysb.2011.10.021`.

[166] M. Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Physics Letters B* 710.4-5 (Apr. 2012), pp. 612–622. DOI: `10.1016/j.physletb.2012.03.013`.

[167] P. Bärnreuther, M. Czakon, and A. Mitov. "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t}+X$". In: *Physical Review Letters* 109.13 (Sept. 25, 2012). DOI: `10.1103/physrevlett.109.132001`.

[168] M. Czakon and A. Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *Journal of High Energy Physics* 2012.12 (Dec. 2012). DOI: `10.1007/jhep12(2012)054`.

[169] M. Czakon and A. Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *Journal of High Energy Physics* 2013.1 (Jan. 2013). DOI: `10.1007/jhep01(2013)080`.

[170] M. Czakon, P. Fiedler, and A. Mitov. "The total top quark pair production cross-section at hadron colliders through $O(\alpha_S^4)$". In: *Physical Review Letters* 110.25 (June 18, 2013). DOI: `10.1103/PhysRevLett.110.252004`. arXiv: `1303.6254v1 [hep-ph]`.

[171] M. Czakon and A. Mitov. "Top++: a program for the calculation of the top-pair cross-section at hadron colliders". In: *Computer Physics Communications* 185.11 (Dec. 23, 2011), pp. 2930–2938. DOI: `10.1016/j.cpc.2014.06.021`. arXiv: `1112.5675v4 [hep-ph]`.

[172] T. Ježo et al. "New NLOPS predictions for $t\bar{t}$ + bjet production at the LHC". In: *The European Physical Journal C* 78.6 (June 2018). DOI: `10.1140/epjc/s10052-018-5956-0`.

[173] M. Schröder. *Comparison of tt + bb for the LHCXSWG*. CMS HIG PAG meeting: `https://indico.cern.ch/event/847928/#8-lhchxsecwg-comparison-of-tth`.

[174] LHC Higgs Cross Section Working Group. *LHCHXSWG working meeting: discussion on tt+bb (MC comparisons)*. Oct. 1, 2019. URL: `https://indico.cern.ch/event/850548/`.

[175] S. Alioli et al. "NLO single-top production matched with shower in POWHEG:s-andt-channel contributions". In: *Journal of High Energy Physics* 2009.09 (Sept. 2009), pp. 111–111. DOI: `10.1088/1126-6708/2009/09/111`.

[176] E. Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *The European Physical Journal C* 71.2 (Feb. 2011). DOI: `10.1140/epjc/s10052-011-1547-z`. arXiv: `1009.2450v1 [hep-ph]`.

[177] N. Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-". In: *Physical Review D* 82.5 (May 24, 2010). DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451v1 [hep-ph].

[178] M. Aliev et al. "HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR". In: *Computer Physics Communications* 182.4 (Apr. 2011), pp. 1034–1046. DOI: 10.1016/j.cpc.2010.12.040.

[179] P. Kant et al. "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions". In: *Computer Physics Communications* 191 (June 2015), pp. 74–89. DOI: 10.1016/j.cpc.2015.02.001.

[180] M. Botje et al. *The PDF4LHC Working Group Interim Recommendations*. Tech. rep. Jan. 3, 2011. arXiv: 1101.0538 [hep-ph].

[181] A. D. Martin et al. "Parton distributions for the LHC". In: *The European Physical Journal C* 63.2 (July 2009), pp. 189–285. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-009-1072-5. URL: https://doi.org/10.1140/epjc/s10052-009-1072-5.

[182] A. D. Martin et al. "Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections". In: *The European Physical Journal C* 64.4 (Oct. 2009), pp. 653–680. DOI: 10.1140/epjc/s10052-009-1164-2.

[183] H.-L. Lai et al. "New parton distributions for collider physics". In: *Physical Review D* 82.074024 (Oct. 26, 2010). DOI: 10.1103/physrevd.82.074024.

[184] J. Gao et al. "CT10 next-to-next-to-leading order global analysis of QCD". In: *Physical Review D* 89.033009 (Feb. 19, 2014). DOI: 10.1103/physrevd.89.033009.

[185] R. D. Ball et al. "Parton distributions with LHC data". In: *Nuclear Physics B* 867.2 (Feb. 2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003.

[186] R. Frederix and S. Frixione. "Merging meets matching in MC@NLO". In: *Journal of High Energy Physics* 2012.12 (Sept. 27, 2012). DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215v1 [hep-ph].

[187] J. M. Campbell, R. K. Ellis, and C. Williams. "Vector boson pair production at the LHC". In: *Journal of High Energy Physics* 2011.7 (Apr. 29, 2011). DOI: 10.1007/JHEP07(2011)018. arXiv: 1105.0020v1 [hep-ph].

[188] D. M. Morse and S. Cooper. *Corrected Cross Sections for Binned V+jet Samples*. Oct. 16, 2017. URL: https://indico.cern.ch/event/673253/#2-corrected-cross-sections-for.

[189] CMS Collaboration. *Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/SummaryTable1G25ns r152.

[190] CMS Collaboration. *Reference muon id, isolation and trigger efficiencies for 2016 legacy re-reco data*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco r10.

[191] CMS Collaboration. *Reference muon id, isolation and trigger efficiencies for 2017 data*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017 r30.

[192] CMS Collaboration. *Reference muon id, isolation and trigger efficiencies for 2018 data*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018 r8.

[193] CMS Collaboration. *Rochester corrections*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/RochcorMuon r29.

[194] CMS Collaboration. *E/gamma RECO: Electron Scale Factors*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations#E_gamma_RECO r13.

[195] CMS Collaboration. *E/gamma IDs: Electrons IDs: Fall17v2*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations#Fall17v2 r13.

[196] A. Datta et al. *Trigger Efficiencies and Scale Factors for the $t\bar{t}H$ ($H\rightarrow b\bar{b}$) Measurement in Single Lepton and Dilepton Channels with the Full Run-2 Data Sample*. CMS internal documentation: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2019/008. Oct. 25, 2020.

[197] CMS Collaboration. *Introduction to Jet Energy Corrections at CMS*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC r6. Oct. 4, 2021. (Visited on 10/04/2021).

[198] CMS Collaboration. *Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018*. Tech. rep. CMS Collaboration, Apr. 23, 2020. URL: http://cds.cern.ch/record/2715872.

[199] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *Journal of Instrumentation* 12.P02014 (July 13, 2016). DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663v2 [hep-ex].

[200] CMS Collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (July 21, 2011). DOI: 10.1088/1748-0221/6/11/P11002. arXiv: 1107.4277v1 [physics.ins-det].

[201] CMS Collaboration. *Jet Energy Resolution*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution#JER_Scaling_factors_and_Uncertai r81.

[202] CMS Collaboration. *Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13TeV with Phase 1 CMS detector*. Tech. rep. CMS Collaboration, Apr. 9, 2019. URL: http://cds.cern.ch/record/2646773.

[203] CMS Collaboration. *B-tagging performance of the CMS Legacy dataset 2018*. Tech. rep. Mar. 2021. URL: https://cds.cern.ch/record/2759970.

[204] CMS Collaboration. *Performance of b-tagging algorithms in 2016 data at 13 TeV*. CMS internal documentation: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2017/018. 2017. (Visited on 02/09/2018).

[205] R. Frederix et al. "The automation of next-to-leading order electroweak calculations". In: *Journal of High Energy Physics* 2018.185 (July 31, 2018). DOI: 10.1007/JHEP07(2018)185. arXiv: 1804.10017 [hep-ph].

[206] P. Keicher and K. Mimasu. *Private Communication*. Aug. 31, 2021.

[207] A. Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *The European Physical Journal C* 75.132 (Mar. 20, 2015). DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[208] G. Ossola, C. G. Papadopoulos, and R. Pittau. "CutTools: a program implementing the OPP reduction method to compute one-loop amplitudes". In: *Journal of High Energy Physics* 2008.42 (Mar. 13, 2008). DOI: 10.1088/1126-6708/2008/03/042. arXiv: 0711.3596 [hep-ph].

[209] T. Peraro. "Ninja: Automated Integrand Reduction via Laurent Expansion for One-Loop Amplitudes". In: *Computer Physics Communications* 185 (Mar. 5, 2014), pp. 2771–2797. DOI: 10.1016/j.cpc.2014.06.017. arXiv: 1403.1229 [hep-ph].

[210] A. Denner, S. Dittmaier, and L. Hofer. "Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations". In: *Computer Physics Communications* 212 (Nov. 4, 2016), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph].

[211] P. de Aquino et al. "ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman diagram computations". In: *Computer Physics Communications* 183 (Aug. 9, 2011), pp. 2254–2263. DOI: 10.1016/j.cpc.2012.05.004. arXiv: 1108.2041 [hep-ph].

[212] H. Murayama, I. Watanabe, and K. Hagiwara. *HELAS: HELicity amplitude subroutines for Feynman diagram evaluations*. Tech. rep. Jan. 1992.

[213] P. Nason. *MINT: a Computer Program for Adaptive Monte Carlo Integration and Generation of Unweighted Distributions*. Comp. software. Sept. 13, 2007. arXiv: 0709.2085 [hep-ph].

[214] T. Binoth et al. "A proposal for a standard interface between Monte Carlo tools and one-loop programs". In: *Computer Physics Communications* 181 (Jan. 8, 2010), pp. 1612–1622. DOI: 10.1016/j.cpc.2010.05.016. arXiv: 1001.1307 [hep-ph].

[215] J. Butterworth et al. "Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report". In: *Standard Model Working Group of the 2013 Les Houches Workshop, Physics at TeV Colliders, Les houches*. May 5, 2014. arXiv: 1405.1067 [hep-ph].

[216] C. Bierlich et al. "Robust Independent Validation of Experiment and Theory: Rivet version 3". In: *SciPost Physics* 8 (2 Dec. 11, 2019), p. 26. DOI: 10.21468/SciPostPhys.8.2.026. arXiv: 1912.05451 [hep-ph].

[217] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.1554 (Feb. 9, 2011). DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727v3 [physics.data-an].

[218] G. Cowan. *Topics in statistical data analysis for high-energy physics*. Tech. rep. Dec. 16, 2010, pp. 197–218. arXiv: 1012.3589v1 [physics.data-an].

[219] G. Cowan. *Statistical Data Analysis*. Oxford university press, 1998.

[220] O. Behnke et al. *Data Analysis in High Energy Physics*. Wiley-VCH GmbH, June 26, 2013. ISBN: 3527410589. URL: https://www.ebook.de/de/product/20188461/data_analysis_in_high_energy_physics.html.

[221] LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011. URL: https://cds.cern.ch/record/1379837.

[222] R. Bellman. *Dynamic programming*. Mineola, N.Y: Dover Publications, 2003. ISBN: 9780486428093.

[223] R. Bellman. *Adaptive control processes : a guided tour*. Princeton, N.J: Princeton University Press, 2015. ISBN: 9780691079011.

[224] J. Neyman and E. S. Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 231.694-706 (Feb. 16, 1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009. URL: http://rsta.royalsocietypublishing.org/content/231/694-706/289.

[225] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (Mar. 1938), pp. 60–62. ISSN: 00034851. DOI: 10.1214/aoms/1177732360.

[226] J. S. Conway. *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*. Tech. rep. Mar. 2, 2011. arXiv: 1103.0354v1 [physics.data-an].

[227] N. Shadskiy. "Treating Uncertainties with Bayesian Neural Networks in the measurement of ttH(bb) production". MA thesis. Karlsruhe Institute of Technology (KIT), 2020. URL: https://publish.etp.kit.edu/record/21982.

[228] J. Schindler. "Reducing background modeling uncertainties in the CMS analysis of the associated production of a Higgs boson and a top quark-antiquark pair with adversarial neural networks". MA thesis. Karlsruhe Institute of Technology (KIT), 2019. URL: https://publish.etp.kit.edu/record/21672.

[229] S. Ehnle. "Uncertainty studies in a measurement of $t\bar{t}H$ production in the $H \to b\bar{b}$ Channel at CMS". MA thesis. Karlsruhe Institute of Technology (KIT), 2020. URL: https://publish.etp.kit.edu/record/21972.

[230] M. Zakaria, M. A. Al-Shebany, and S. Sarhan. "Artificial Neural Network : A Brief Overview". In: 2014.

[231] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. URL: http://www.deeplearningbook.org.

[232] M. A. Nielsen. *Neural networks and deep learning*. Vol. 25. Determination press San Francisco, CA, 2015. URL: http://neuralnetworksanddeeplearning.com/index.html.

[233] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain." In: *Psychological Review* 65.6 (1958), pp. 386–408. DOI: 10.1037/h0042519.

[234] Wikimedia Commons. *A Neural network with multiple layers*. Apr. 10, 2015. URL: https://commons.wikimedia.org/wiki/File:Multi-Layer_Neural_Network-Vector.svg (visited on 10/13/2021).

[235] J. S. Bridle. "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition". In: *Neurocomputing* (1990). Ed. by F. F. Soulié and J. Hérault, pp. 227–236. DOI: 10.1007/978-3-642-76153-9_28.

[236] E. B. Baum and F. Wilczek. "Supervised learning of probability distributions by neural networks". In: *Neural information processing systems*. Vol. 12. New York, NY. 1998, pp. 52–61.

[237] E. Levin and M. Fleisher. "Accelerated learning in layered neural networks". In: *Complex systems* 2 (1988), pp. 625–640.

[238] E. Gordon-Rodriguez et al. "Uses and Abuses of the Cross-Entropy Loss: Case Studies in Modern Deep Learning". In: *Proceedings on "I Can't Believe It's Not Better!" at NeurIPS Workshops*. Ed. by J. Zosa Forde et al. Vol. 137. Proceedings of Machine Learning Research. PMLR, Dec. 2020, pp. 1–10. URL: https://proceedings.mlr.press/v137/gordon-rodriguez20a.html.

[239] J. Duchi, E. Hazan, and Y. Singer. "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization". In: *Journal of Machine Learning Research* 12 (July 2011), pp. 2121–2159. ISSN: 1532-4435. URL: https://www.jmlr.org/papers/volume12/duchi11a/duchi11a.pdf.

[240] A. Lydia and S. Francis. "Adagrad—an optimizer for stochastic gradient descent". In: *International Journal of Information and Computing Science* 6.5 (May 5, 2019).

[241] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. "Learning Representations by Back-Propagating Errors". In: *Nature* 323 (Oct. 9, 1986), pp. 533–536. DOI: 10.1038/323533a0.

[242] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (Oct. 2011), pp. 2825–2830. URL: https://www.jmlr.org/papers/v12/pedregosa11a.html.

[243] R. Caruana, S. Lawrence, and L. Giles. "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping". In: *Advances in neural information processing systems* (2001), pp. 402–408. ISSN: 10495258.

[244] N. Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[245] A. Y. Ng. "Feature selection, L1 vs. L2 regularization, and rotational invariance". In: ACM Press, 2004. DOI: 10.1145/1015330.1015435.

[246] CMS Collaboration. *TOP Systematic Uncertainties (Run2)*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopSystematics r183.

[247] J. Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *Journal of Physics G: Nuclear and Particle Physics* 43.2 (Jan. 6, 2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. URL: https://doi.org/10.1088/0954-3899/43/2/023001.

[248] J. Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *The European Physical Journal C* 53 (Dec. 15, 2007), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569v2 [hep-ph].

[249] CMS Collaboration. *Lumi POG*. CMS internal url: https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM#SummaryTable r126.

[250] CMS Collaboration. *JECUncertaintySources TWiki*. CMS internal url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources#Run_2_reduced_set_of_uncertainty r51.

[251] R. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (Mar. 27, 1993), pp. 219–228. DOI: 10.1016/0010-4655(93)90005-w.

[252] CMS Collaboration. *CMS Combine Tool*. Oct. 1, 2021. URL: https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/ (visited on 10/01/2021).

[253] CMS Collaboration et al. "First measurement of the cross section for top quark pair production with additional charm jets using dileptonic final states in pp collisions at s=13TeV". In: *Physics Letters B* 820 (2021). ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2021.136565. URL: https://www.sciencedirect.com/science/article/pii/S0370269321005050.

[254] P. Keicher. *Full combination: fit results and bias tests*. CMS internal url: https://indico.cern.ch/event/963617/#167-full-combination-fit-resul. Dec. 9, 2019.

[255] L. Reuter. *Combined ANN for 2016-18*. CMS internal url: https://indico.cern.ch/event/867117/contributions/3664523/attachments/1959128/3255356/dnnstudiesupdate.pdf. Dec. 9, 2019.

[256] R. D. Cousins. *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms*. Tech. rep. Mar. 3, 2013. URL: http://www.physics.ucla.edu/%5C~cousins/stats/cousins_saturated.pdf.

[257] L. Bianchini et al. *Search for the associated ttH(H→bb) production using the Matrix Element Method*. CMS internal documentation: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2013/313. CMS Analysis Note. 2013.

[258] S. Wunsch et al. "Identifying the relevant dependencies of the neural network response on characteristics of the input space". In: *Computing and Software for Big Science* 2.5 (Sept. 17, 2018). DOI: 10.1007/s41781-018-0012-1. arXiv: 1803.08782 [physics.data-an].

[259] A. Ghorbani and J. Zou. "Data Shapley: Equitable Valuation of Data for Machine Learning". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by K. Chaudhuri and R. Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, June 2019, pp. 2242–2251. URL: http://proceedings.mlr.press/v97/ghorbani19c.html.

[260] P. Keicher. *Status Update SL*. CMS internal url: https://indico.cern.ch/event/996691/#164-status-update-sl. Dec. 9, 2019.

[261] CMS Collaboration et al. "Measurement of the cross section for $t\bar{t}$ production with additional jets and b jets in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Journal of High Energy Physics* 2020.7 (July 2020). DOI: 10.1007/jhep07(2020)125. URL: https://doi.org/10.1007/jhep07(2020)125.

[262] P. Keicher. *Results of Step 1 in Unblinding (GoF, pulls and impacts)*. CMS internal url: https://twiki.cern.ch/twiki/pub/CMS/HIG19011Review/ttHbb_unblinding_status_update.pdf. July 5, 2021.

[263] P. Krämer. "Measurement of $t\bar{t}H$ Simplified Template Cross Sections in the $H \to b\bar{b}$ channel at the CMS experiment". MA thesis. Karlsruhe Institute of Technology (KIT), 2020. URL: https://publish.etp.kit.edu/record/21939.

[264] ATLAS Collaboration. *Measurement of the Higgs boson decaying to b-quarks produced in association with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2020-058. Geneva: CERN, Nov. 2020. URL: https://cds.cern.ch/record/2743685.

[265] CMS Collaboration et al. "Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2021.7 (July 2021). DOI: 10.1007/jhep07(2021)027.

[266] CMS Collaboration. *Sensitivity projections for Higgs boson properties measurements at the HL-LHC*. Tech. rep. 2018. URL: https://inspirehep.net/literature/1703974.

[267] CMS collaboration. "Identification of b-quark jets with the CMS experiment". In: *Journal of Instrumentation* 8.04 (Apr. 2013), P04013–P04013. DOI: 10.1088/1748-0221/8/04/p04013. URL: https://doi.org/10.1088/1748-0221/8/04/p04013.

[268] D. Guest et al. "Jet flavor classification in high-energy physics with deep neural networks". In: *Physical Review D* 94.112002 (11 Dec. 2016). DOI: 10.1103/PhysRevD.94.112002. URL: https://link.aps.org/doi/10.1103/PhysRevD.94.112002.

[269] CMS Collaboration. *Lumi POG - Year-to-year combinations and corrleations*. CMS internal url: `https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM#LumiComb` r143.

Appendix

A Trigger paths for the SL channel

Table A.1: **Trigger paths for the SL channel**. Multiple triggers within the same run period are combined with a logical OR.

Trigger paths	Run era
Single-muon channel	
`HLT_IsoMu24_v*`	2016 B-H
`HLT_IsoTkMu24_v*`	2016 B-H
`HLT_IsoMu27_v*`	2017 B-F
`HLT_IsoMu24_v*`	2018 A-D
Single-electron channel	
`HLT_Ele27_WPTight_Gsf_v*`	2016 B-H
`HLT_Ele32_WPTight_Gsf_L1DoubleEG_v*` AND `ele32DoubleL1ToSingleL1Flag_v*`	2017 B-F
`HLT_Ele28_eta2p1_WPTight_Gsf_HT150_v*`	2017 B-F
`HLT_Ele32_WPTight_Gsf_v*`	2018 A-D
`HLT_Ele28_eta2p1_WPTight_Gsf_HT150_v*`	2018 A-D

B Additional distributions for the b-tagging scale factor correction factors

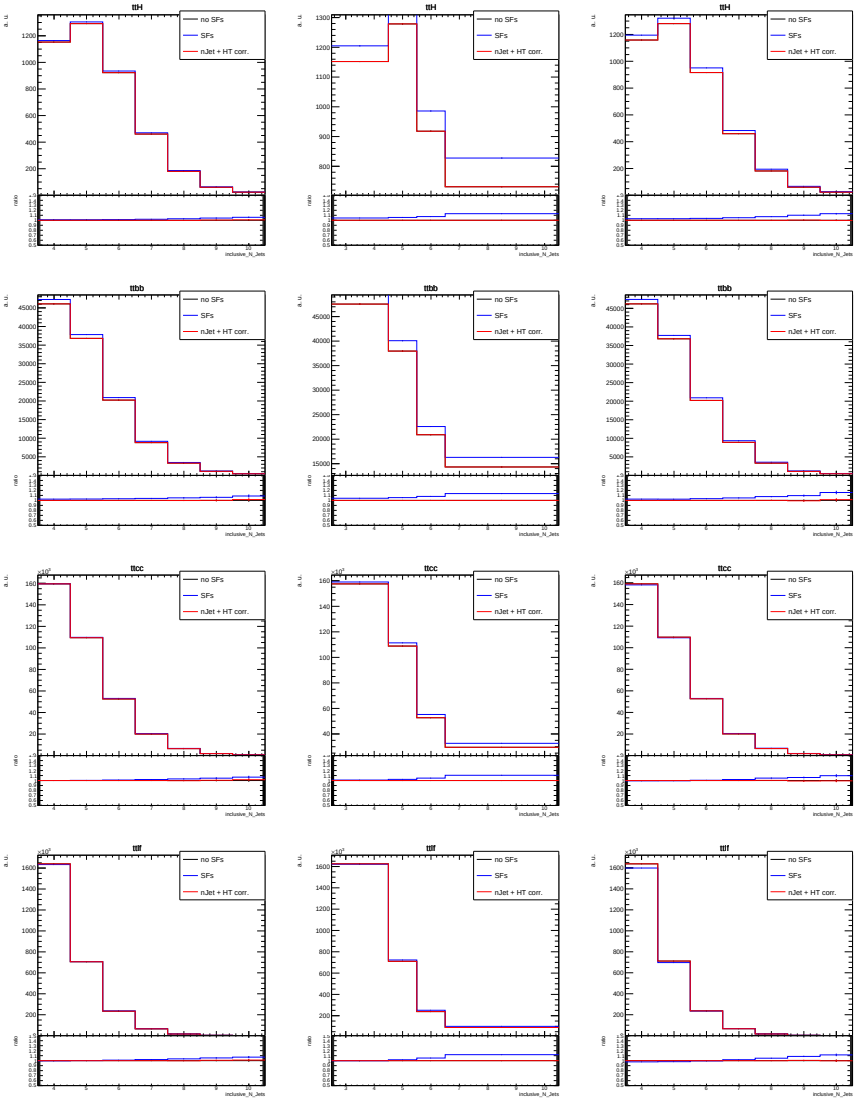


Figure B.1: Closure check for the number of jets before applying any b-tagging selection for 2016, 2017 and 2018 (from left to right) for the $t\bar{t}H$, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$ and $t\bar{t} + lf$ (from top to bottom) processes. The black line shows the distribution before applying the b-tagging SF. The blue line shows the distribution after applying the b-tagging SF and the red line shows the distribution after applying the b-tagging SF as well as the normalization scale factor. As expected the red distribution matches the black one.

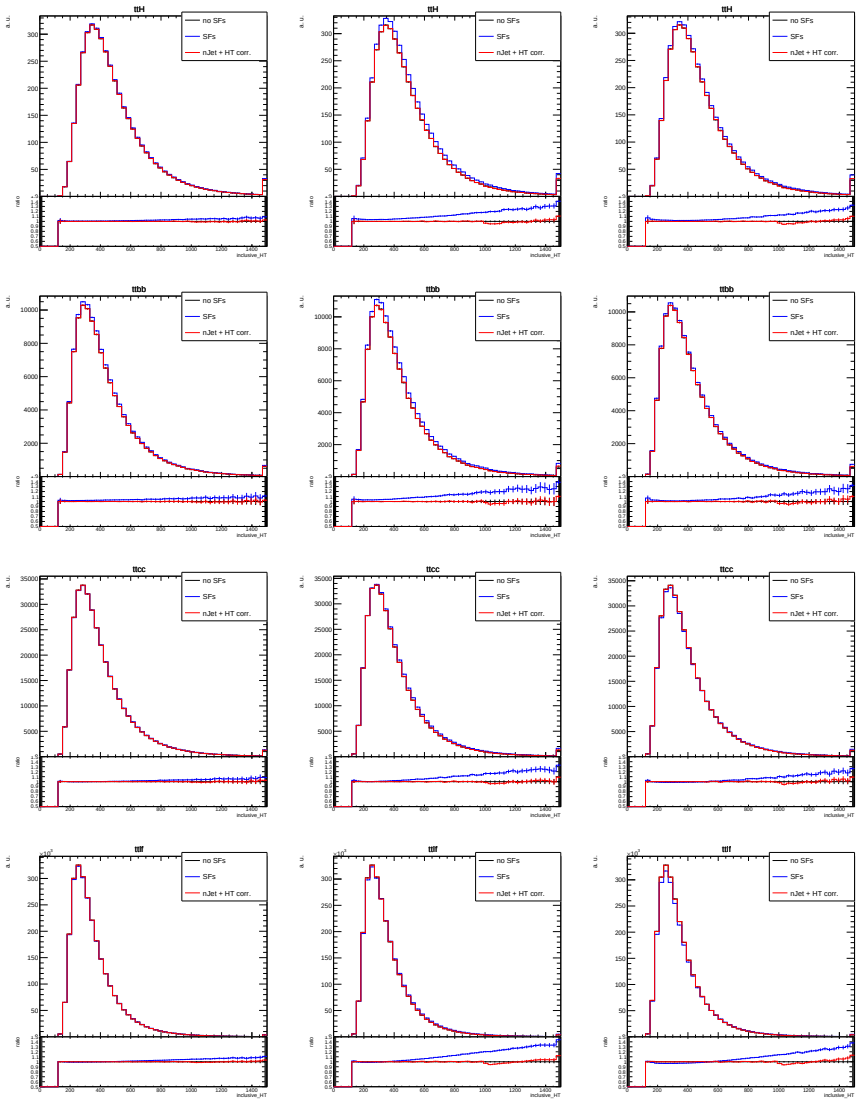


Figure B.2: Closure check for HT of the event before applying any b-tagging selection for 2016, 2017 and 2018 (from left to right) for the $t\bar{t}H$, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$ and $t\bar{t} + lf$ (from top to bottom) processes. The black line shows the distribution before applying the b-tagging SF. The blue line shows the distribution after applying the b-tagging SF and the red line shows the distribution after applying the b-tagging SF as well as the normalization scale factor. As expected the red distribution matches the black one.

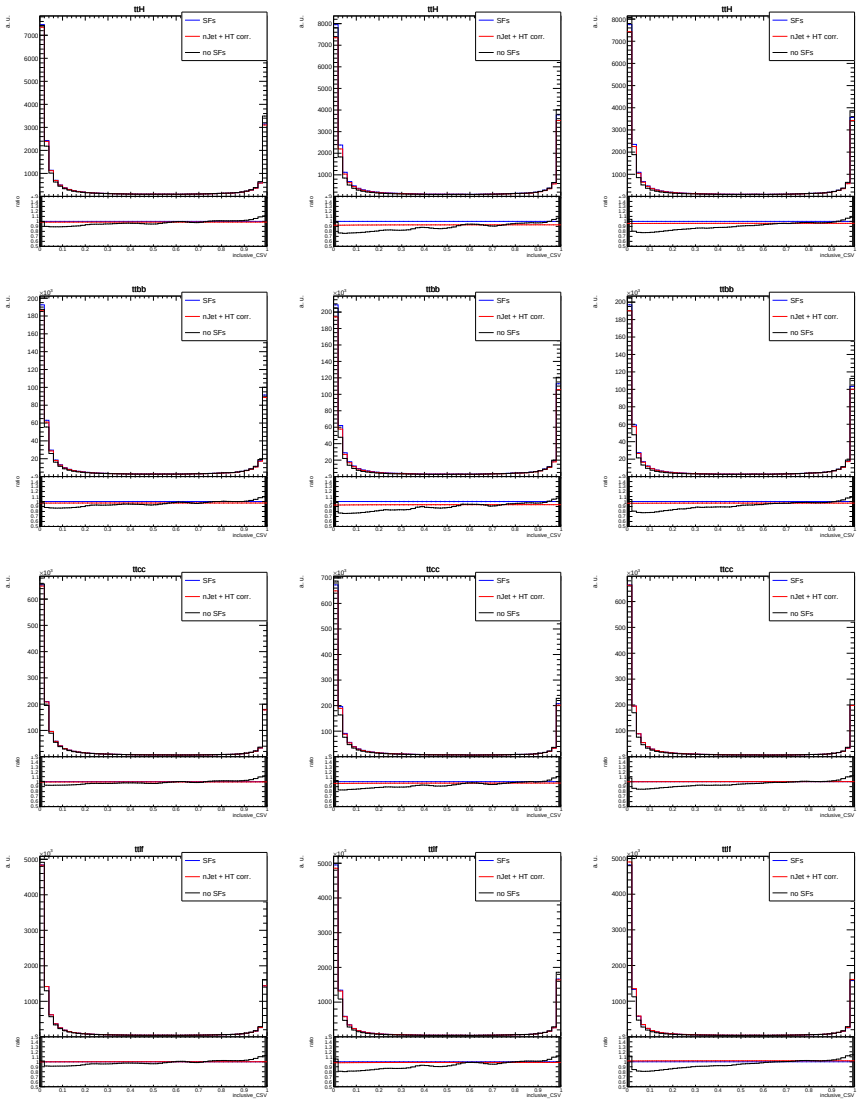


Figure B.3: Closure check for b-tagging value of all jets before applying any b-tagging selection for 2016, 2017 and 2018 (from left to right) for the $t\bar{t}H$, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$ and $t\bar{t}$ + lf (from top to bottom) processes. The black line shows the distribution before applying the b-tagging SF. The blue line shows the distribution after applying the b-tagging SF and the red line shows the distribution after applying the b-tagging SF as well as the normalization scale factor. These plots show that the application of the b-tagging normalization scale factor does not change the shape of the corrected b-tagging distribution.

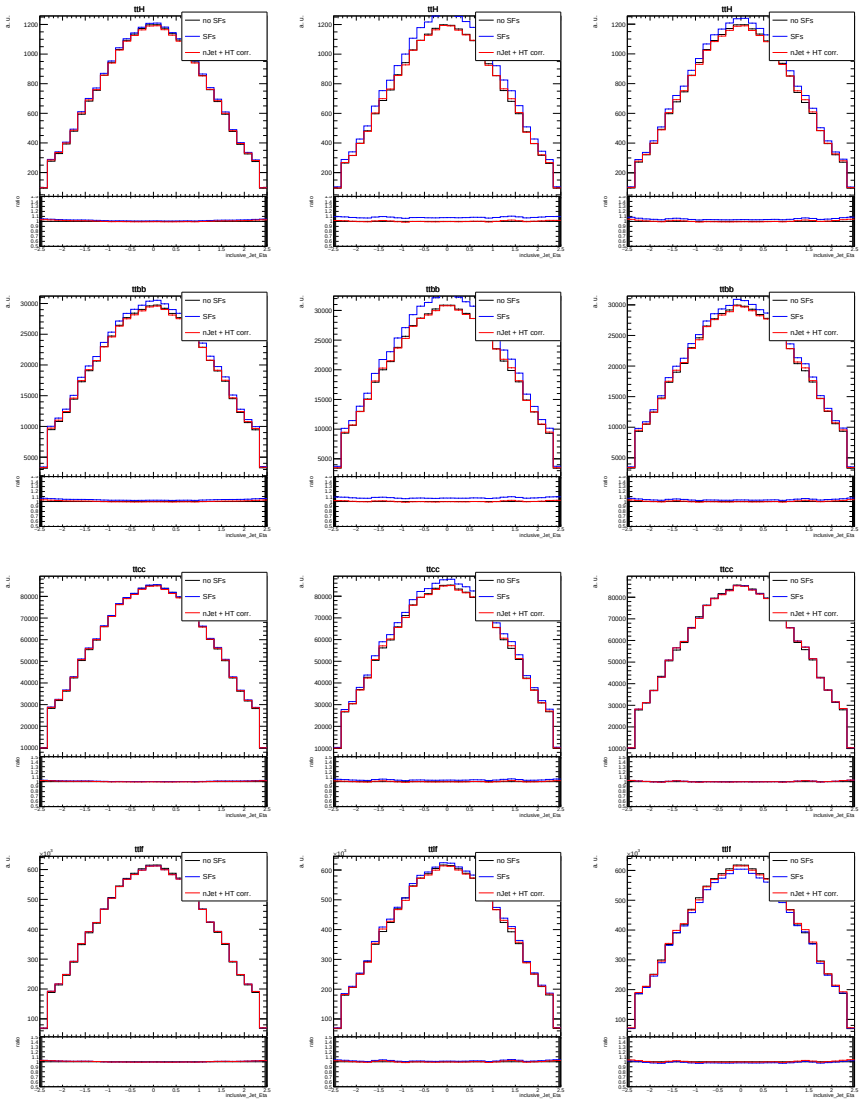


Figure B.4: Closure check for pseudo rapidity of all jets before applying any b-tagging selection for 2016, 2017 and 2018 (from left to right) for the $t\bar{t}H$, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$ and $t\bar{t} + lf$ (from top to bottom) processes. The black line shows the distribution before applying the b-tagging SF. The blue line shows the distribution after applying the b-tagging SF and the red line shows the distribution after applying the b-tagging SF as well as the normalization scale factor. These plots show that the application of the b-tagging normalization scale factor does not change the shape of the pseudo rapidity distribution of the jets.

C Details on the systematic uncertainties

Table C.2: Inclusive cross-section uncertainties due to the renormalisation and factorisation scales on the signal (top) and the background (bottom) processes. Uncertainties in the same column for two different processes (different rows) are treated as correlated. Values are taken from Ref. [138, 139, 158, 161, 162].

Process	Scales($t\bar{t}H$) in %	Scales(tHq) in %	Scales(tHW) in %
$t\bar{t}H$	+5.8/-9.2		
tHq		+6.5/-14.9[(1)]	
tHW			+4.9/-6.7

Process	Scales($t\bar{t}$) in percent	Scales(t) %	Scales(V) in %	Scales(VV) in %
$t\bar{t}$	+2.4/-3.5			
single t [(2)]		+3.1/-2.1		
W + jets			±3.8	
Z + jets			±2	
$t\bar{t}$ + W	+25.5/-16.4			
$t\bar{t}$ + Z	+8.1/-9.3			
Diboson				±3

[(1)]: includes flavour-scheme dependence

[(2)]: only dominant t channel considered

Table C.3: Inclusive cross-section uncertainties due to the PDF+α_s on the signal (top) and the background (bottom) processes for different initial states. Uncertainties for the same initial state (same column) for two different processes (different rows) are treated as correlated. Values are taken from Ref. [138, 139, 158, 161, 162].

	$gg_{t\bar{t}H}$ in %	qb_{tHq} in %	gb_{tHW} in %
$t\bar{t}H$	3.6		
tHq		3.7	
tHW			6.3

	gg in %	$q\bar{q}$ in %	gq in %
$t\bar{t}$	± 4.2		
single t [(1)]			± 2.8
W + jets		$-0.4/+0.8$	
Z + jets		± 0.2	
$t\bar{t}$ + W		± 3.6	
$t\bar{t}$ + Z	± 3.5		
Diboson		5	

[(1)]: only dominant t channel considered

Table C.4: Luminosity uncertainty per year. In the fits for the individual years, the luminosity uncertainty is implemented as shown in the first row. For the combined fit corresponding to 137.5 fb^{-1}, the uncertainties are partially correlated as recommended in Ref. [269]

Year	2016 in %	2017 in %	2018 in %
Uncorrelated 2016	2.2	0.0	0.0
Uncorrelated 2017	0.0	2.0	0.0
Uncorrelated 2018	0.0	0.0	1.5
Correlated between 2017 and 2018	0.0	0.6	0.2
Correlated between all years	0.6	0.9	2.0

D Architecture of ANNs

D.1 Inclusive $t\bar{t}H(b\bar{b})$ measurement

Table D.5: **Hyperparameters of Network Architecture**.

	$\geq$ 4j, 3t	$\geq$ 4j, $\geq$ 4t
Optimizer	Adagrad (learning rate = 0.01)	
Loss function	categorical cross-entropy	
Number of input variables	8	21
Hidden Layers	4	
Nodes per Hidden Layer	[2048,182,1024,64]	[1024,2048,512,512]
Node Activation	leakyelu	
Output Node activation	Softmax	
L2 regularisation	10^{-5}	10^{-5}
L1 regularisation	10^{-5}	10^{-5}
Dropout Percentage	0.5	0.5
early stopping min epochs	50	50
batchsize	2048	4096
trainable parameters	644 797	3 437 063

Table D.6: **Number of Training Events for Network Training**.

	$\geq$ 4j, 3t	$\geq$ 4j, $\geq$ 4t
events for training	3 786 298	1 248 927
events for validation	420 699	138 769
events for testing	420 699	138 769
events per process		
$t\bar{t}H$	2 654 442	1 170 335
tHq	253 217	107 151
tHW	148 823	36 707
$t\bar{t}b\bar{b}$/b	210 572	45 017
$t\bar{t}$ + 2b	62 936	5910
$t\bar{t} + c\bar{c}$	177 024	10 995
$t\bar{t}$ + lf	699 983	11 581

Table D.7: Input variables used in the ANNs in the SL channels in each category. Variables used in a given category are denoted by a "+" and unused variables by a "−".

Variable	Definition	≥ 4j, 3t	≥ 4j, ≥ 4t
MEM	matrix element method discriminant	+	+
⟨b-tag score⟩	average b-tagging value (DeepJet) of all jets	+	+
⟨b-tag score(tagged)⟩	average b-tagging value (DeepJet) of all tagged jets	−	+
b-tag score (jet #3)	third highest b-tagging value (DeepJet) amongst all jets	−	+
mean((b-tag score − ⟨b-tag score⟩)^2)	mean of squared difference between the b-tagging values (DeepJet) of jets and the average b-tagging value of all jets	+	−
BLR (transformed) = $\log\left[\frac{\mathrm{BLR}}{1-\mathrm{BLR}}\right]$	transformed likelihood ratio discriminating between events with 4 b-quark jets and 2 b-quark jets	−	+
$\left\langle \Delta\eta^{\mathrm{jets}} \right\rangle$	average $\Delta\eta$ between two jets	−	+
$\left\langle \Delta\eta^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average $\Delta\eta$ between two b-tagged jets	−	+
$\left\langle \mathrm{m}^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average mass of tagged jets	−	+
$\left\langle \mathrm{m}^{\mathrm{jets}} \right\rangle$	average mass of jets	+	−
m(min(ΔR^{jets}(tagged)))	mass of pair of two b-tagged jets closest in ΔR	−	+
$\left\langle p_{\mathrm{T}}^{\mathrm{jets}} \right\rangle$	average p_{T} of all jets	−	+
$\left\langle p_{\mathrm{T}}^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average p_{T} of all tagged jets	−	+
p_{T}(min(ΔR^{jets}(tagged)))	sum p_{T} of pair of closest b-tagged jets	−	+
$\sum p_{\mathrm{T}}^{\mathrm{jets}}$	sum of all jet p_{T}	+	−
$\sum p_{\mathrm{T}}^{\mathrm{jets}}$(tagged)	sum of all b-tagged jet p_{T}	−	+
N(Jets)	number of reconstructed jets	−	+
b-tag score($\mathrm{b_t}$)	b-tagging value (DeepJet) of b-jet of top quark decay from reconstruction of tHW system with JA-BDTs	+	+
$\frac{p_{\mathrm{T}}(\mathrm{t})+p_{\mathrm{T}}(\mathrm{H})+p_{\mathrm{T}}(\mathrm{Wb})}{H_{\mathrm{T}}+p_{\mathrm{T}}(MET)+p_{\mathrm{T}}(\mathrm{lep})}$	transverse momentum fraction assigned to the tHW process divided by the sum of all transverse momenta in the event from reconstruction of tHW system with JA-BDTs	−	+
$\lvert\eta^{\mathrm{lep+jets}}\rvert$	$\lvert\eta\rvert$ of forward jet from reconstruction of tHq system with JA-BDTs	−	+
m($\mathrm{t_{lep}}$)	reconstructed mass of leptonically decaying top quark of $\mathrm{t\bar{t}H}$ system with JA-BDT	−	+
Reconstruction BDT (tHq)	output value of JA-BDT for reconstruction of tHq system	+	+
Reconstruction BDT ($\mathrm{t\bar{t}H}$)	output value of JA-BDT for reconstruction of $\mathrm{t\bar{t}H}$ system	+	+
Reconstruction BDT ($\mathrm{t\bar{t}}$)	output value of JA-BDT for reconstruction of $\mathrm{t\bar{t}}$ system	−	+

D.2 STXS cross section measurement

Table D.8: **Hyperparameters of Network Architecture for STXS cross section measurement**.

	≥ 4j, 3t	≥ 4j, ≥ 4t
Optimizer	Adagrad(learning rate = 0.01)	
Loss function	categorical cross-entropy	
Number of input variables	12	16
Hidden Layers	4	3
Nodes per Hidden Layer	[2048,2048,512,1024]	[2048,2048,64]
Node Activation	leakyelu	
Output Node activation	Softmax	
L2 regularisation	1×10^{-5}	1×10^{-5}
L1 regularisation	1×10^{-5}	1×10^{-5}
Dropout Percentage	0.5	0.5
early stopping min epochs	50	50
batchsize	2048	1024
trainable parameters	5 802 501	4 362 629

Table D.9: **Number of Training Events for Network Training for STXS cross section measurement**.

	$\geq$ 4j, 3t	$\geq$ 4j, $\geq$ 4t
events for training	297 355	199 306
events for validation	74 338	49 826
events for testing	74 338	49 826
events per process		
$p_{\mathrm{T}}^{\mathrm{H}} \leq 60$	60 425	39 552
$60 \leq p_{\mathrm{T}}^{\mathrm{H}} \leq 120$	112 765	72 628
$120 \leq p_{\mathrm{T}}^{\mathrm{H}} \leq 200$	106 675	71 645
$200 \leq p_{\mathrm{T}}^{\mathrm{H}} \leq 300$	58 958	41 190
$300 \leq p_{\mathrm{T}}^{\mathrm{H}}$	32 870	24 117

Table D.10: **Input features for the STXS network**. Listed are all observables used for the training of the ANNs for the cross section measurement.

Variable	Definition	$\geq$ 4j, 3t	$\geq$ 4j, $\geq$ 4t
MEM	matrix element method discriminant	+	+
$\left\langle p_{\mathrm{T}}^{\mathrm{jets}} \right\rangle$	average p_{T} of all jets	+	+
$\left\langle p_{\mathrm{T}}^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average p_{T} of all b-tagged jets	+	+
$\left\langle \Delta \mathrm{R}^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average Δ R of all b-tagged jets	+	+
$\left\langle \mathrm{m}^{\mathrm{jets}}(\mathrm{tagged}) \right\rangle$	average mass of all b-tagged jets	−	+
$p_{\mathrm{T}}(\min(\Delta R^{\mathrm{jets}}(\mathrm{tagged})))$	sum p_{T} of pair of closest b-tagged jets	−	+
$\mathrm{t\bar{t}H}$ reco $\Delta R(\mathrm{H})$	ΔR of Higgs decay products from reconstruction of $\mathrm{t\bar{t}H}$ system with JA-BDTs	+	+
$\mathrm{t\bar{t}H}$ reco $\mathrm{m_H}$	reconstructed mass of Higgs from reconstruction of $\mathrm{t\bar{t}H}$ system with JA-BDTs	−	+
$\mathrm{t\bar{t}H}$ reco $\log(p_{\mathrm{T}}^{\mathrm{H}})$	logarithm of p_{T} of the reconstructed Higgs from reconstruction of $\mathrm{t\bar{t}H}$ system with JA-BDTs	−	+
$\mathrm{t\bar{t}H}$ reco $\log(\mathrm{m}(\mathrm{t_{lep}}))$	logarithm mass of leptonically decaying top quark from reconstruction of $\mathrm{t\bar{t}H}$ system with JA-BDTs	−	+
tHq reco $p_{\mathrm{T}}^{\mathrm{H}}$	p_{T} of Higgs from reconstruction of tHq system with JA-BDTs	+	+
tHq reco $\log(p_{\mathrm{T}}^{\mathrm{H}})$	log of p_{T} of Higgs from reconstruction of tHq system with JA-BDTs	+	+
tHq reco $\log(\min(p_{\mathrm{T},1}^{\mathrm{H\ dec\ prod}}, p_{\mathrm{T},2}^{\mathrm{H\ dec\ prod}}))$	log of p_{T} of softer Higgs decay product from reconstruction of tHq system with JA-BDTs	+	+
tHq reco $\log(\mathrm{m_H})$	log of mass of Higgs from reconstruction of tHq system with JA-BDTs	+	−
tHW reco $p_{\mathrm{T}}^{\mathrm{H}}$	p_{T} of Higgs from reconstruction of tHW system with JA-BDTs	+	+
tHW reco $\log(p_{\mathrm{T}}^{\mathrm{H}})$	log of p_{T} of Higgs from reconstruction of tHW system with JA-BDTs	+	+
tHW reco $\Delta R(\mathrm{H})$	ΔR of Higgs decay products from reconstruction of tHW system with JA-BDTs	−	+
tHW reco $\log(\mathrm{m_H})$	log of mass of Higgs from reconstruction of tHW system with JA-BDTs	+	−

E Input feature validation

E.1 GoF Summary

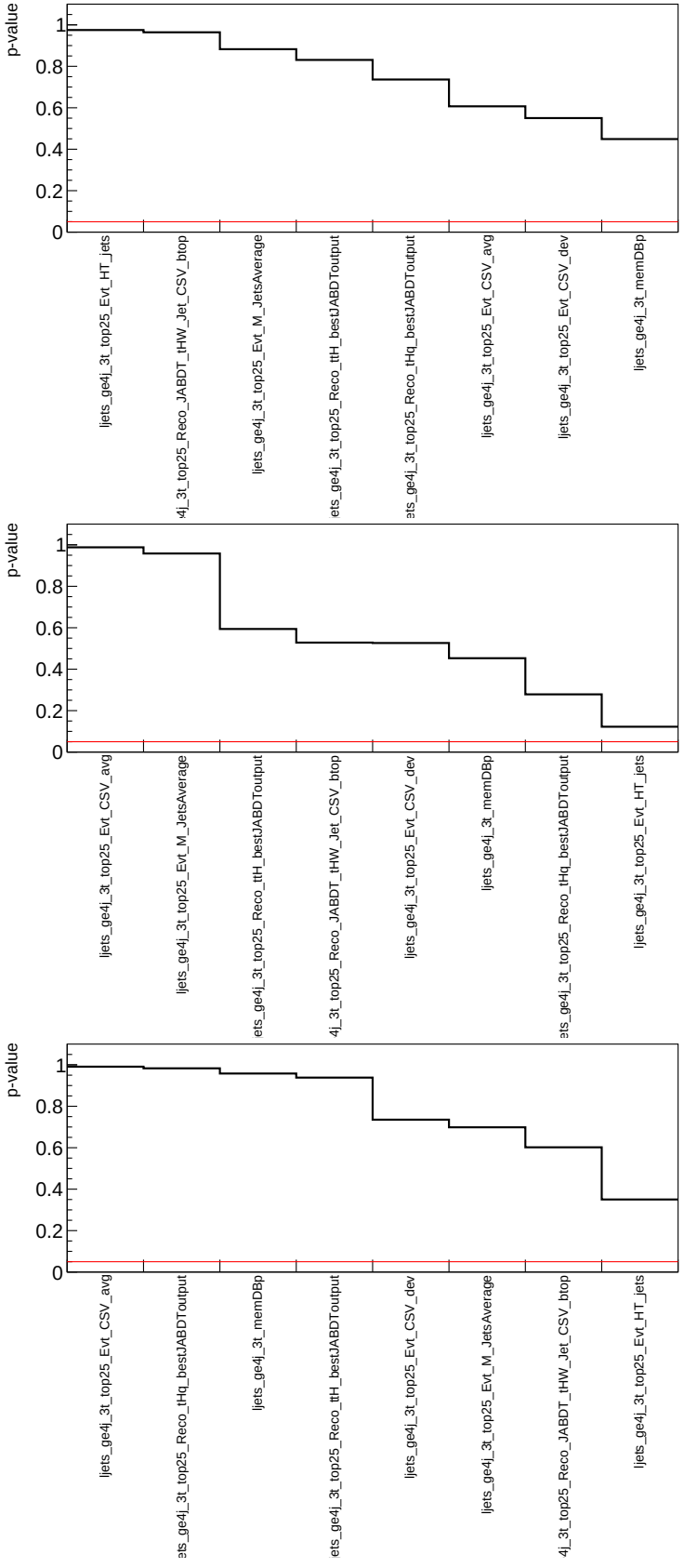


Figure E.5: **Summary of GoF Tests in 1D in $\geq$ 4j, 3t.** The figure shows the p values calculated for the $\geq$ 4j, 3t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. The red horizontal line at 5 % indicates the exclusion limit for the validation. Observables showing a p value below the threshold are excluded from the analysis and are not considered further.

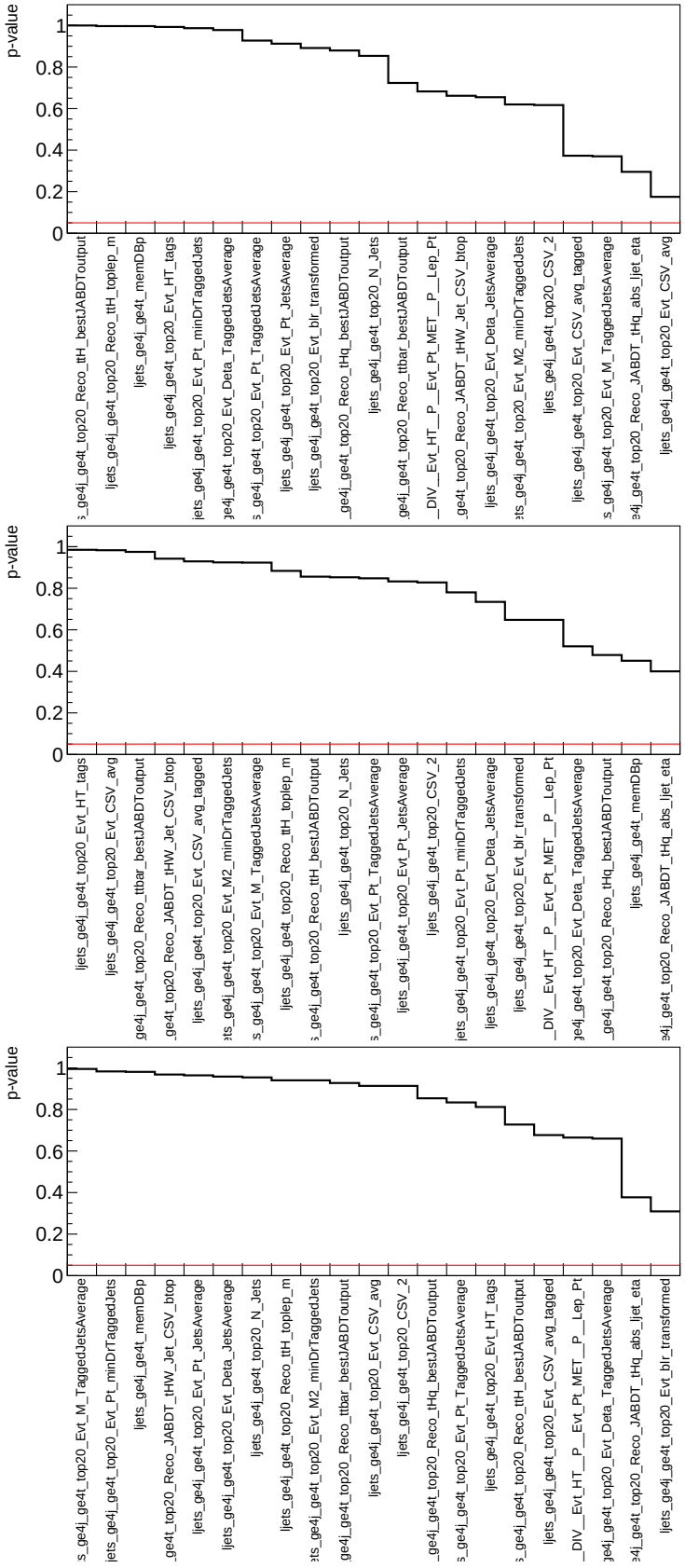


Figure E.6: **Summary of GoF Tests in 1D in $\geq$ 4j, $\geq$ 4t.** The figure shows the p values calculated for the $\geq$ 4j, $\geq$ 4t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. The red horizontal line at 5 % indicates the exclusion limit for the validation. Observables showing a p value below the threshold are excluded from the analysis and are not considered further.

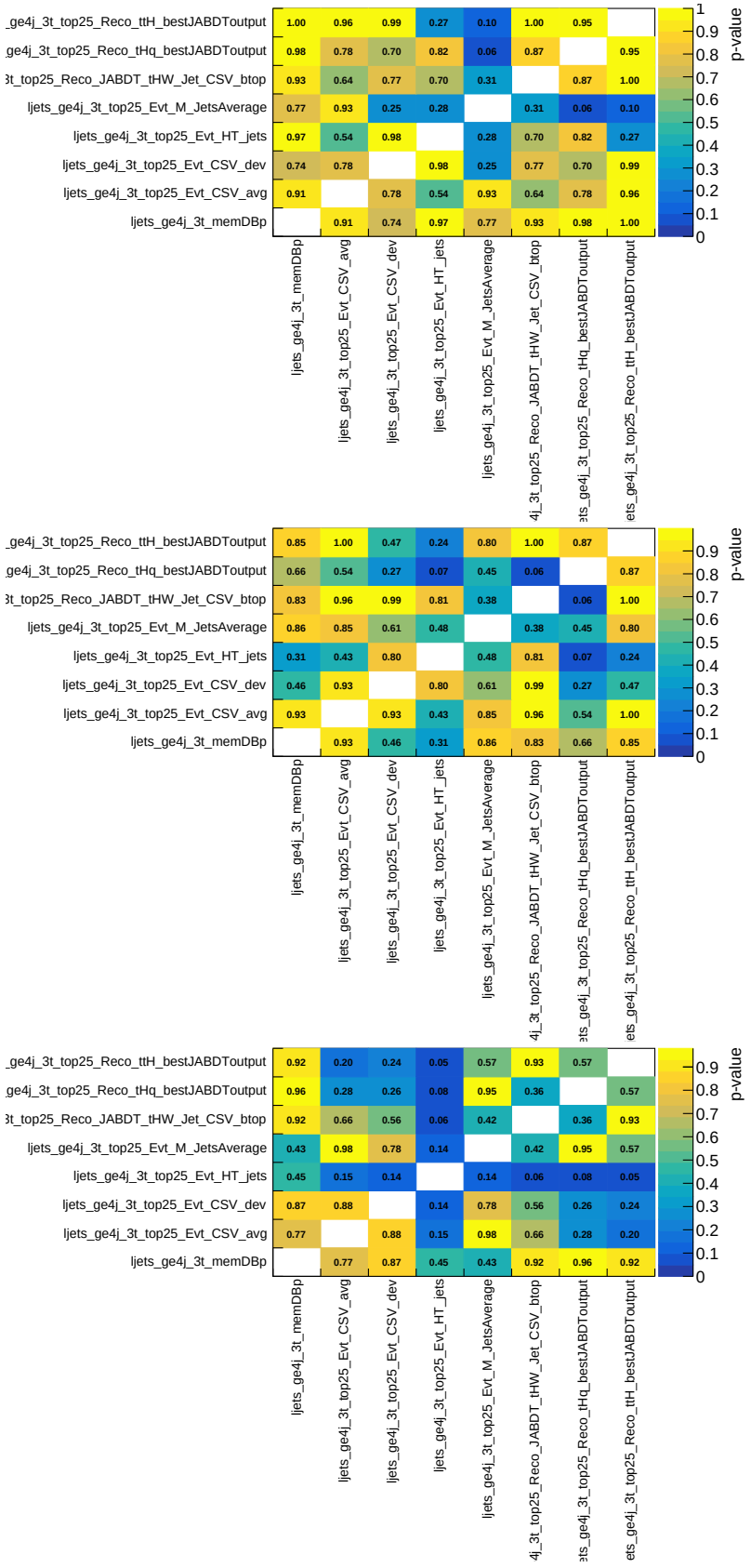


Figure E.7: **Summary of GoF Tests in 2D in $\geq$ 4j, 3t.** The figure shows the p values calculated for the correlations of the $\geq$ 4j, 3t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. Observables with correlations showing a p value below the 5 % threshold are excluded from the analysis and are not considered further.

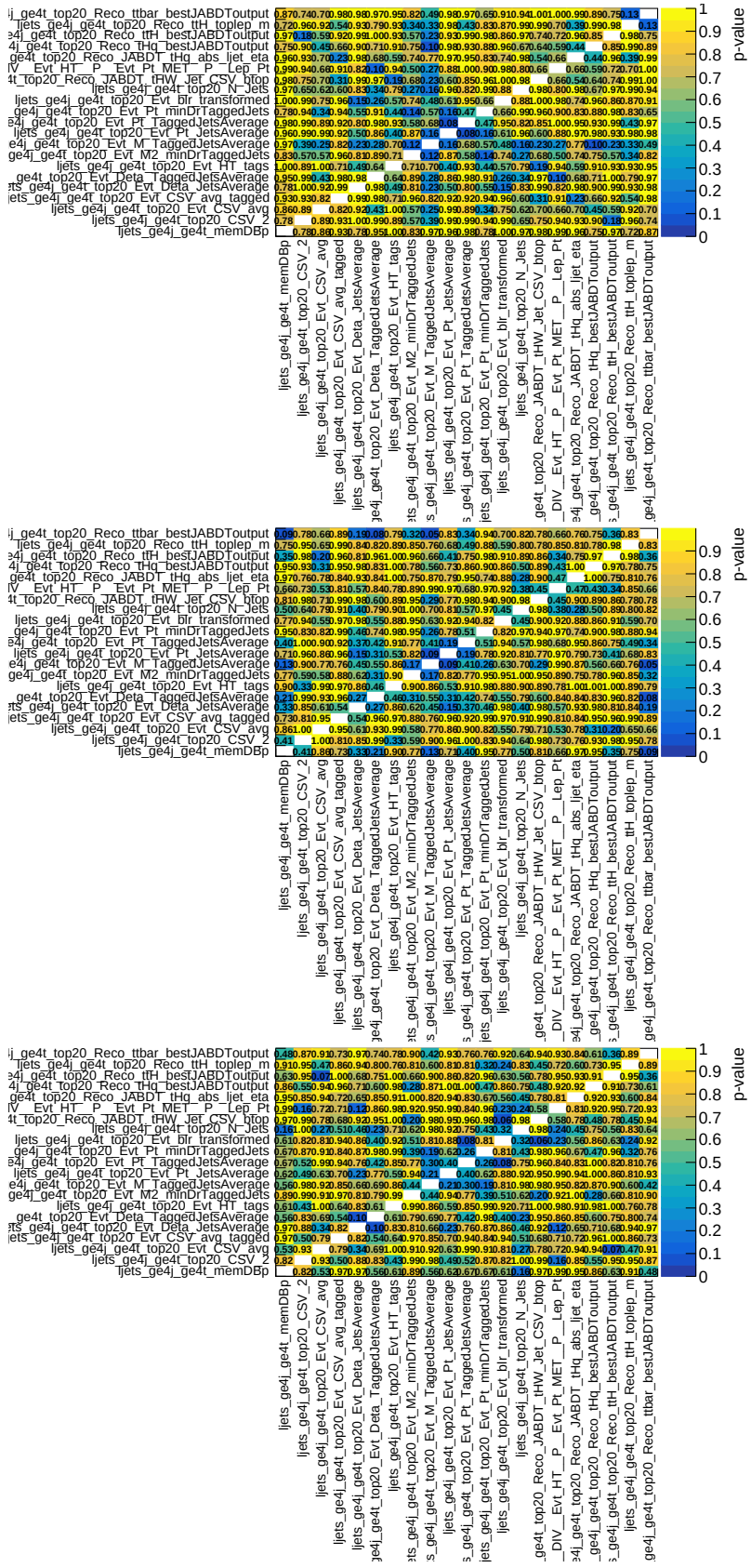


Figure E.8: **Summary of GoF Tests in 2D in $\geq$ 4j, $\geq$ 4t.** The figure shows the p values calculated for the correlations of the $\geq$ 4j, $\geq$ 4t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. Observables with correlations showing a p value below the 5 % threshold are excluded from the analysis and are not considered further.

F Input feature validation for the cross section measurement

F.1 GoF Summary

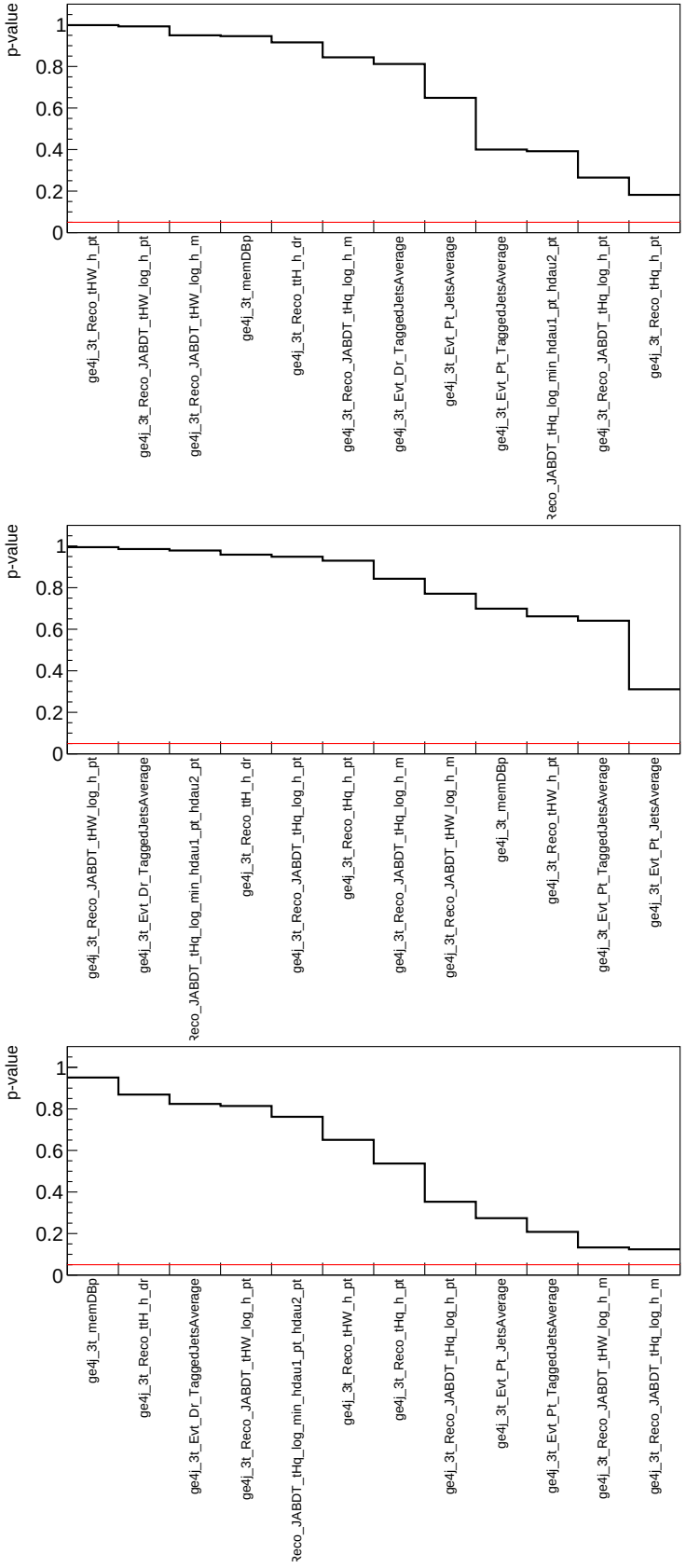


Figure F.9: **Summary of GoF Tests in 1D STXS in $\geq$ 4j, 3t.** The figure shows the p values calculated for the $\geq$ 4j, 3t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. The red horizontal line at 5 % indicates the exclusion limit for the validation. Observables showing a p value below the threshold are excluded from the analysis and are not considered further.

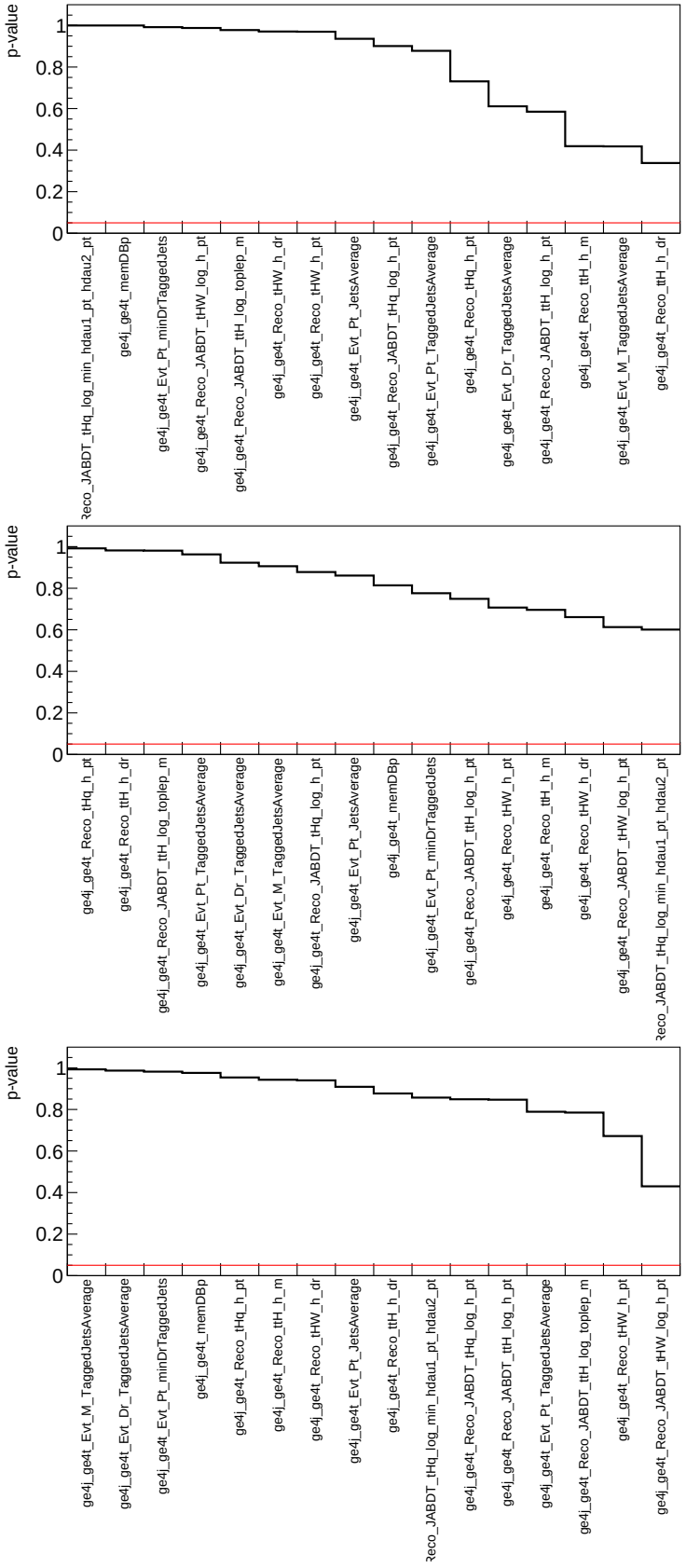


Figure F.10: **Summary of GoF Tests in 1D STXS in $\geq$ 4j, $\geq$ 4t.** The figure shows the p values calculated for the $\geq$ 4j, $\geq$ 4t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. The red horizontal line at 5 % indicates the exclusion limit for the validation. Observables showing a p value below the threshold are excluded from the analysis and are not considered further.

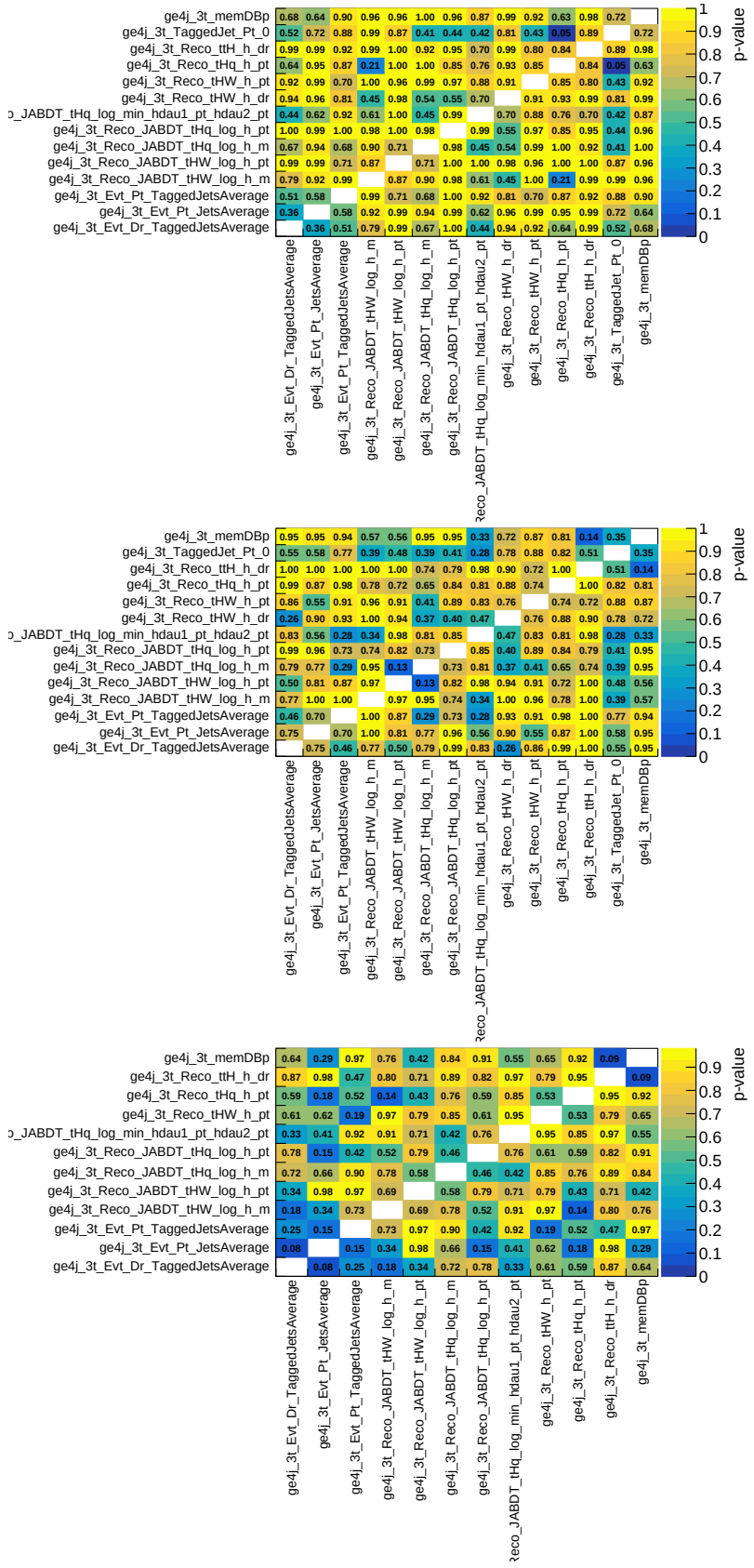


Figure F.11: **Summary of GoF Tests in 2D STXS in $\geq$ 4j, 3t.** The figure shows the p values calculated for the correlations of the $\geq$ 4j, 3t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. Observables with correlations showing a p value below the 5 % threshold are excluded from the analysis and are not considered further.

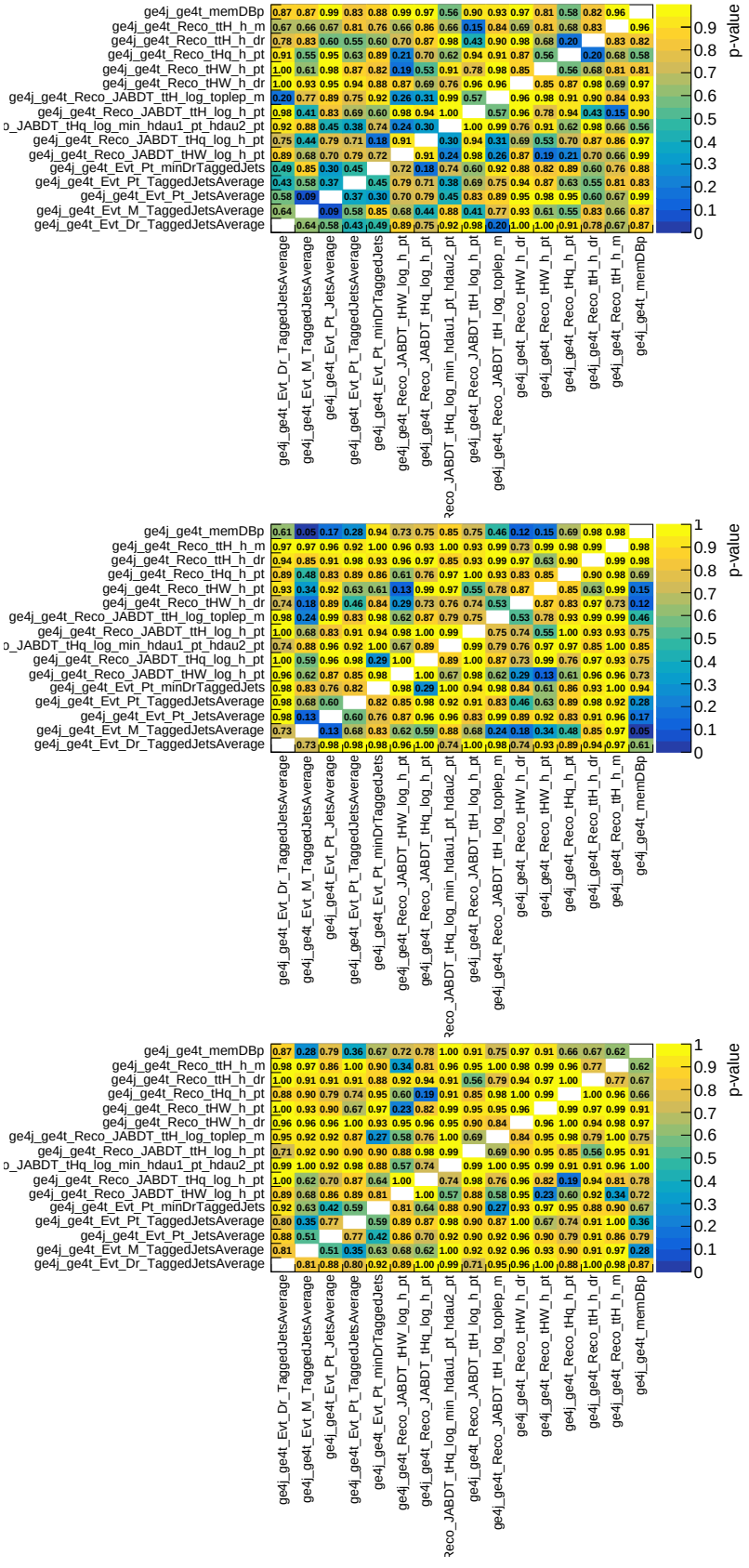


Figure F.12: **Summary of GoF Tests in 2D STXS in $\geq$ 4j, $\geq$ 4t.** The figure shows the p values calculated for the correlations of the $\geq$ 4j, $\geq$ 4t ANN input features for 2016 (top), 2017 (middle) and 2018 (bottom) data. Observables with correlations showing a p value below the 5 % threshold are excluded from the analysis and are not considered further.

Table G.11: **Comparison of the $S/\sqrt{B}$ of the discriminant distributions**. The first row shows the number of bins of the final distributions based directly on the $t\bar{t}H$ and $t\bar{t}mb$ nodes (middle column) and using the ratio observable as defined in Eq. (8.2) in the $\geq$ 4j, $\geq$ 4t region. All years of data taking are considered. The subsequent rows show the largest values for the $S/\sqrt{B}$ metric. The final row displays the average $S/\sqrt{B}$ of the bins.

	DNN output	Ratio Observable
$\sum_{\text{years}}$ Bins	87	56
highest $S/\sqrt{B}$	2.46	1.60
second $S/\sqrt{B}$	1.89	1.60
third $S/\sqrt{B}$	1.56	1.51
forth $S/\sqrt{B}$	1.45	1.41
fifth $S/\sqrt{B}$	1.14	1.27
$\left\langle S/\sqrt{B} \right\rangle$	0.50	0.83

G Additional tests for the ratio observable

This section briefly summarizes the tests performed to validated the ratio observable introduced in Sec. 8.2. All tests are based on the parameter estimation using the test statistic introduced in Chap. 6 in a fit to pseudo data corresponding to the SM prediction, which is referred to as the Asimov data set in the following. The following metrics are used to validate the impact of this observable on the final sensitivity:

- comparison of the signal-over-square-root-of-background ratio ($S/\sqrt{B}$) in the bins of the final distributions
- parameter estimation of the signal strength modifier $\mu_{t\bar{t}H}$
- computation of the expected significance with respect to the hypothesis excluding the $t\bar{t}H(b\bar{b})$ process (background-only hypothesis)
- comparison of the impacts on the signal inference as explained in Sec. 8.3.2
- comparison of the constraints placed on the nuisance parameters
- bias tests with different $t\bar{t} + b\bar{b}$ hypotheses in Asimov data and dedicated pseudo experiments generated from the corresponding statistical model.

These tests are further discussed in the following. The results presented in this section were used to finalize the strategy employed for the semileptonic $t\bar{t}H(b\bar{b})$ analysis. However, the statistical model was further optimized since then. Specifically, the binning of the distributions used in the final statistical model is coarser. Additionally, the purity uncertainties for the heavy- and light-flavor b-tagging scale factors as well as the charm-quark-related b-tagging uncertainties are decorrelated between 2016 and the other years in the final statistical model to account for the upgrade of the pixel detector in 2017. Therefore, the results presented in this section do not match exactly with the results presented in Sec. 8.3. However, the differences between the versions are small and the observations presented in this section are applicable to the final version of the analysis.

Tab. G.11 shows the comparison of the $S/\sqrt{B}$ metric, which quantifies the compatibility of the signal contribution in a given bin with statistical fluctuations expected for the sum of the background yields. The distributions constructed from the $t\bar{t}H$ and $t\bar{t}mb$ nodes,

Table G.12: **Comparison of the expected sensitivity to the signal strength modifier of statistical models with and without the ratio observable**. Shown are the post-fit estimations for $\mu_{t\bar{t}H}$ together with the total (second column) and the statistical uncertainty (third column). The last column displays the expected significance with respect to the background-only hypothesis.

Fitted Observables	Stat+Syst	Stat-Only	Significance
Pure ANN output	1.0 +0.37/-0.35	1.0 +0.21/-0.21	2.9
Ratio Observable ($t\bar{t}H$ + $t\bar{t}mb$ merged in both analysis regions)	1.0 +0.33/-0.29	1.0 +0.21/-0.21	3.6
Ratio Observable ($t\bar{t}H$ + $t\bar{t}mb$ merged in $\geq$ 4j, $\geq$ 4t)	1.0 +0.32/-0.29	1.0 +0.21/-0.20	3.7

respectively, have two bins with significantly larger $S/\sqrt{B}$ values. However, the distributions of the ratio observable yield larger $S/\sqrt{B}$ on average, which means that these distributions have more bins with a higher signal purity in total.

Tab. G.12 shows the expected sensitivity obtained using distributions of the separate ANN nodes or the ratio observables. The total uncertainty in the parameter estimation decreases by approximately 14 % when using the ratio observable in both analysis regions or only in the $\geq$ 4j, $\geq$ 4t category. The statistical uncertainty is the same in all statistical models used for the signal inference. However, since the models using the ratio observables consider fewer bins for the parameter estimation, the discrimination power of the distributions of the ratio observable must be larger. This is in line with the behavior observed in the discussion of the $S/\sqrt{B}$ metric. The improvement of the total uncertainty in the parameter estimation must therefore originate from a better control over the systematic uncertainties. The largest expected significance is obtained using the ratio observable in the $\geq$ 4j, $\geq$ 4t region.

Fig. G.13 shows a comparison of the ranking based on the impact of the nuisance parameters on the signal inference. Overall, the signal-strength estimation in the model based on the ratio observable is less susceptible to the nuisance parameters, which is indicated by the smaller order of magnitude of the impacts compared to the statistical model based directly on the ANN discrimination. Moreover, the number of systematic uncertainties among the 30 parameters with the largest impact is smaller in the estimation based on the ratio observable. The ranking is dominated by uncertainties modeling the effects of the limited size of the simulated samples as well as modeling uncertainties on the signal processes and the dominant $t\bar{t}$ + jets background. Given the previously-discussed large discrimination power of the distribution of the ratio observable, the large impact of the parameters directly modifying the corresponding bin contents is reasonable. Moreover, the high placement of the modeling uncertainties in the impact ranking is plausible and expected, as discussed in Sec. 8.3.2. Consequently, the impact ranking indicates a better control over the systematic uncertainties other than those related to the limited size of the simulated samples, thus reducing their impact on the signal inference.

Table G.13 shows the comparison of the constraints extracted from the estimation of the nuisance parameters. The constraints show a reduction of up to 45 %, which originates from the coarser binning of the distributions of the ratio observable. The reduction of the constraints indicates a greater flexibility of the statistical model, which is desired to facilitate the mitigation of differences between the simulated samples and observed data. Moreover, the usage of fewer bins can help to avoid spurious constraints of the parameters that can arise from the statistical fluctuation of the distributions used to construct them.

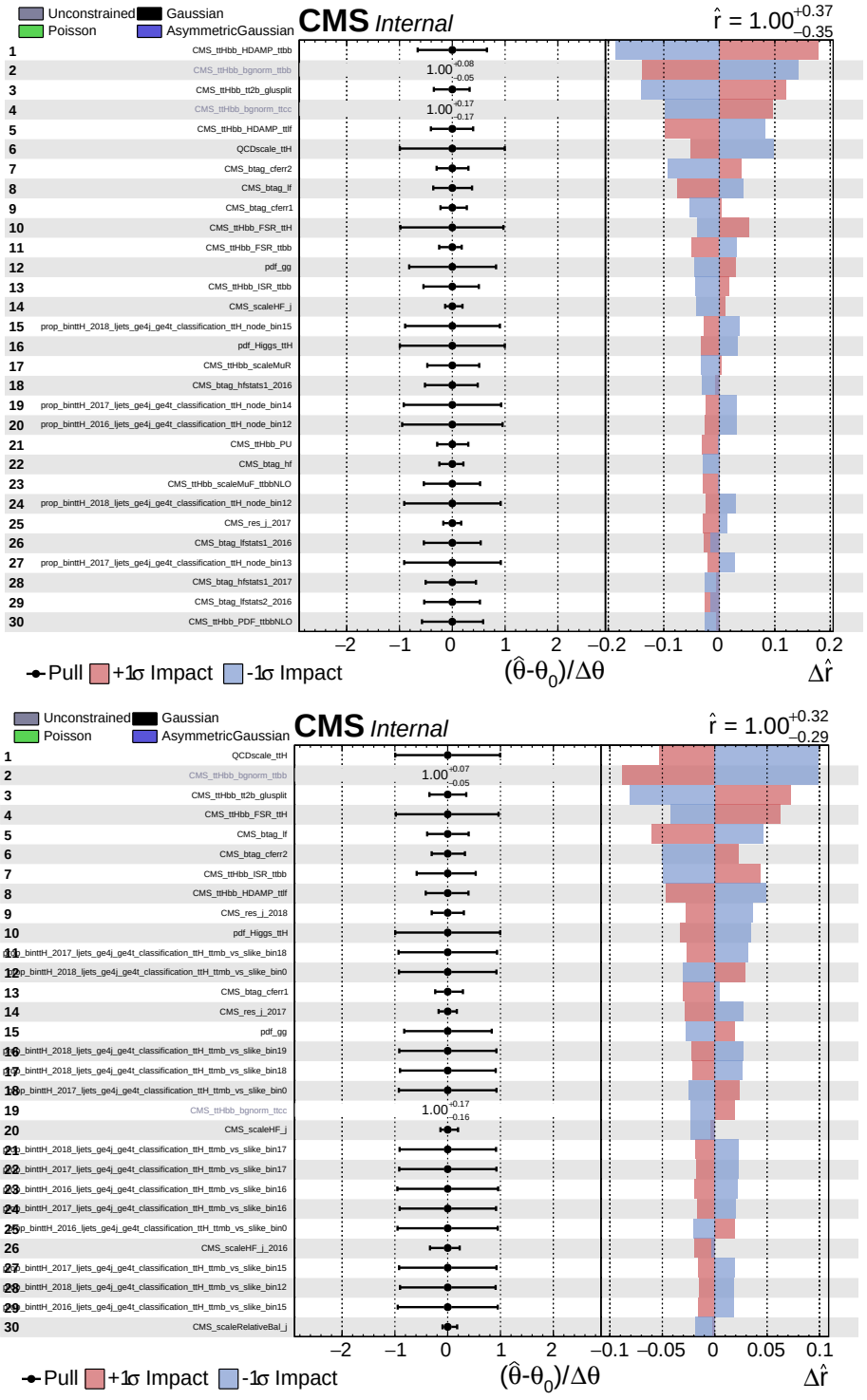


Figure G.13: **Comparison of impacts of nuisance parameters on signal inference**. Shown are the rankings based on the impacts of the nuisance parameters on the signal inference in a statistical model based directly on the ANN discriminant distributions (top) and the ratio observable in the $\geq$ 4j, $\geq$ 4t region (bottom).

Table G.13: **Comparison of the parameter constraints originating from the parameter estimation of $\mu_{\mathbf{t\bar{t}H}}$.** Shown are the 15 parameters with the largest relative difference of the constraints estimated using a statistical model based directly on the ANN discriminants (second column) and the ratio observable in the $\geq$ 4j, $\geq$ 4t (third column). The relative difference in the last column is calculated with respect to the values in the second column.

Parameter	Pure ANN Discr.	Ratio Observable	Difference	Rel. Difference
CMS_effTrigger_m_2018	$^{+0.673}_{-0.67}$	$^{+0.976}_{-0.974}$	$^{+0.303}_{-0.304}$	0.45
CMS_eff_m_2018	$^{+0.673}_{-0.671}$	$^{+0.977}_{-0.974}$	$^{+0.303}_{-0.303}$	0.45
CMS_effTrigger_e_2018	$^{+0.666}_{-0.671}$	$^{+0.968}_{-0.964}$	$^{+0.302}_{-0.293}$	0.45
CMS_btag_lfstats1_2018	$^{+0.679}_{-0.649}$	$^{+0.948}_{-0.96}$	$^{+0.269}_{-0.31}$	0.44
CMS_eff_e_2018	$^{+0.656}_{-0.67}$	$^{+0.945}_{-0.951}$	$^{+0.289}_{-0.281}$	0.43
CMS_btag_lfstats2_2018	$^{+0.635}_{-0.604}$	$^{+0.838}_{-0.834}$	$^{+0.203}_{-0.23}$	0.35
CMS_btag_hfstats1_2018	$^{+0.586}_{-0.648}$	$^{+0.808}_{-0.817}$	$^{+0.223}_{-0.169}$	0.32
CMS_btag_hfstats2_2018	$^{+0.576}_{-0.606}$	$^{+0.791}_{-0.764}$	$^{+0.215}_{-0.158}$	0.32
CMS_scaleHEM1516_j	$^{+0.613}_{-0.36}$	$^{+0.839}_{-0.387}$	$^{+0.226}_{-0.027}$	0.22
CMS_res_j_2018	$^{+0.277}_{-0.268}$	$^{+0.306}_{-0.302}$	$^{+0.029}_{-0.034}$	0.12
CMS_ttHbb_HDAMP_ttbb	$^{+0.656}_{-0.655}$	$^{+0.699}_{-0.752}$	$^{+0.043}_{-0.097}$	0.11
CMS_scaleHF_j_2018	$^{+0.369}_{-0.414}$	$^{+0.393}_{-0.461}$	$^{+0.025}_{-0.047}$	0.09
CMS_scaleEC2_j_2018	$^{+0.4}_{-0.37}$	$^{+0.444}_{-0.395}$	$^{+0.044}_{-0.025}$	0.09
CMS_btag_lfstats1_2017	$^{+0.502}_{-0.492}$	$^{+0.537}_{-0.541}$	$^{+0.035}_{-0.049}$	0.08
CMS_eff_m_2017	$^{+0.498}_{-0.5}$	$^{+0.54}_{-0.541}$	$^{+0.042}_{-0.041}$	0.08

However, a coarser binning of the distributions of the ANN discriminants in the $t\bar{t}H$ and $t\bar{t}mb$ node also results in a larger contribution of the background processes in the individual bins and therefore to a loss in sensitivity to the $t\bar{t}H(b\bar{b})$ process. The statistical model based on the ratio observable is able to profit from the positive effects of usaging fewer bins without losing sensitivity to the signal strength modifier, which is generally desirable.

The stability of the signal strength estimation is estimated by performing the fit to pseudo data that is generated based on the Asimov data set considering the statistical fluctuations expected from Poisson statistics. Two sets of pseudo experiments are tested. The first set uses the $t\bar{t} + b\bar{b}$ processes as it is predicted in the simulation. The second set tests employs a $t\bar{t} + b\bar{b}$ prediction that is increased by 30 %, which corresponds to the latest measurements of the $t\bar{t} + b\bar{b}$ process [253, 261] and is expected to yield a more realistic test scenario. The injected $t\bar{t}H(b\bar{b})$ signal strength used for the generation of the pseudo experiments is one. The results are shown in Tab. G.14. The injected signal strength value of one is covered within the 68 % confidence interval in the mean estimated signal strength in both sets of pseudo experiments. The same is true for the parameter estimation performed with the statistical model based on the ratio observable. However, the central values of the estimated signal strength is much closer to the true injected value used for the generation of the pseudo experiments. This indicates that the signal strength estimation in the statistical model employing the ratio observable is more stable under statistical fluctuation.

Table G.14: **Comparison of the signal strength estimation in fits to pseudo data**. Shown are the mean values of the estimated signal strength in fits to 400 pseudo experiments generated from the statistical model specified in the first column. Additionally, the signal strength estimation is performed with different $t\bar{t} + b\bar{b}$ hypotheses as indicated by the corresponding rows. For convenience, the estimation in a fit to the Asimov data set is shown in the third column.

Statistical Model	$t\bar{t} + b\bar{b}$ Hypothesis	Fit to Asimov	Fit to pseudo data
Pure ANN discr.	$t\bar{t} + b\bar{b}$ (4FS)	1.00 ± 0.36	0.73 ± 0.33
	$t\bar{t} + b\bar{b}$ (4FS) $\times 1.3$	1.00 ± 0.41	0.67 ± 0.37
Ratio Observable	$t\bar{t} + b\bar{b}$ (4FS)	1.00 ± 0.30	0.85 ± 0.28
	$t\bar{t} + b\bar{b}$ (4FS) $\times 1.3$	1.00 ± 0.34	0.82 ± 0.31

To further quantify the stability, the bias b in the signal strength estimation is calculated for each pseudo experiment following

$$b = \frac{\mu^{\text{post-fit}} - \mu^{\text{injected}}}{\sigma_{\mu}^{\text{post-fit}}}. \tag{12.1}$$

The injected signal strength μ^{injected} is one and $\sigma_{\mu}^{\text{post-fit}}$ denotes the uncertainty on the estimated signal strength modifier. Consequently, this metric is zero if the parameter estimation works perfectly. The distributions of this quantity in the fits to the different sets of pseudo experiments are shown in Fig. G.14. The ideal value of zero is not compatible within the 68 % confidence interval in the parameter estimations using the statistical model based directly on the ANN discriminants. In the parameter estimation employing the ratio observable, the ideal value is indeed compatible within the 68 % interval. Moreover, the mean value of the expected bias in the signal strength estimation decreases by 23 % in the fit to pseudo experiments based on the nominal $t\bar{t} + b\bar{b}$ prediction and by 28 % in the fits to the scaled $t\bar{t} + b\bar{b}$ scenario. Additionally, the mean expected bias values between the two sets of pseudo experiments are compatible within the uncertainty of gaussian fit, which is not the case in the statistical model based directly on the ANN discriminants. This further indicates that the parameter estimation employing the ratio observable is more robust against statistical fluctuation, which is why this statistical model is used in the semileptonic $t\bar{t}H(b\bar{b})$ analysis.

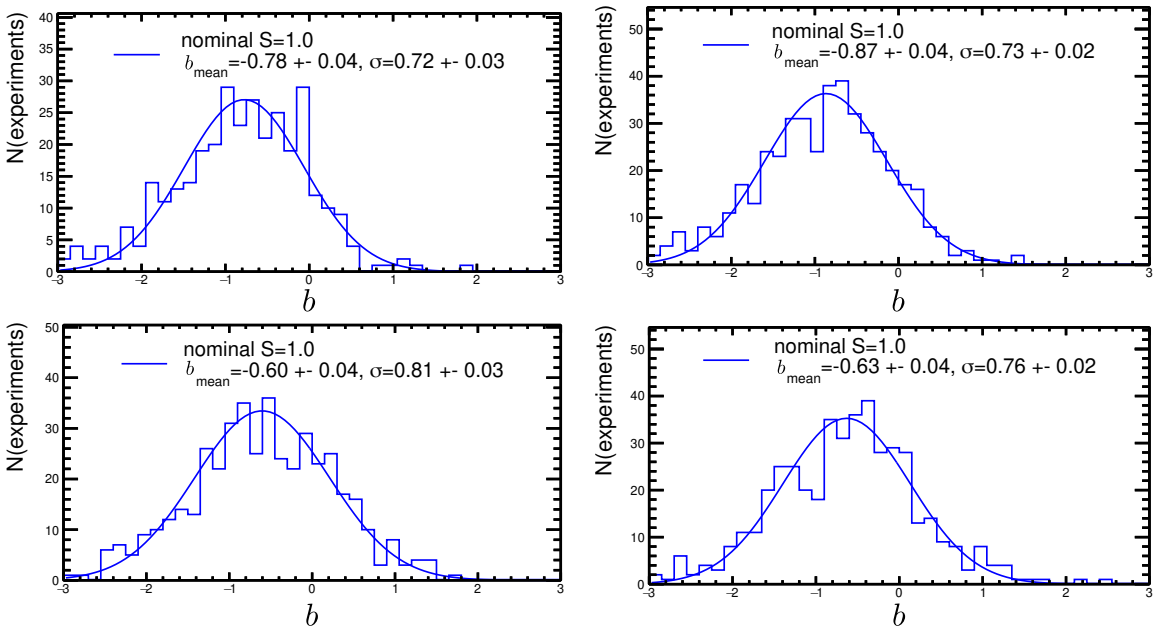


Figure G.14: **Distribution of the bias in fits to pseudo data**. The x-axis shows the bias defined in Eq. (12.1). The pseudo data is generated from the statistical model based directly on the ANN discriminants (top) and the ratio observable in the $\geq$ 4j, $\geq$ 4t region (bottom). Two sets of $t\bar{t} + b\bar{b}$ hypotheses are tested, where the $t\bar{t} + b\bar{b}$ hypothesis implemented in the corresponding statistical model is used as it is (left) or scaled by 30 % (right). The estimations for the mean bias and the standard deviation of the respective distribution are extracted from a gaussian fit (blue line) to the distributions.

H Supplementary results

H.1 Inclusive $t\bar{t}H(b\bar{b})$ analysis

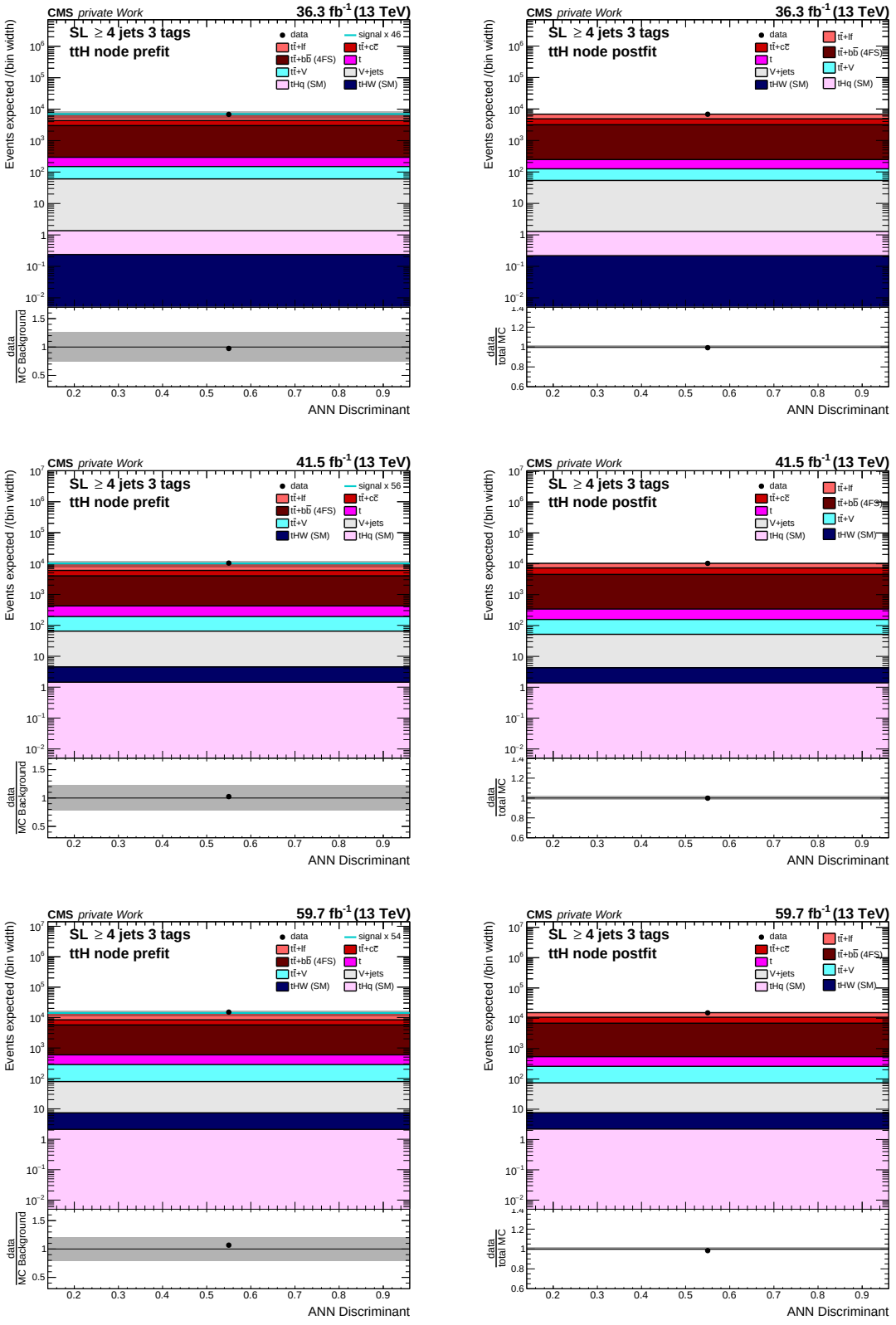


Figure H.15: **Discriminants in the SL channel in the ≥ 4j, 3t category $t\bar{t}H$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

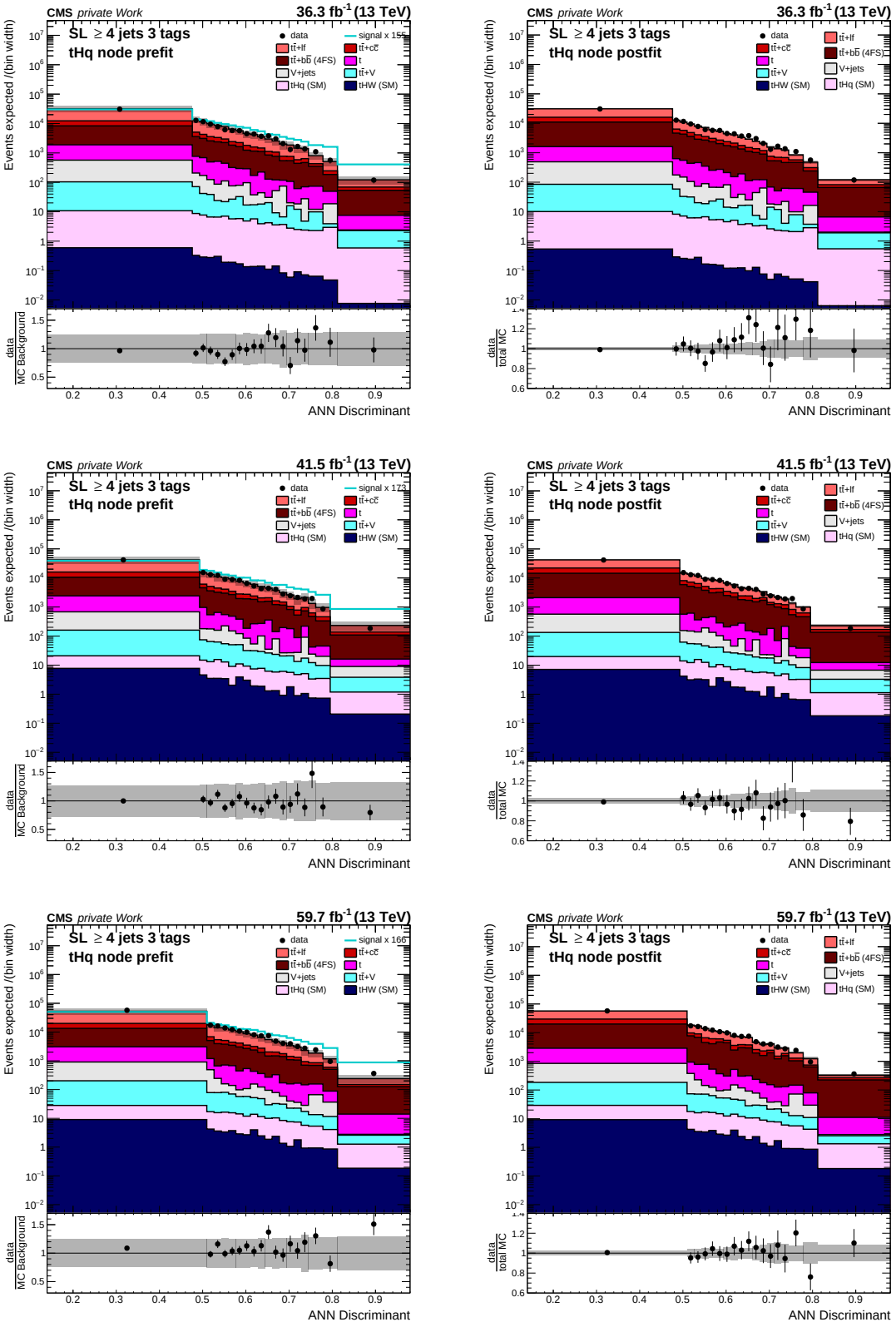


Figure H.16: **Discriminants in the SL channel in the ≥ 4j, 3t category tHq node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

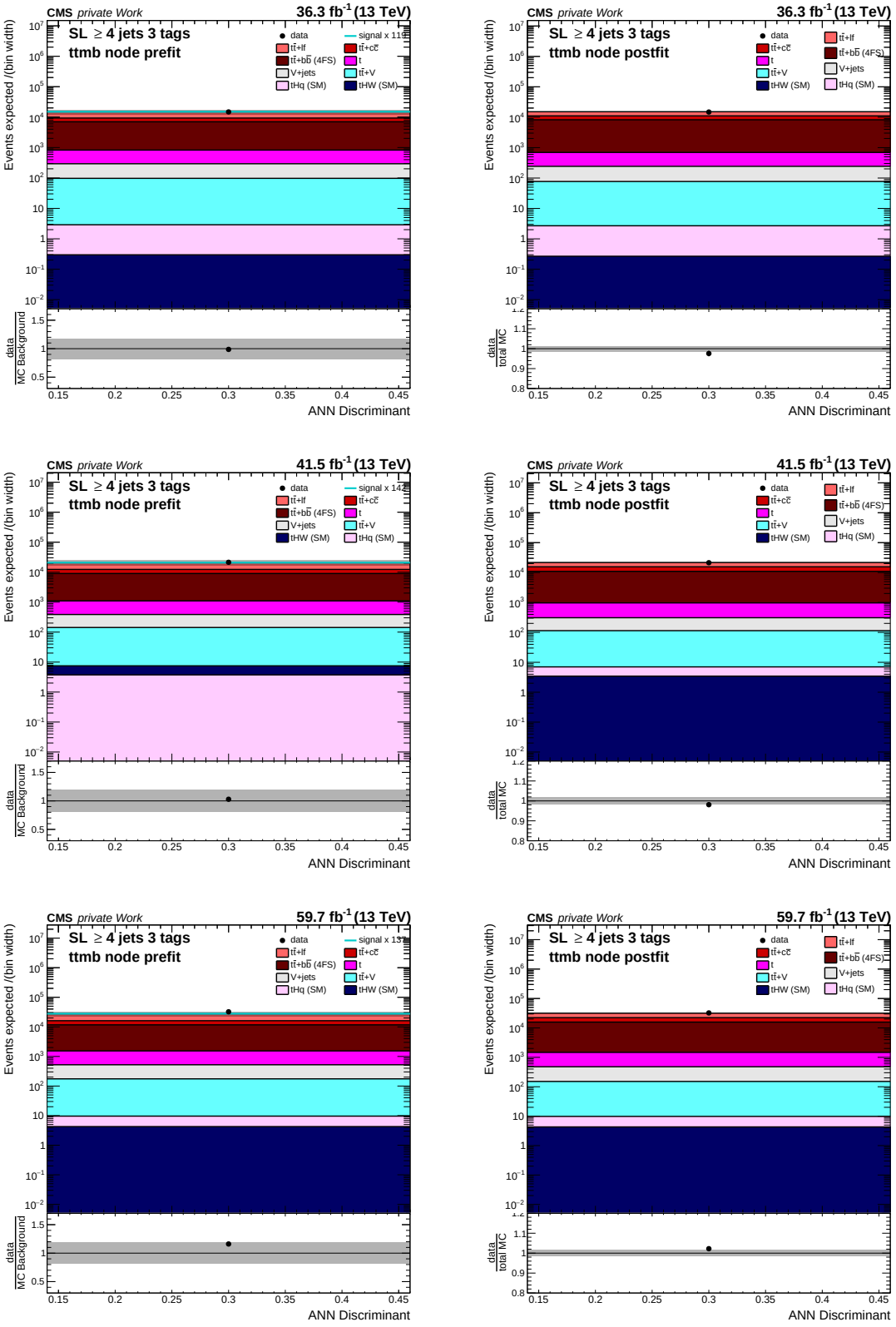


Figure H.17: **Discriminants in the SL channel in the ≥ 4j, 3t category $t\bar{t}b\bar{b}$/b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

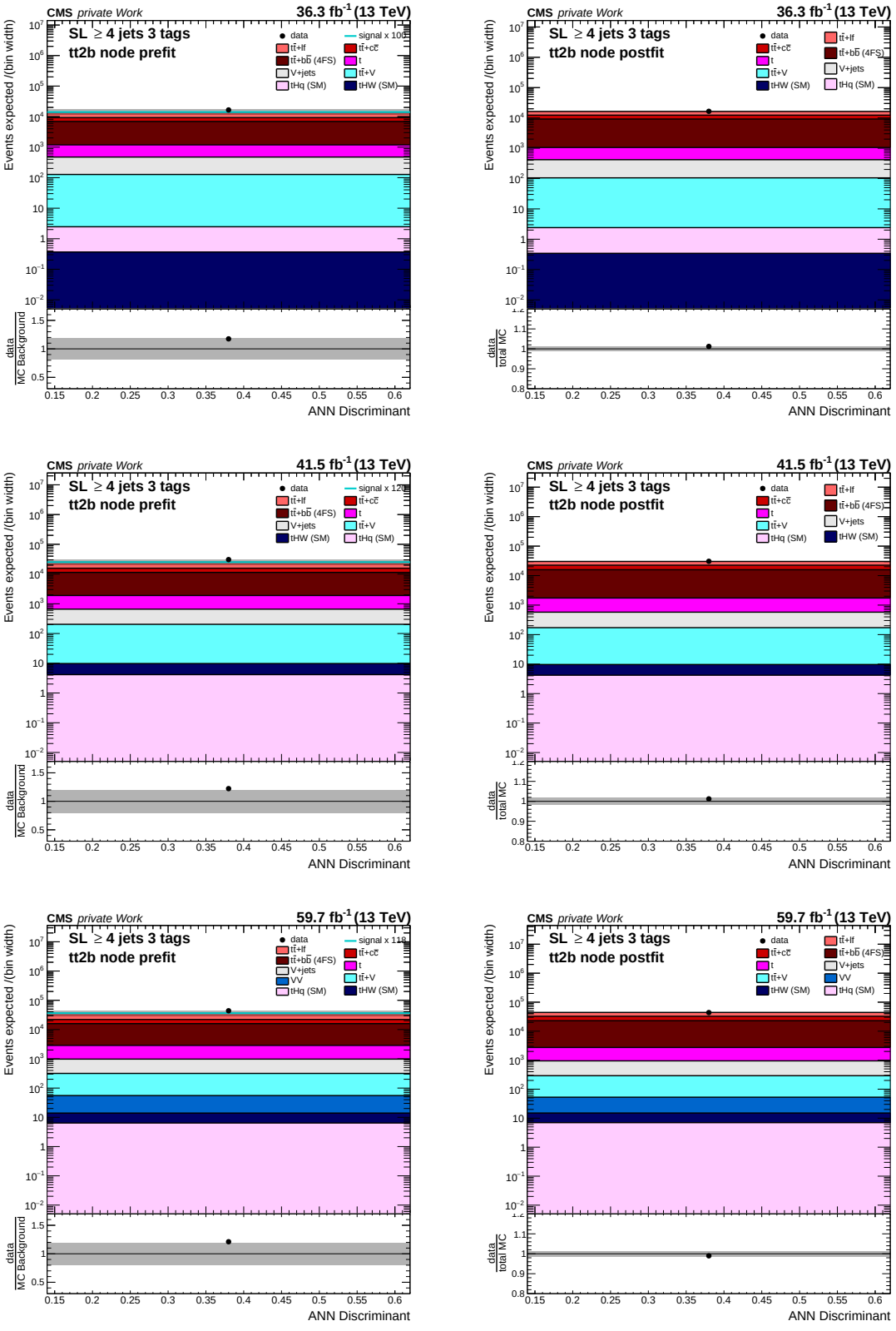


Figure H.18: **Discriminants in the SL channel in the $\geq$ 4j, 3t category $\mathrm{t\bar{t}}+$2b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

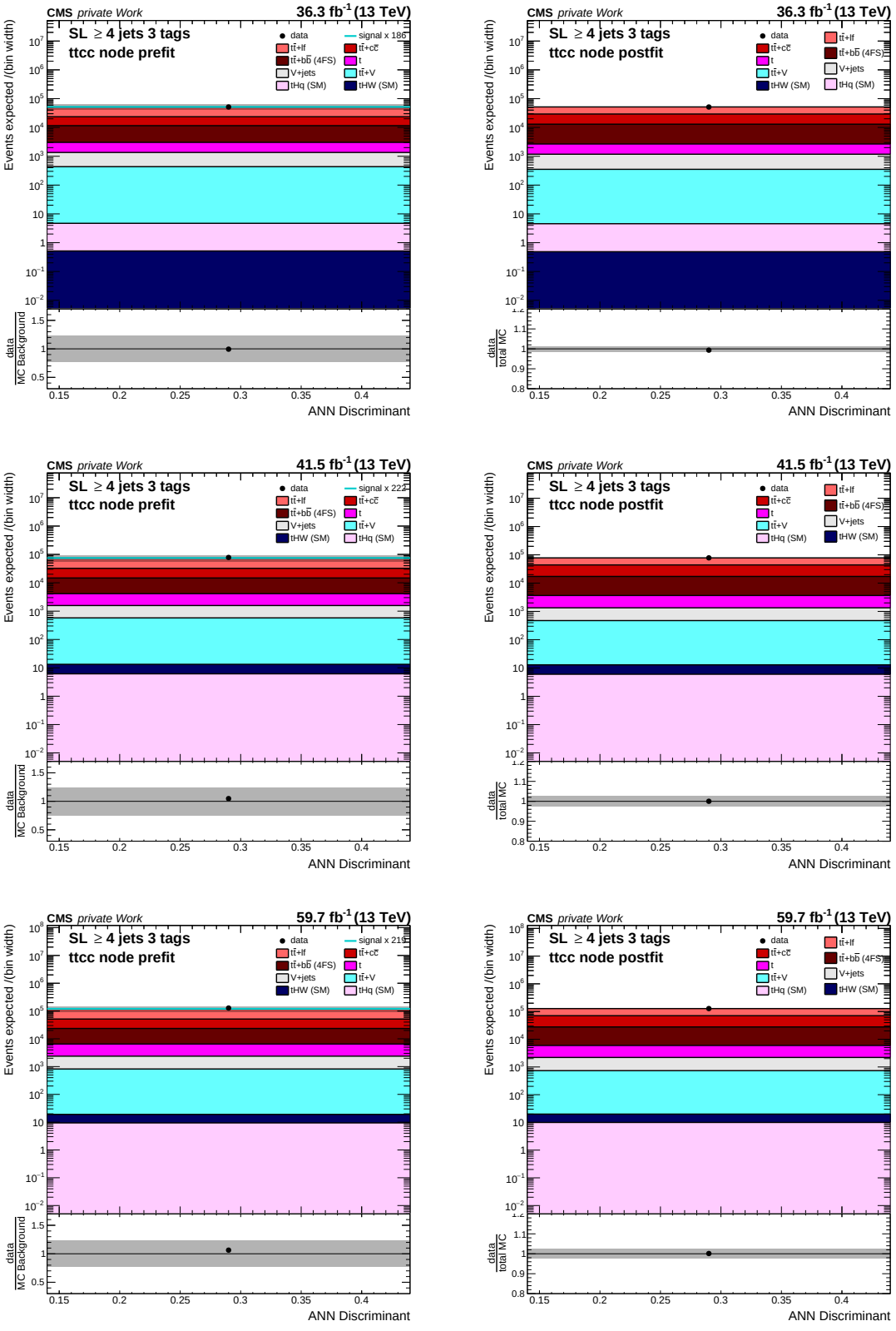


Figure H.19: **Discriminants in the SL channel in the ≥ 4j, 3t category $t\bar{t} + c\bar{c}$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

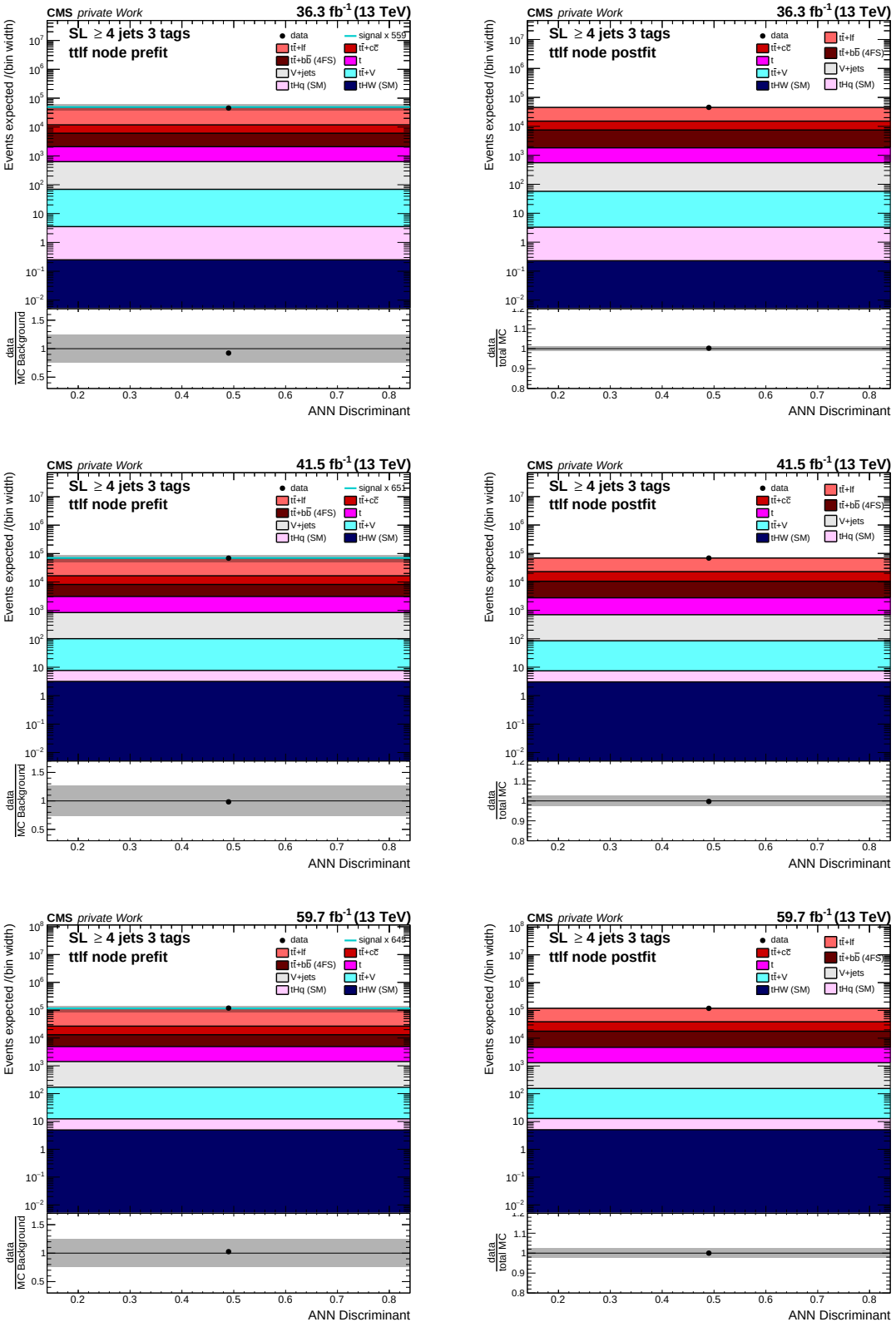


Figure H.20: **Discriminants in the SL channel in the $\geq$ 4j, 3t category $\mathrm{t\bar{t}}$ + lf node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

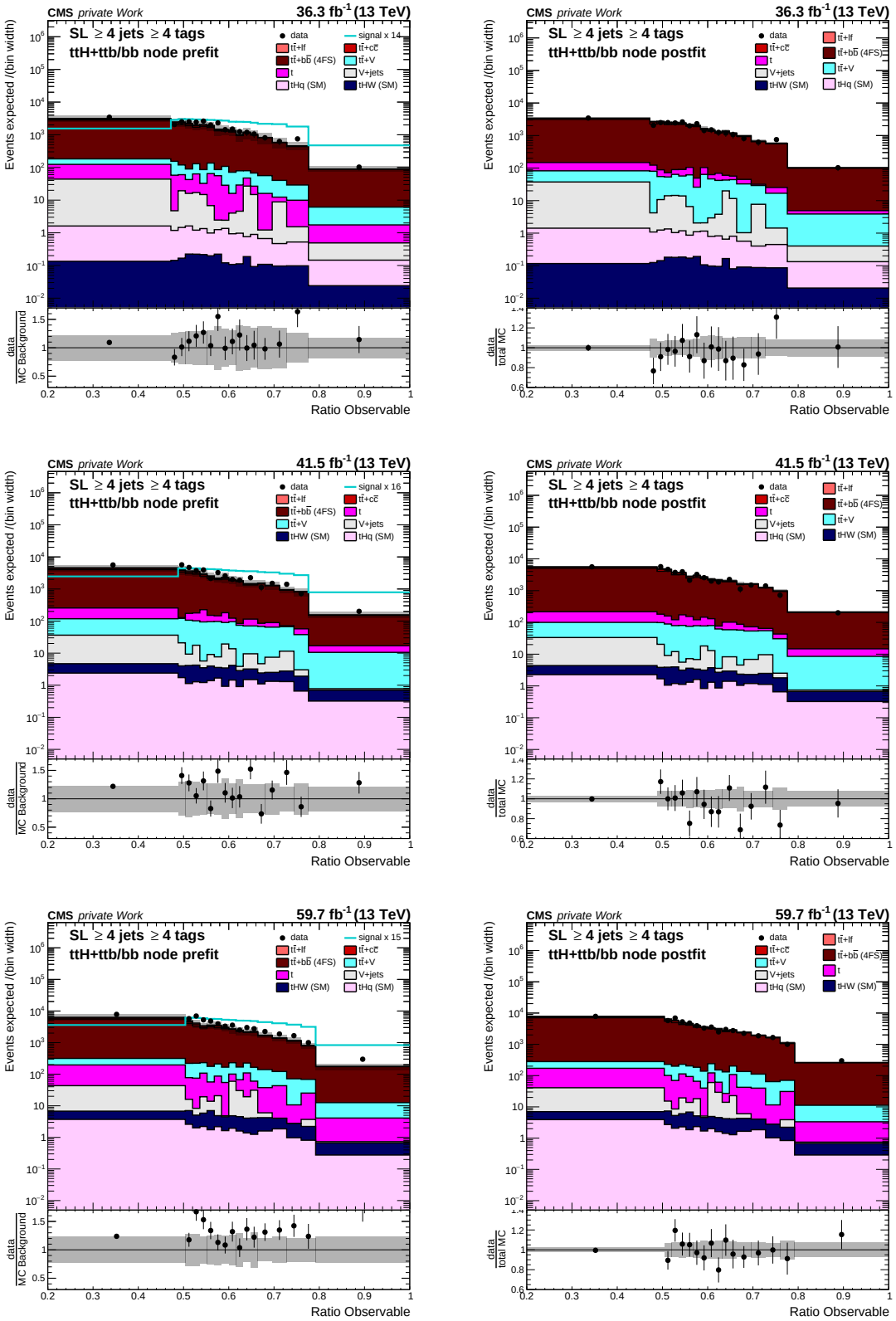


Figure H.21: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ttH + ttbb/b ratio observable category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

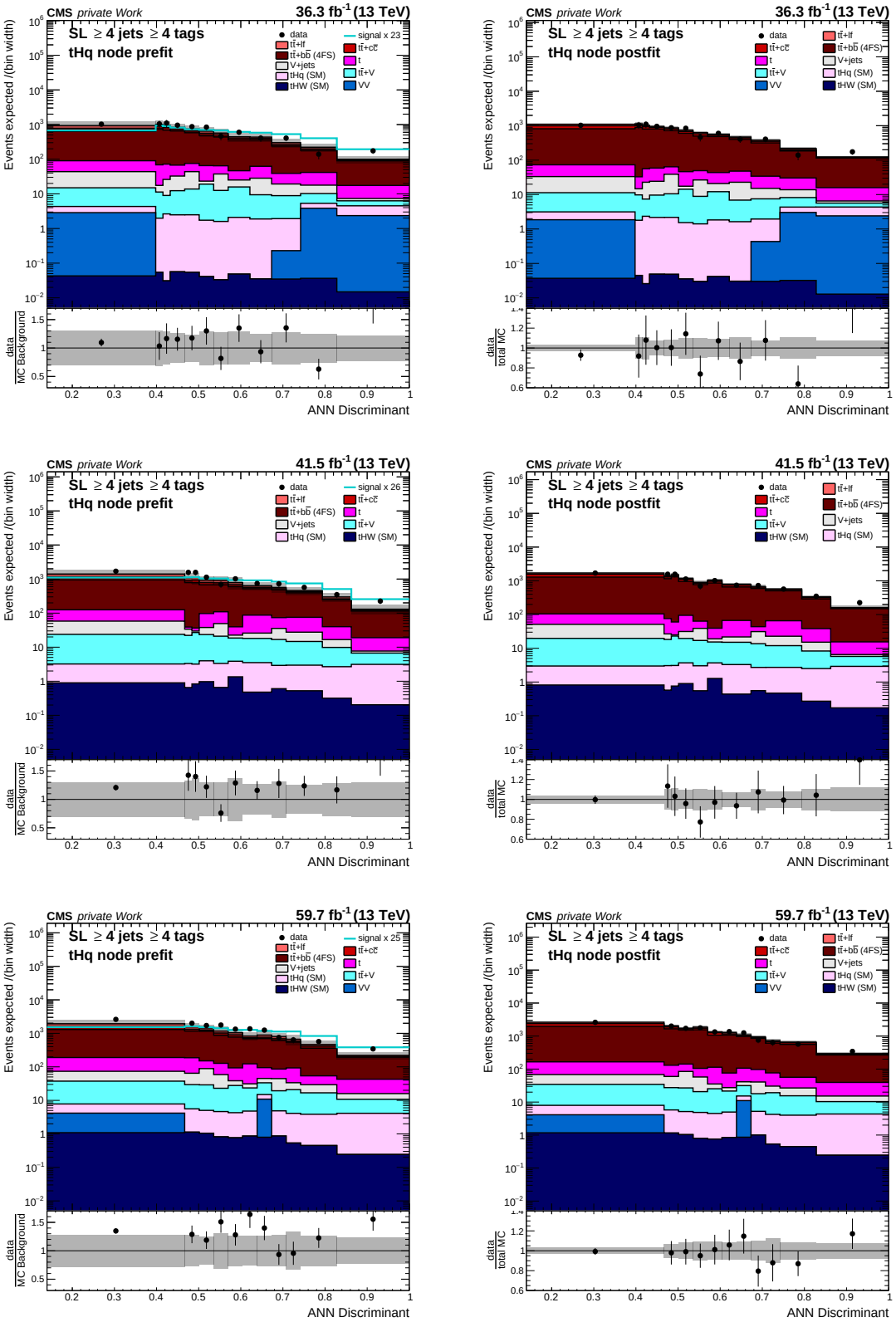


Figure H.22: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category tHq node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

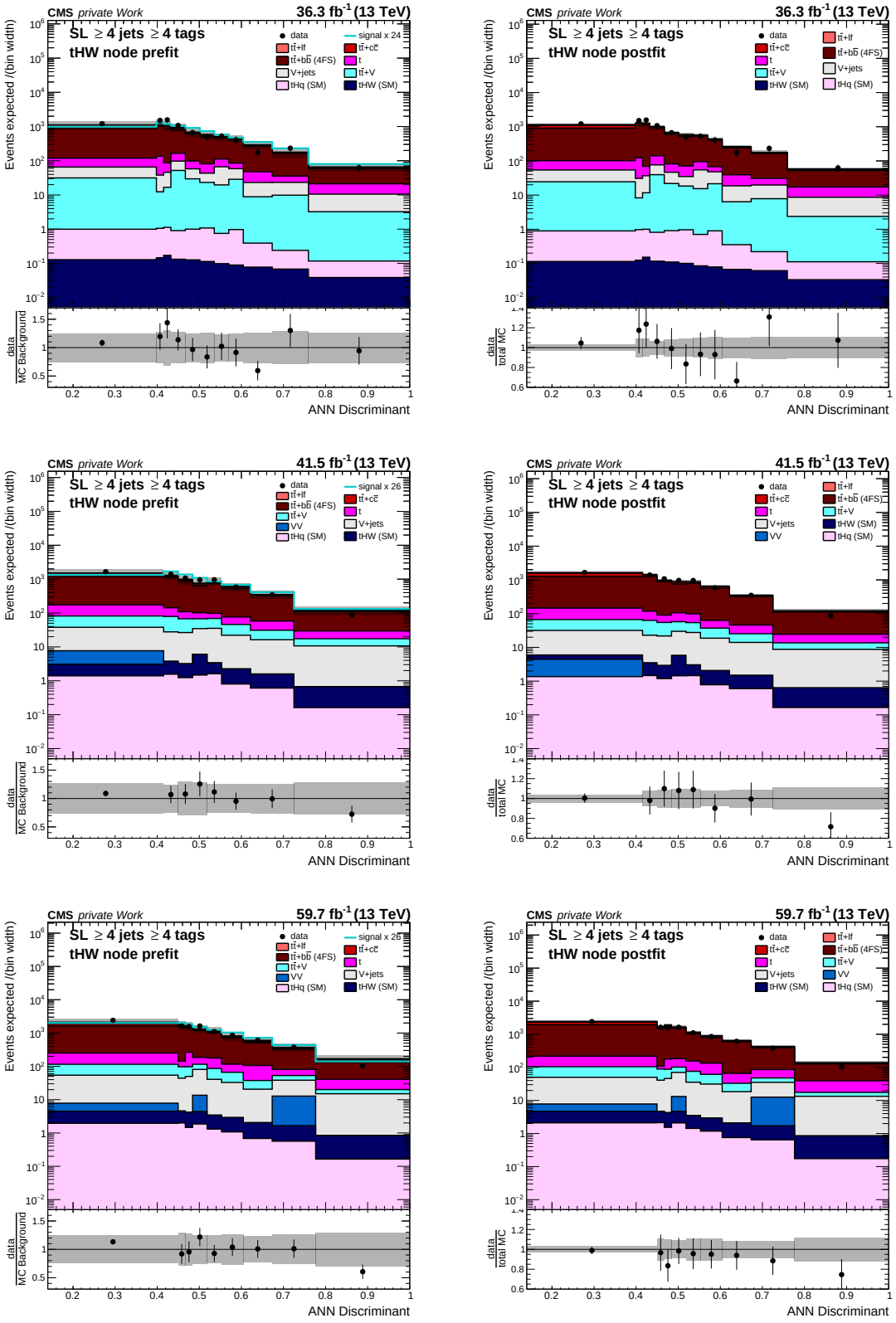


Figure H.23: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category tHW node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

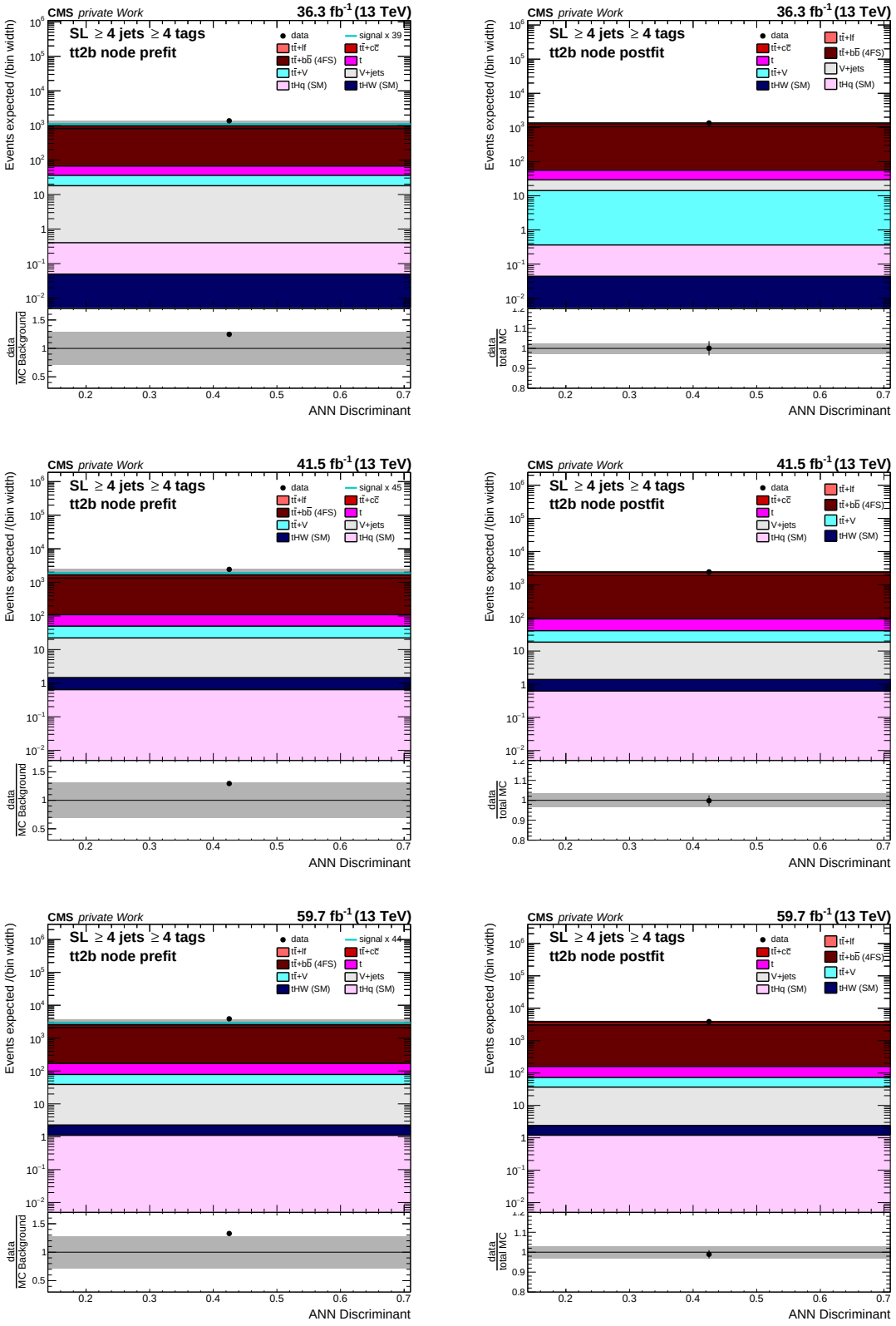


Figure H.24: **Discriminants in the SL channel in the $\geq$ 4j, $\geq$ 4t category $\mathrm{t\bar{t}}$ + 2b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

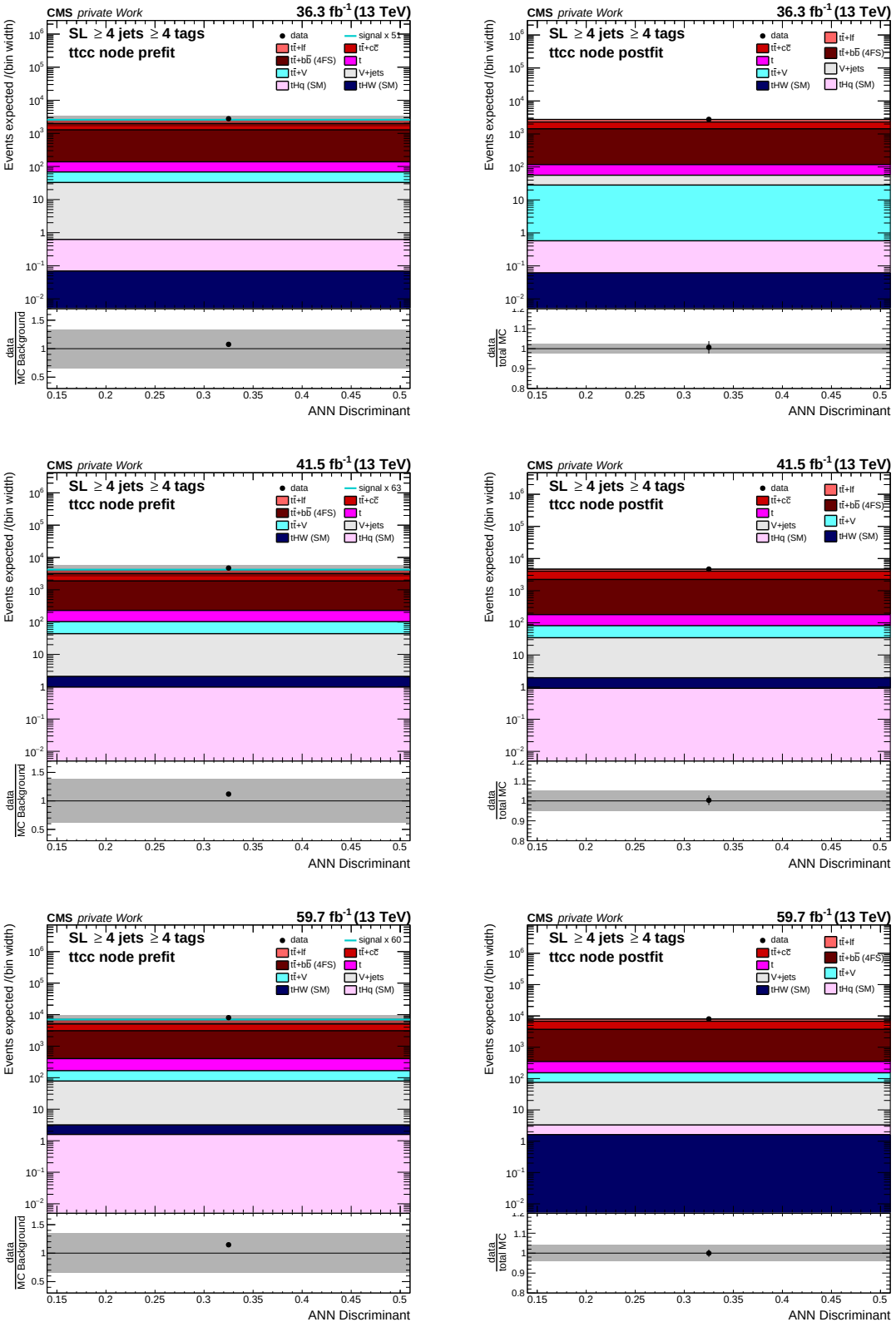


Figure H.25: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category $t\bar{t} + c\bar{c}$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

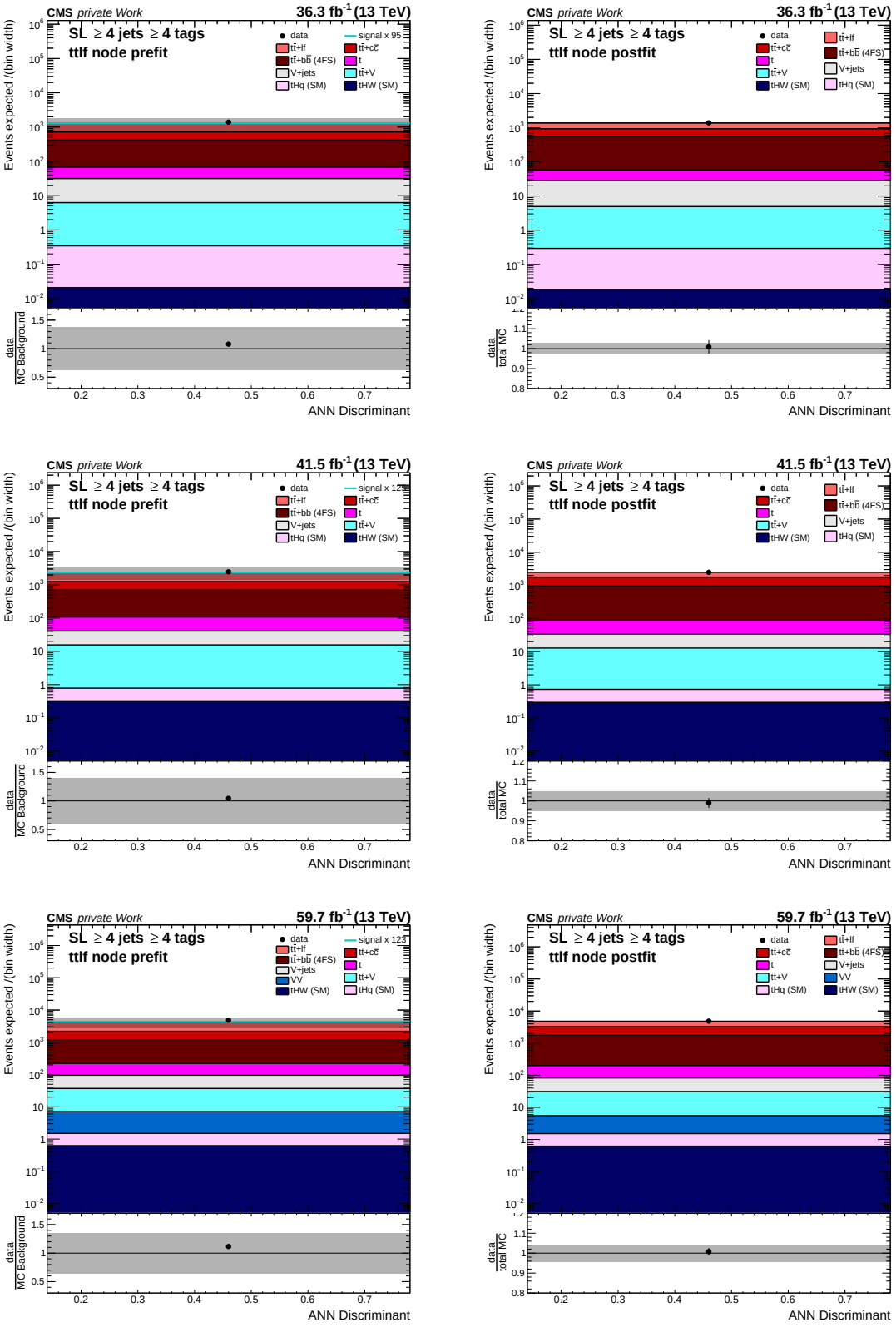


Figure H.26: **Discriminants in the SL channel in the $\geq$ 4j, $\geq$ 4t category $\mathrm{t\bar{t}}+\mathrm{lf}$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

Table H.15: **Prefit (post-fit) yields in the ≥ 4j, 3t category in the $t\bar{t}H$ and $t\bar{t}$ + jets nodes for the 2016 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields									
Process	$t\bar{t}H$		$t\bar{t}+b\bar{b}$		$t\bar{t}+2b$		$t\bar{t}+c\bar{c}$		$t\bar{t}+lf$	
$t\bar{t}+lf$	2160	(1630)	1816	(1310)	2311	(1919)	8413	(6703)	26146	(21101)
$t\bar{t}+c\bar{c}$	1103	(1397)	731	(906)	1111	(1506)	3601	(4889)	4056	(5386)
$t\bar{t}+b\bar{b}$	2219	(2388)	1987	(2416)	2764	(3912)	2555	(3099)	2809	(3970)
single t	119	(100)	170	(144)	342	(307)	502	(446)	1014	(882)
V + jets	49	(43)	63	(53)	166	(151)	276	(248)	393	(348)
$t\bar{t}+V$	73	(59)	30	(24)	60	(50)	129	(104)	46	(38)
Diboson	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
tH	1	(1)	0.9	(0.9)	1	(1)	1	(1)	2	(2)
Total bkg.	5723	(5618)	4798	(4854)	6755	(7845)	15478	(15491)	34465	(31727)
$\pm$ tot unc.	$\pm$1497	($\pm$81)	$\pm$859	($\pm$65)	$\pm$1279	($\pm$97)	$\pm$3528	($\pm$201)	$\pm$8328	($\pm$356)
$t\bar{t}H$	124	(0.0)	40	(0.0)	67	(0.0)	83	(0.0)	62	(0.0)
$\pm$ tot unc.	$\pm$14	($\pm$0.0)	$\pm$5	($\pm$0.0)	$\pm$8	($\pm$0.0)	$\pm$11	($\pm$0.0)	$\pm$7	($\pm$0.0)
Data	5579		4737		7940		15390		31803	

Table H.16: **Prefit (post-fit) yields in the ≥ 4j, 3t category in the tH nodes for the 2016 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	7535	(5749)	2234	(1679)
$t\bar{t}$ + $c\bar{c}$	1512	(1895)	1120	(1445)
$t\bar{t}$ + $b\bar{b}$	2522	(3661)	1593	(1929)
single t	507	(432)	461	(405)
V + jets	172	(150)	301	(269)
$t\bar{t}$ + V	37	(30)	72	(57)
Diboson	0.0	(0.0)	0.0	(0.0)
tH	5	(5)	2	(2)
Total bkg.	12290	(11922)	5783	(5787)
± tot unc.	±3059	(±165)	±1249	(±84)
$t\bar{t}$H	79	(0.0)	57	(0.0)
± tot unc.	±10	(±0.0)	±7	(±0.0)
Data	11866		5793	

Table H.17: **Prefit (post-fit) yields in the $\geq$ 4j, $\geq$ 4t category in the $t\bar{t}H$ and $t\bar{t}$ + jets nodes for the 2016 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields							
Process	Ratio Observable ($t\bar{t}H$, $t\bar{t}mb$)		$t\bar{t}$ + 2b		$t\bar{t}$ + $c\bar{c}$		$t\bar{t}$ + lf	
$t\bar{t}$ + lf	50	(37)	67	(50)	219	(165)	365	(284)
$t\bar{t}$ + $c\bar{c}$	83	(96)	84	(102)	259	(322)	193	(243)
$t\bar{t}$ + $b\bar{b}$	1070	(1235)	431	(589)	423	(487)	224	(317)
single t	29	(24)	18	(15)	26	(23)	22	(19)
V + jets	14	(12)	10	(9)	12	(10)	16	(15)
$t\bar{t}$ + V	31	(24)	10	(8)	13	(10)	4	(3)
Diboson	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
tH	0.8	(0.7)	0.2	(0.2)	0.2	(0.2)	0.2	(0.2)
Total bkg.	1277	(1428)	620	(773)	953	(1018)	824	(882)
$\pm$ tot unc.	$\pm$282	($\pm$41)	$\pm$179	($\pm$21)	$\pm$326	($\pm$26)	$\pm$308	($\pm$26)
$t\bar{t}H$	87	(0.0)	16	(0.0)	19	(0.0)	9	(0.0)
$\pm$ tot unc.	$\pm$14	($\pm$0.0)	$\pm$3	($\pm$0.0)	$\pm$3	($\pm$0.0)	$\pm$2	($\pm$0.0)
Data	1409		774		1025		890	

Table H.18: **Prefit (post-fit) yields in the ≥ 4j, ≥ 4t category in the tH nodes for the 2016 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	67	(47)	38	(25)
$t\bar{t}$ + $c\bar{c}$	58	(72)	53	(63)
$t\bar{t}$ + $b\bar{b}$	298	(393)	335	(360)
single t	26	(22)	29	(25)
V + jets	14	(11)	19	(17)
$t\bar{t}$ + V	7	(5)	16	(12)
Diboson	1	(1)	0.0	(0.0)
tH	2	(1)	0.5	(0.5)
Total bkg.	474	(552)	490	(503)
± tot unc.	±127	(±16)	±117	(±14)
$t\bar{t}$H	20	(0.0)	20	(0.0)
± tot unc.	±3	(±0.0)	±3	(±0.0)
Data	532		522	

Table H.19: **Prefit (post-fit) yields in the ≥ 4j, 3t category in the $t\bar{t}H$ and $t\bar{t}+$jets nodes for the 2017 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields									
Process	$t\bar{t}H$		$t\bar{t}+b\bar{b}$		$t\bar{t}+2b$		$t\bar{t}+c\bar{c}$		$t\bar{t}+lf$	
$t\bar{t}+lf$	3354	(2470)	2681	(1970)	4481	(3675)	12669	(10264)	37362	(32179)
$t\bar{t}+c\bar{c}$	1646	(2321)	1067	(1477)	2083	(3227)	5283	(8002)	5852	(8715)
$t\bar{t}+b\bar{b}$	2927	(3427)	2534	(3200)	4538	(6777)	3195	(4078)	3560	(5431)
single t	191	(151)	228	(210)	595	(557)	764	(696)	1574	(1434)
V + jets	50	(39)	78	(62)	219	(194)	300	(260)	517	(430)
$t\bar{t}+V$	105	(86)	43	(34)	93	(78)	169	(139)	65	(54)
Diboson	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
tH	4	(4)	2	(2)	5	(5)	4	(4)	5	(5)
Total bkg.	8276	(8498)	6634	(6957)	12013	(14512)	22383	(23442)	48935	(48249)
± tot unc.	±1839	(±177)	±1301	(±119)	±2396	(±244)	±5530	(±639)	±13097	(±1442)
$t\bar{t}H$	147	(0.0)	47	(0.0)	100	(0.0)	101	(0.0)	75	(0.0)
± tot unc.	±16	(±0.0)	±5	(±0.0)	±11	(±0.0)	±14	(±0.0)	±10	(±0.0)
Data	8484		6826		14689		23447		48112	

Table H.20: **Prefit (post-fit) yields in the $\geq$ 4j, 3t category in the tH nodes for the 2017 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	10156	(7871)	2751	(1981)
$t\bar{t}$ + $c\bar{c}$	2102	(2924)	1502	(2121)
$t\bar{t}$ + $b\bar{b}$	3323	(5051)	1967	(2522)
single t	681	(594)	633	(568)
V + jets	202	(169)	347	(295)
$t\bar{t}$ + V	57	(47)	107	(85)
Diboson	0.0	(0.0)	8	(7)
tH	10	(10)	6	(5)
Total bkg.	16532	(16666)	7320	(7586)
$\pm$ tot unc.	$\pm$4563	($\pm$480)	$\pm$1616	($\pm$193)
$t\bar{t}$H	95	(0.0)	65	(0.0)
$\pm$ tot unc.	$\pm$12	($\pm$0.0)	$\pm$7	($\pm$0.0)
Data	16490		7672	

Table H.21: **Prefit (post-fit) yields in the $\geq$ 4j, $\geq$ 4t category in the $t\bar{t}H$ and $t\bar{t}$ + jets nodes for the 2017 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields							
Process	Ratio Observable ($t\bar{t}H$, $t\bar{t}mb$)		$t\bar{t}+2b$		$t\bar{t}+c\bar{c}$		$t\bar{t}+lf$	
$t\bar{t}+lf$	96	(51)	129	(73)	374	(229)	710	(462)
$t\bar{t}+c\bar{c}$	139	(194)	169	(236)	479	(674)	360	(524)
$t\bar{t}+b\bar{b}$	1588	(2023)	712	(1030)	612	(762)	378	(558)
single t	53	(47)	34	(30)	47	(37)	43	(35)
V + jets	11	(10)	12	(10)	15	(12)	16	(14)
$t\bar{t}+V$	46	(37)	16	(13)	22	(17)	10	(8)
Diboson	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)
tH	2	(2)	0.8	(0.8)	0.8	(0.7)	0.5	(0.5)
Total bkg.	1936	(2364)	1073	(1392)	1551	(1731)	1517	(1601)
$\pm$ tot unc.	$\pm$451	($\pm$70)	$\pm$334	($\pm$57)	$\pm$593	($\pm$80)	$\pm$599	($\pm$75)
$t\bar{t}H$	121	(0.0)	23	(0.0)	25	(0.0)	12	(0.0)
$\pm$ tot unc.	$\pm$20	($\pm$0.0)	$\pm$4	($\pm$0.0)	$\pm$5	($\pm$0.0)	$\pm$2	($\pm$0.0)
Data	2348		1389		1737		1585	

Table H.22: **Prefit (post-fit) yields in the ≥ 4j, ≥ 4t category in the tH nodes for the 2017 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	116	(62)	48	(25)
$t\bar{t}$ + $c\bar{c}$	90	(122)	77	(108)
$t\bar{t}$ + $b\bar{b}$	453	(643)	458	(509)
single t	40	(33)	39	(35)
V + jets	16	(14)	18	(15)
$t\bar{t}$ + V	13	(10)	23	(18)
Diboson	0.0	(0.0)	1	(0.8)
tH	3	(3)	2	(2)
Total bkg.	730	(887)	665	(712)
± tot unc.	±207	(±29)	±166	(±25)
$t\bar{t}$H	27	(0.0)	25	(0.0)
± tot unc.	±4	(±0.0)	±4	(±0.0)
Data	887		707	

Table H.23: **Prefit (post-fit) yields in the ≥ 4j, 3t category in the $t\bar{t}H$ and $t\bar{t}+$jets nodes for the 2018 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields									
Process	$t\bar{t}H$		$t\bar{t}+b\bar{b}$		$t\bar{t}+2b$		$t\bar{t}+c\bar{c}$		$t\bar{t}+lf$	
$t\bar{t}+lf$	4698	(3719)	3736	(2960)	6787	(5591)	20638	(17095)	62390	(55879)
$t\bar{t}+c\bar{c}$	2213	(3228)	1420	(2100)	2952	(4655)	8386	(12899)	9505	(14815)
$t\bar{t}+b\bar{b}$	4199	(5193)	3255	(4579)	6272	(9715)	5118	(6557)	5734	(9214)
single t	253	(231)	324	(317)	910	(862)	1236	(1128)	2448	(2362)
V + jets	58	(55)	112	(105)	317	(314)	466	(440)	879	(808)
$t\bar{t}+V$	170	(152)	53	(45)	126	(115)	240	(215)	110	(100)
Diboson	0.0	(0.0)	0.0	(0.0)	20	(18)	0.0	(0.0)	0.0	(0.0)
tH	6	(6)	3	(3)	7	(7)	6	(6)	9	(9)
Total bkg.	11597	(12583)	8902	(10109)	17390	(21278)	36091	(38339)	81075	(83186)
± tot unc.	±2384	(±222)	±1682	(±174)	±3379	(±291)	±8428	(±838)	±19854	(±1865)
$t\bar{t}H$	212	(0.0)	65	(0.0)	146	(0.0)	165	(0.0)	126	(0.0)
± tot unc.	±24	(±0.0)	±8	(±0.0)	±18	(±0.0)	±24	(±0.0)	±16	(±0.0)
Data	12381		10334		21061		38389		83192	

Table H.24: **Prefit (post-fit) yields in the $\geq$ 4j, 3t category in the tH nodes for the 2018 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	13359	(11092)	4252	(3203)
$t\bar{t}$ + $c\bar{c}$	2628	(3917)	2097	(3067)
$t\bar{t}$ + $b\bar{b}$	4337	(7137)	2770	(3638)
single t	877	(816)	916	(848)
V + jets	284	(265)	542	(493)
$t\bar{t}$ + V	74	(66)	145	(125)
Diboson	0.0	(0.0)	14	(12)
tH	14	(14)	8	(8)
Total bkg.	21572	(23308)	10744	(11394)
$\pm$ tot unc.	$\pm$5220	($\pm$493)	$\pm$2102	($\pm$236)
$t\bar{t}$H	129	(0.0)	94	(0.0)
$\pm$ tot unc.	$\pm$17	($\pm$0.0)	$\pm$11	($\pm$0.0)
Data	23452		11245	

Table H.25: **Prefit (post-fit) yields in the $\geq$ 4j, $\geq$ 4t category in the $t\bar{t}H$ and $t\bar{t}$ + jets nodes for the 2018 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields							
Process	Ratio Observable ($t\bar{t}H$, $t\bar{t}mb$)		$t\bar{t}$ + 2b		$t\bar{t}$ + c$\bar{c}$		$t\bar{t}$ + lf	
$t\bar{t}$ + lf	126	(78)	216	(139)	687	(461)	1379	(970)
$t\bar{t}$ + c$\bar{c}$	167	(238)	226	(344)	757	(1127)	641	(986)
$t\bar{t}$ + b$\bar{b}$	2200	(2893)	1121	(1655)	993	(1249)	620	(997)
single t	60	(53)	52	(49)	87	(73)	81	(75)
V + jets	14	(13)	21	(20)	28	(27)	37	(33)
$t\bar{t}$ + V	67	(61)	22	(21)	33	(29)	19	(16)
Diboson	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	4	(3)
tH	3	(4)	1	(1)	1	(1)	1.0	(1.0)
Total bkg.	2637	(3339)	1660	(2228)	2586	(2967)	2782	(3081)
$\pm$ tot unc.	$\pm$587	($\pm$84)	$\pm$472	($\pm$68)	$\pm$888	($\pm$125)	$\pm$995	($\pm$132)
$t\bar{t}H$	168	(0.0)	37	(0.0)	43	(0.0)	23	(0.0)
$\pm$ tot unc.	$\pm$29	($\pm$0.0)	$\pm$7	($\pm$0.0)	$\pm$9	($\pm$0.0)	$\pm$4	($\pm$0.0)
Data	3330		2205		2966		3105	

Table H.26: **Prefit (post-fit) yields in the ≥ 4j, ≥ 4t category in the tH nodes for the 2018 data set**. The prefit uncertainties do not include freely-floating parameters. The postfit uncertainties consider all nuisances in the fit to observed data. The a posteriori signal yield is set to zero since the corresponding signal strength is not known yet.

	pre-fit (post-fit) yields			
Process	tHq		tHW	
$t\bar{t}$ + lf	164	(106)	76	(43)
$t\bar{t}$ + $c\bar{c}$	120	(179)	110	(143)
$t\bar{t}$ + $b\bar{b}$	610	(965)	653	(782)
single t	64	(55)	66	(57)
V + jets	21	(19)	28	(25)
$t\bar{t}$ + V	17	(15)	29	(26)
Diboson	1	(1)	2	(2)
tH	4	(4)	2	(3)
Total bkg.	1001	(1346)	966	(1081)
$\pm$ tot unc.	$\pm$261	($\pm$39)	$\pm$227	($\pm$33)
$t\bar{t}$H	39	(0.0)	37	(0.0)
$\pm$ tot unc.	$\pm$7	($\pm$0.0)	$\pm$6	($\pm$0.0)
Data	1335		1045	

H.2 Cross section measurement

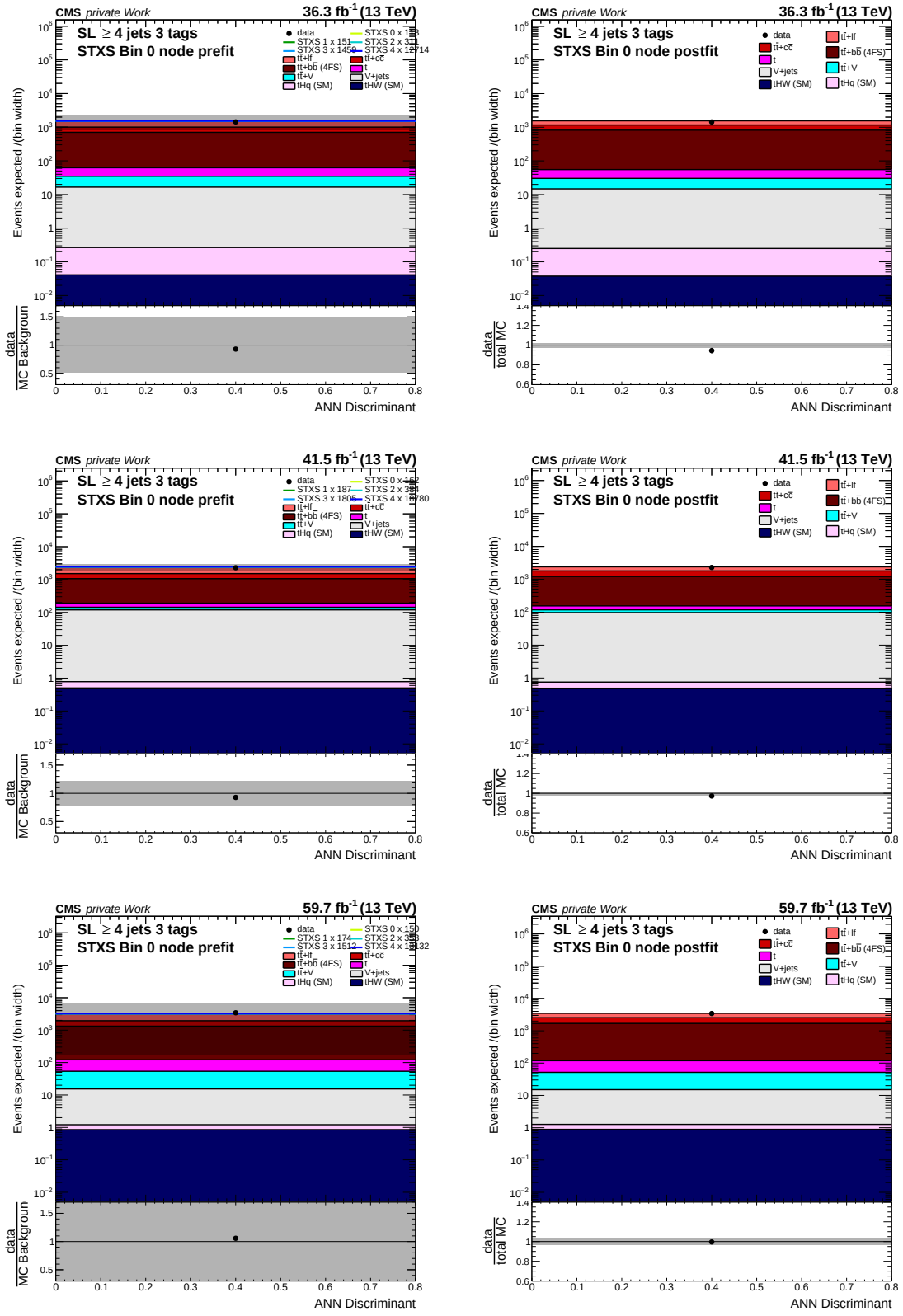


Figure H.27: **Discriminants in the SL channel in the ≥ 4j, 3t category tt̄H node times the STXS ANN output for the first STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

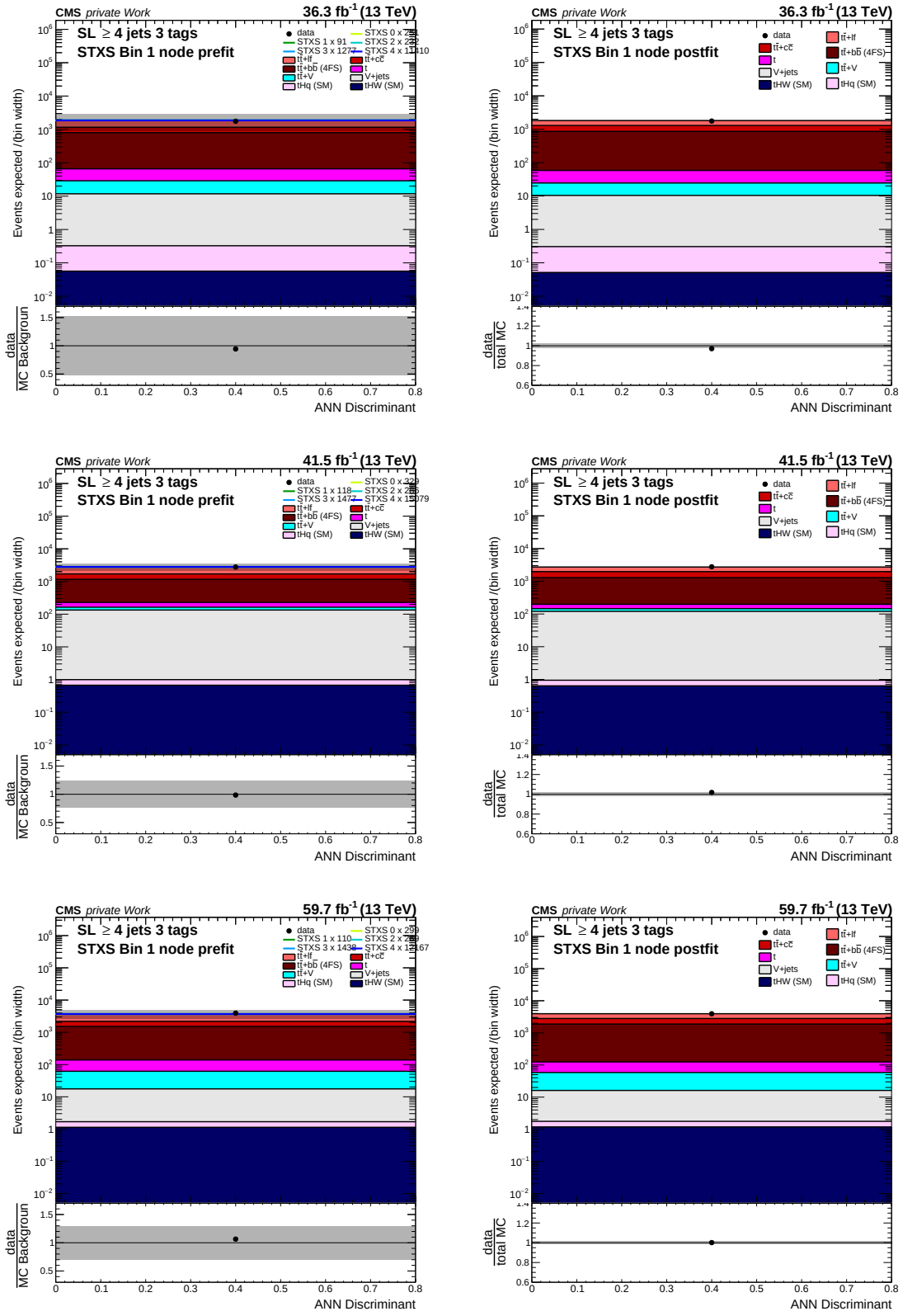


Figure H.28: **Discriminants in the SL channel in the ≥ 4j, 3t category tt̄H node times the STXS ANN output for the second STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

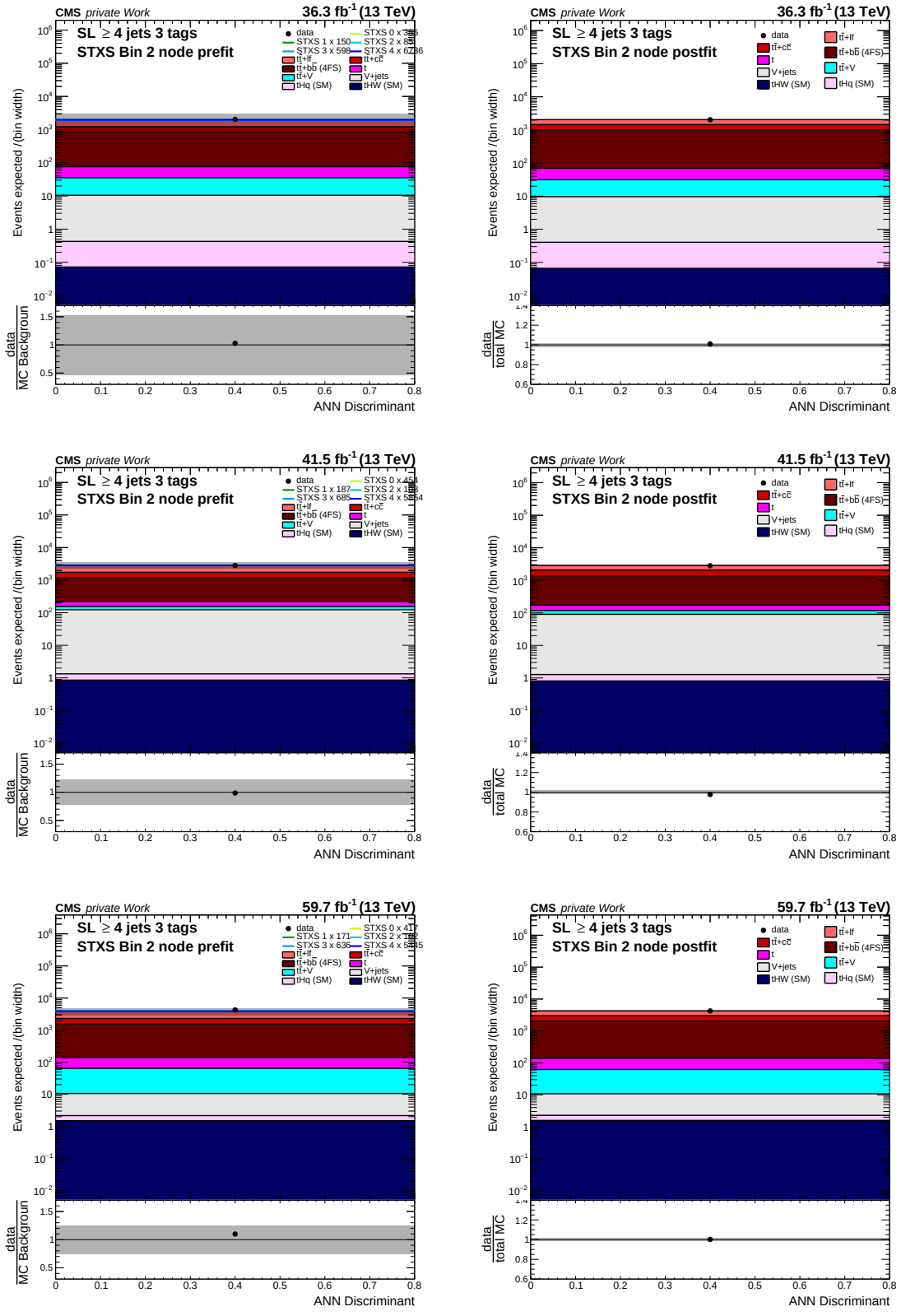


Figure H.29: **Discriminants in the SL channel in the ≥ 4j, 3t category tt̄H node times the STXS ANN output for the third STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

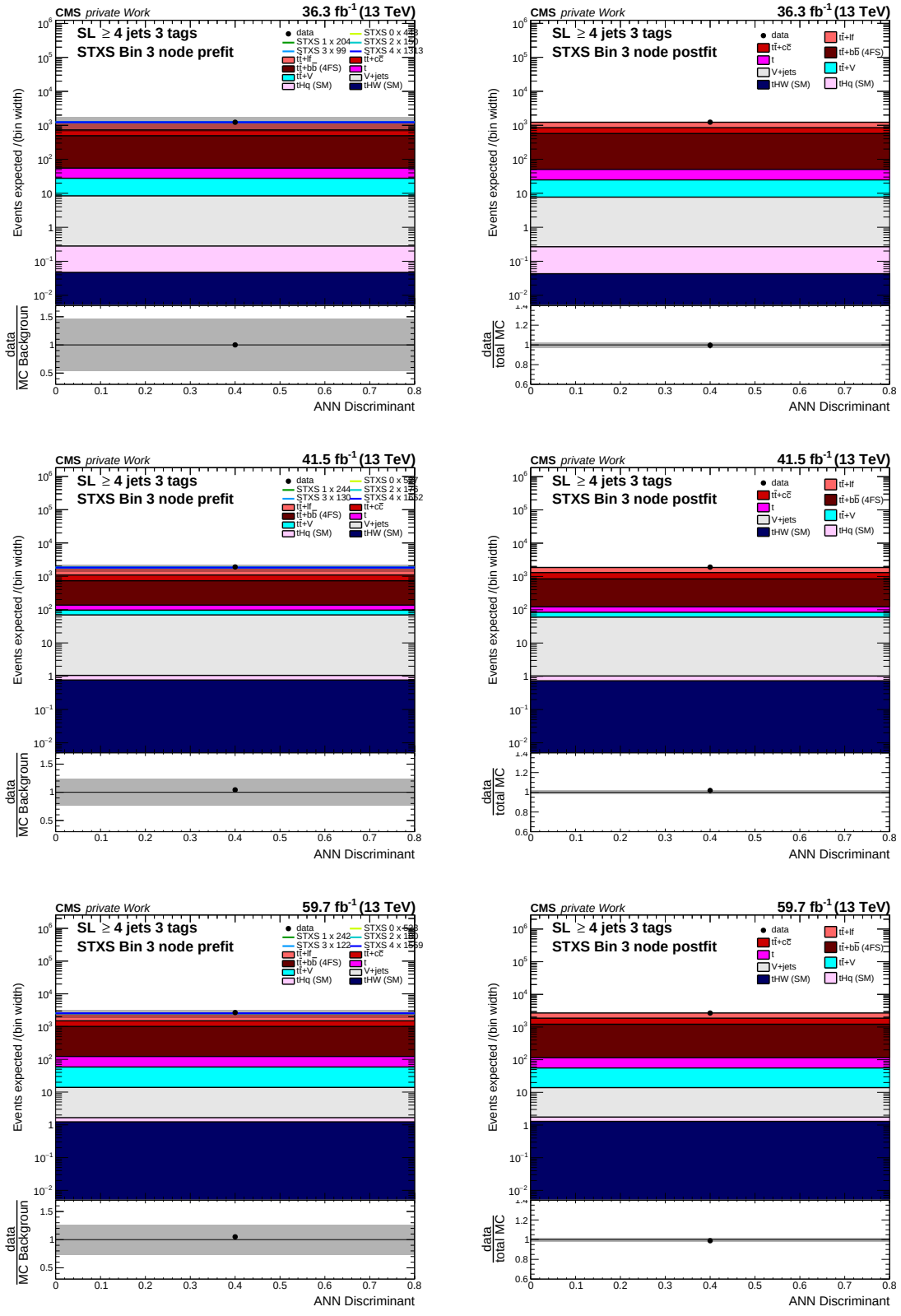


Figure H.30: **Discriminants in the SL channel in the ≥ 4j, 3t category tt̄H node times the STXS ANN output for the forth STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

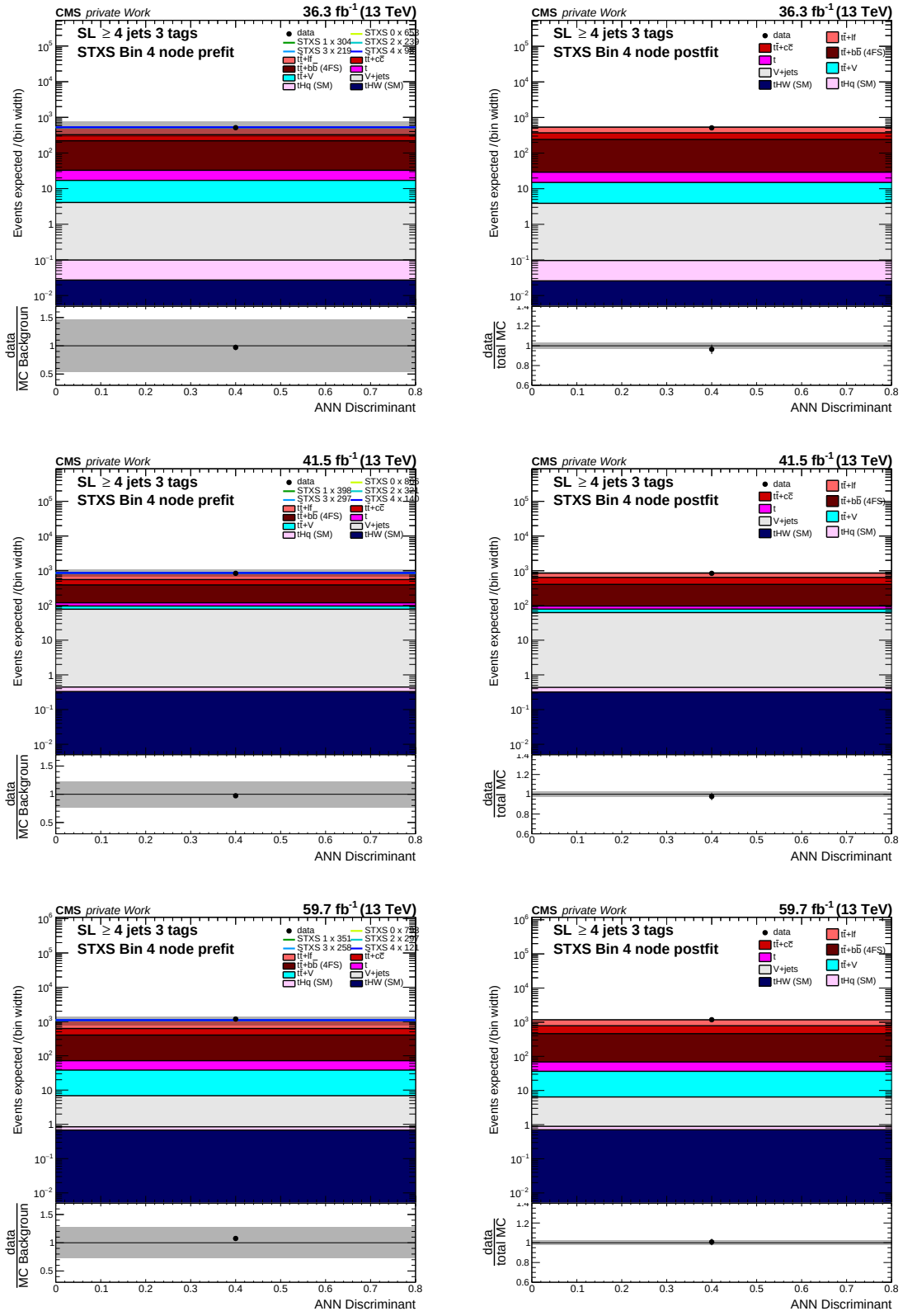


Figure H.31: **Discriminants in the SL channel in the $\geq$ 4j, 3t category $t\bar{t}H$ node times the STXS ANN output for the last STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

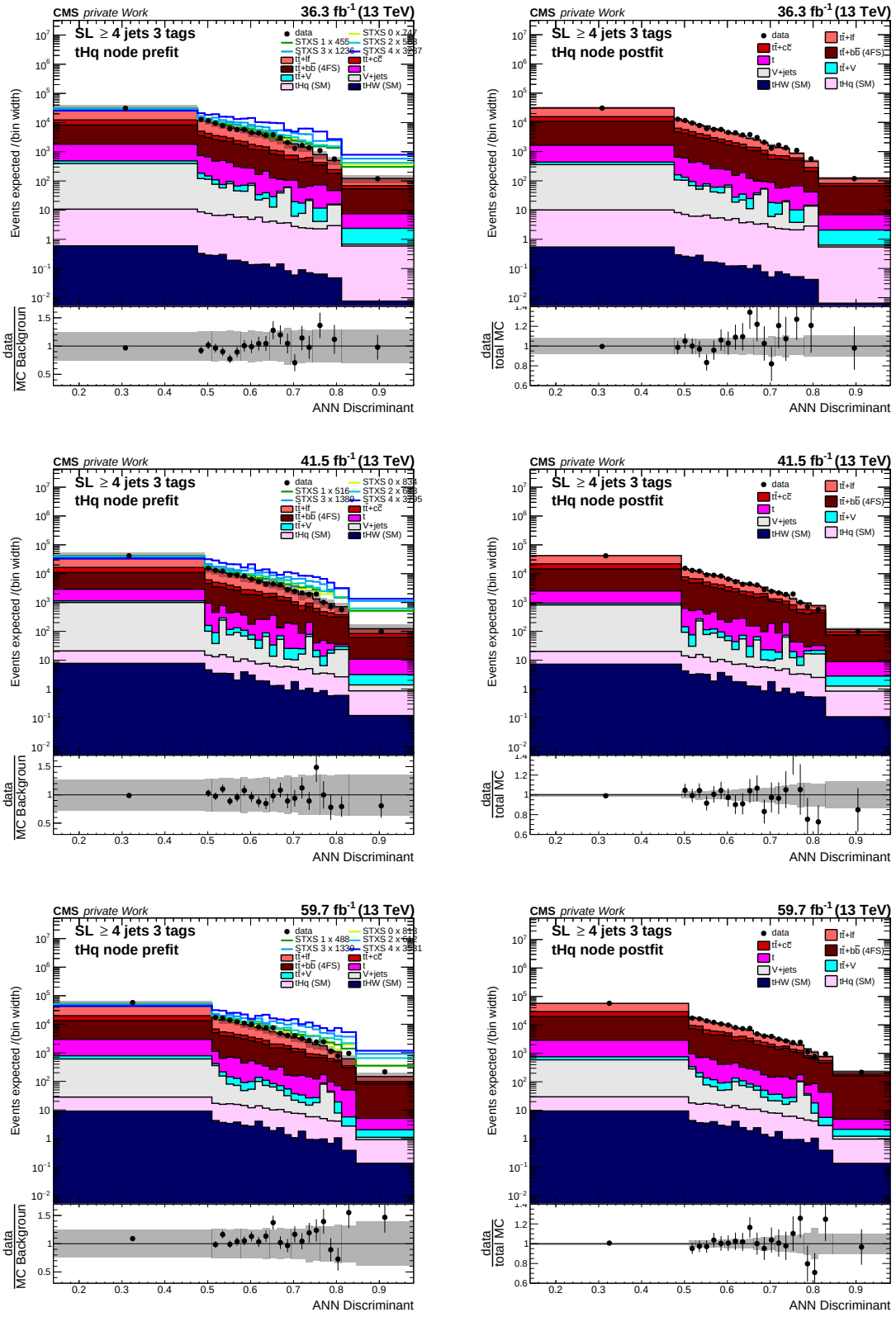


Figure H.32: **Discriminants in the SL channel in the $\geq$ 4j, 3t category tHq node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

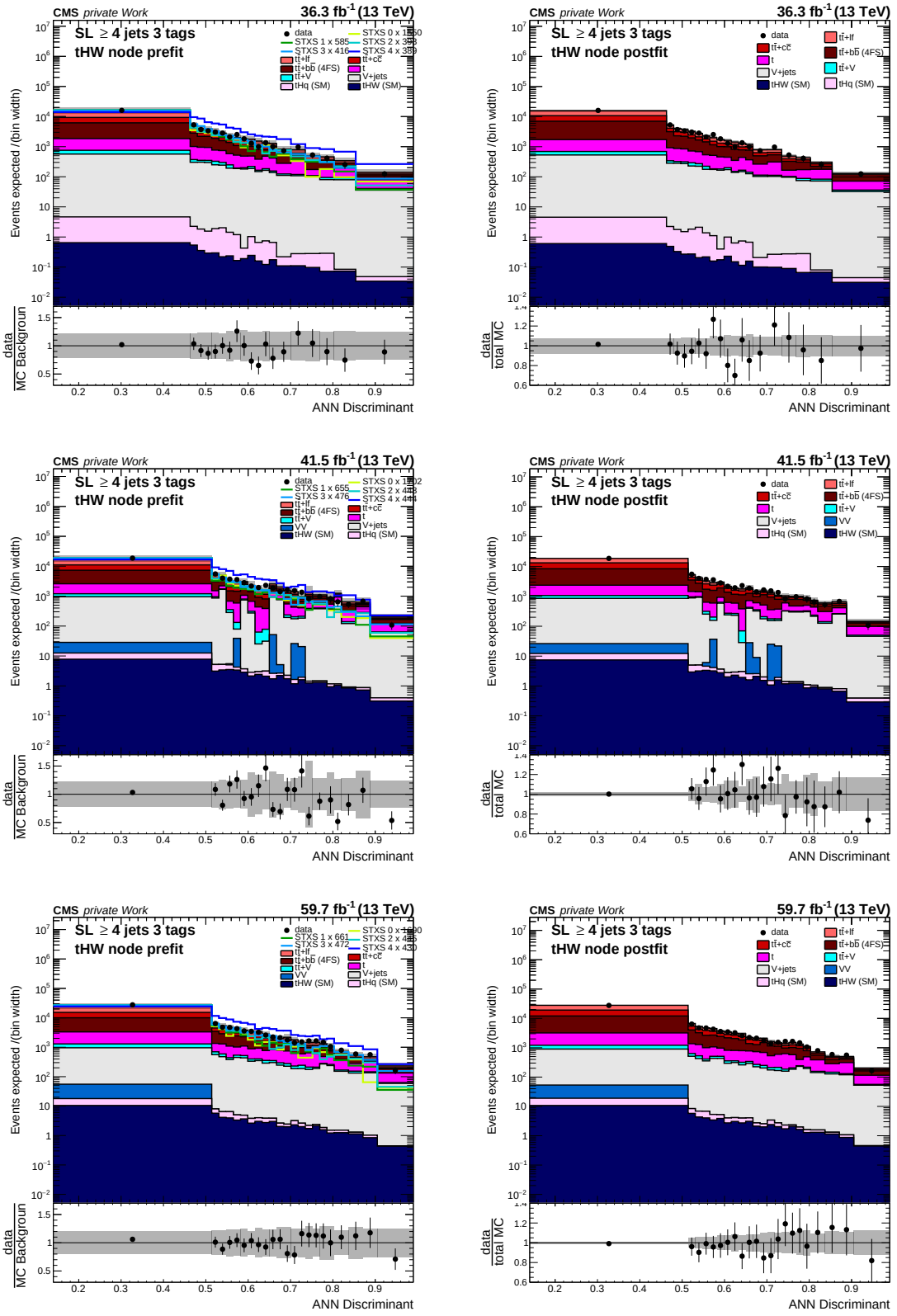


Figure H.33: **Discriminants in the SL channel in the ≥ 4j, 3t category tHW node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal template is scaled to the total background contribution in the pre-fit figures (cyan line).

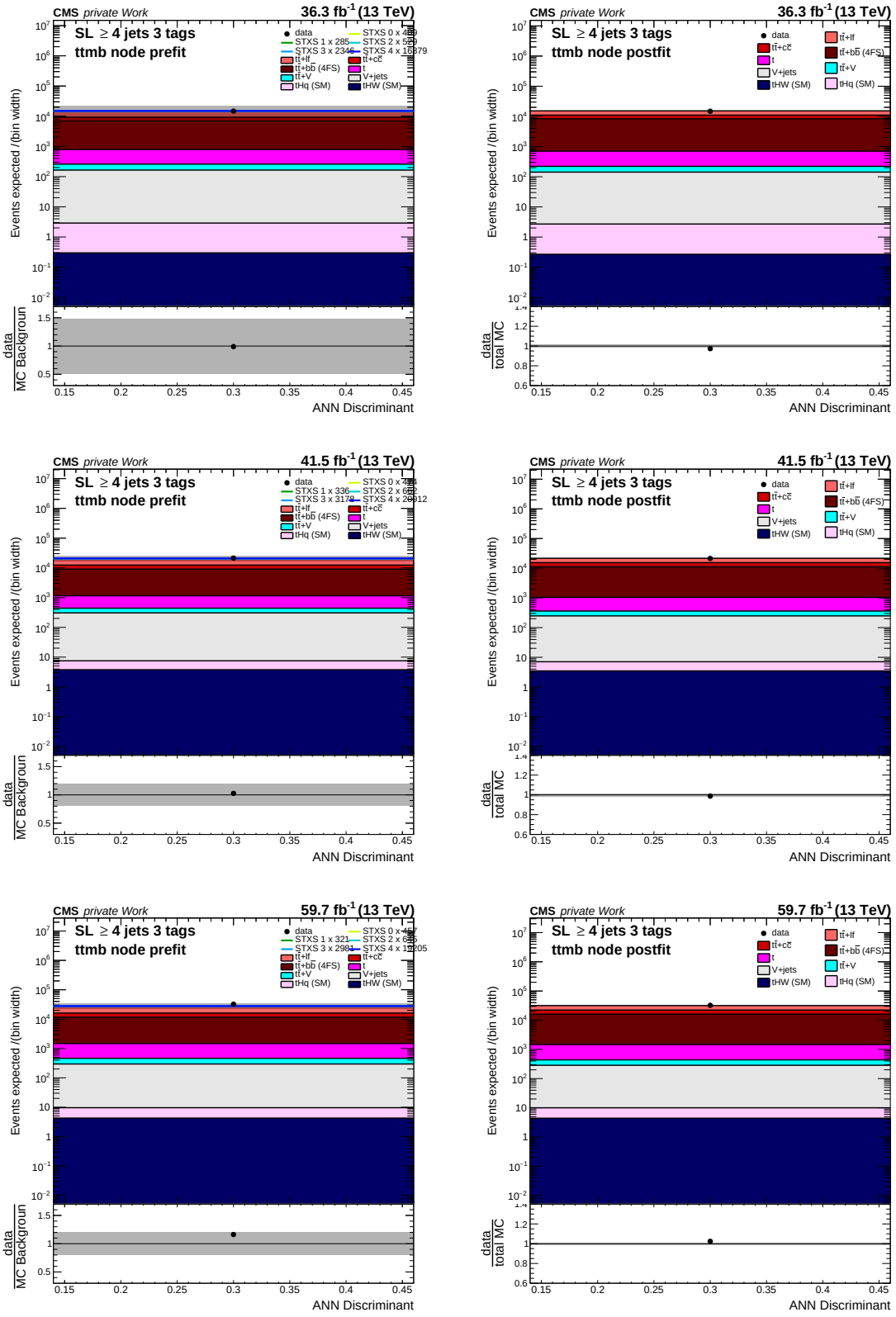


Figure H.34: **Discriminants in the SL channel in the ≥ 4j, 3t category tt̄bb̄/b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

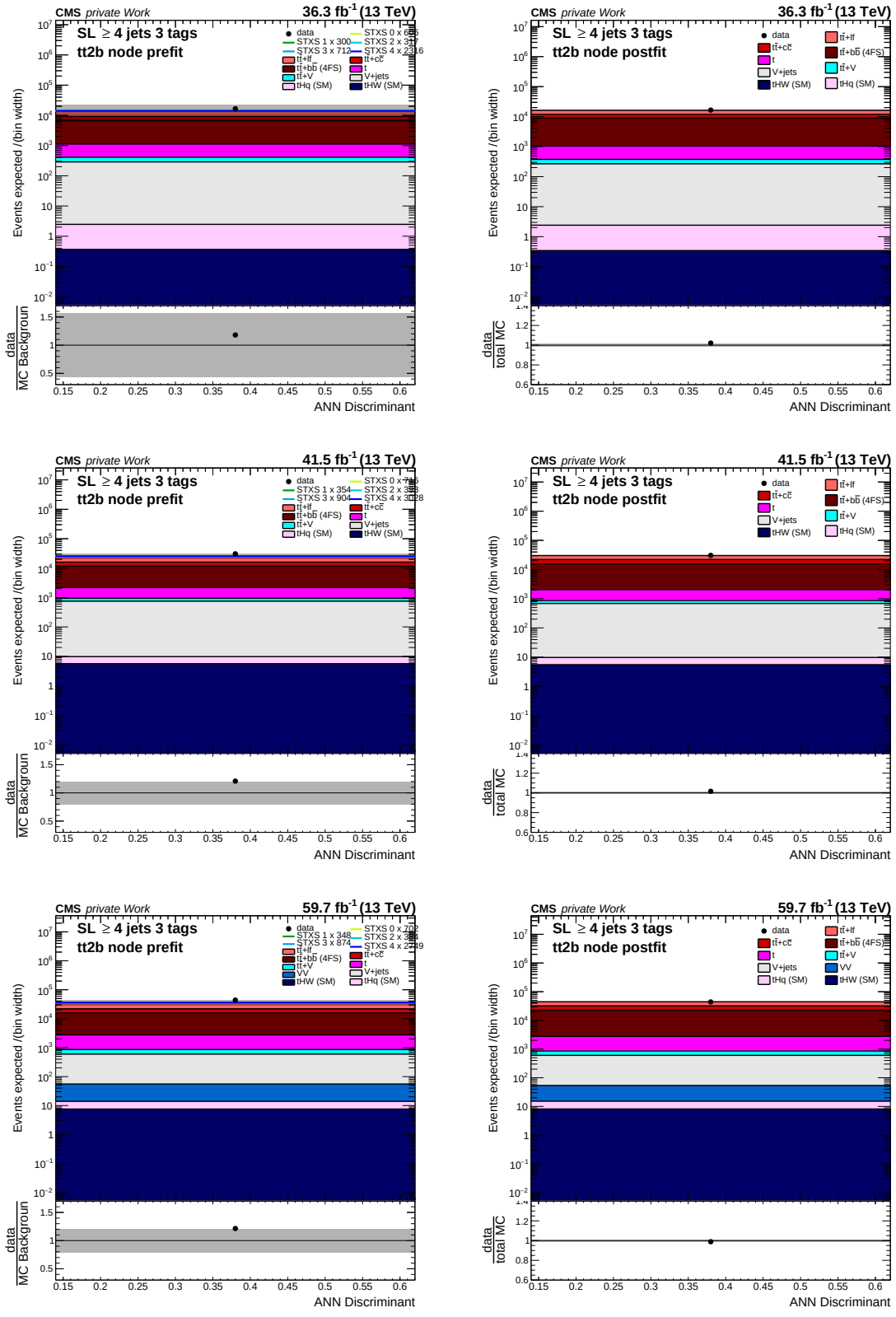


Figure H.35: **Discriminants in the SL channel in the $\geq$ 4j, 3t category $\mathrm{t\bar{t}}+$2b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

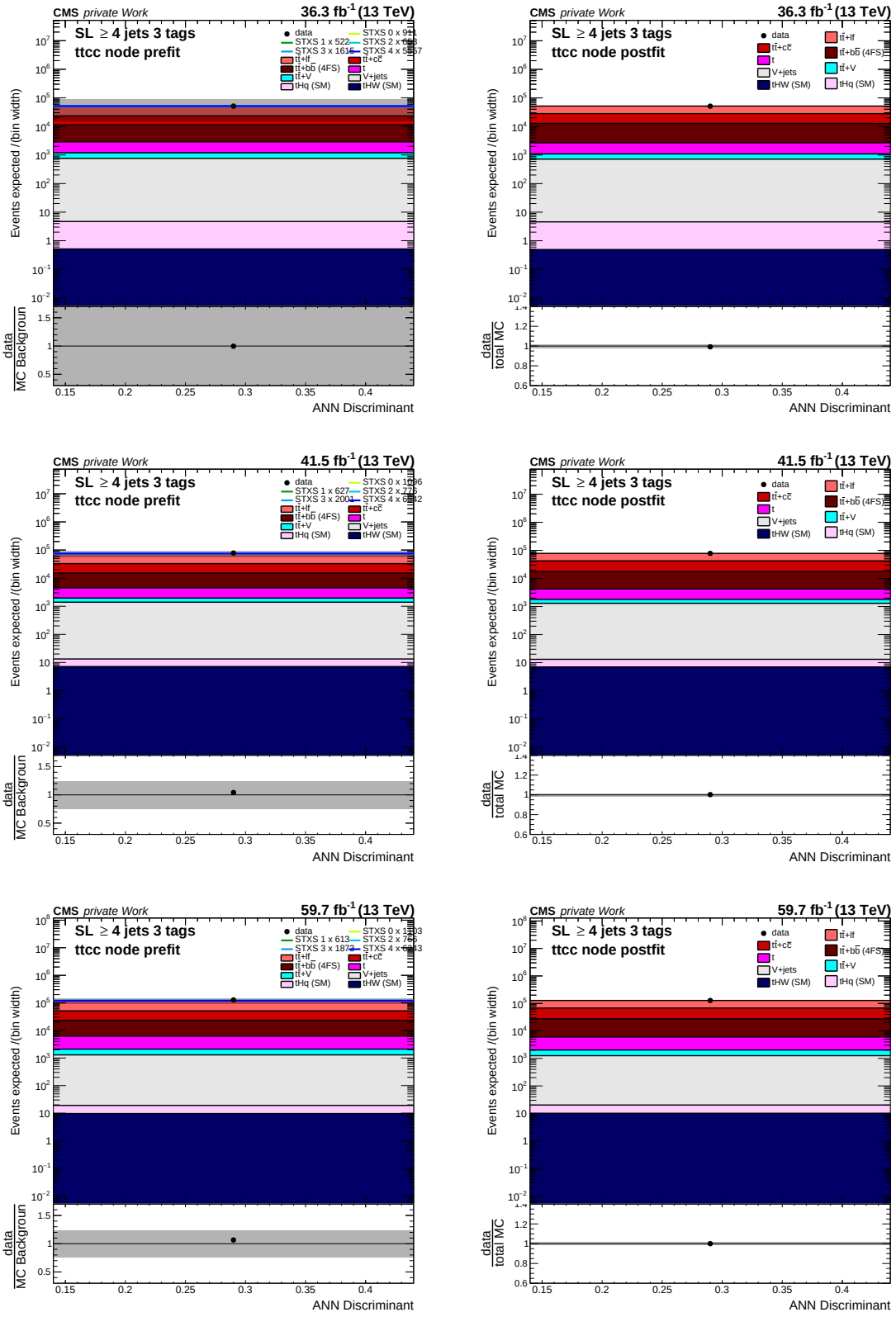


Figure H.36: **Discriminants in the SL channel in the ≥ 4j, 3t category $\mathrm{t\bar{t}} + \mathrm{c\bar{c}}$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

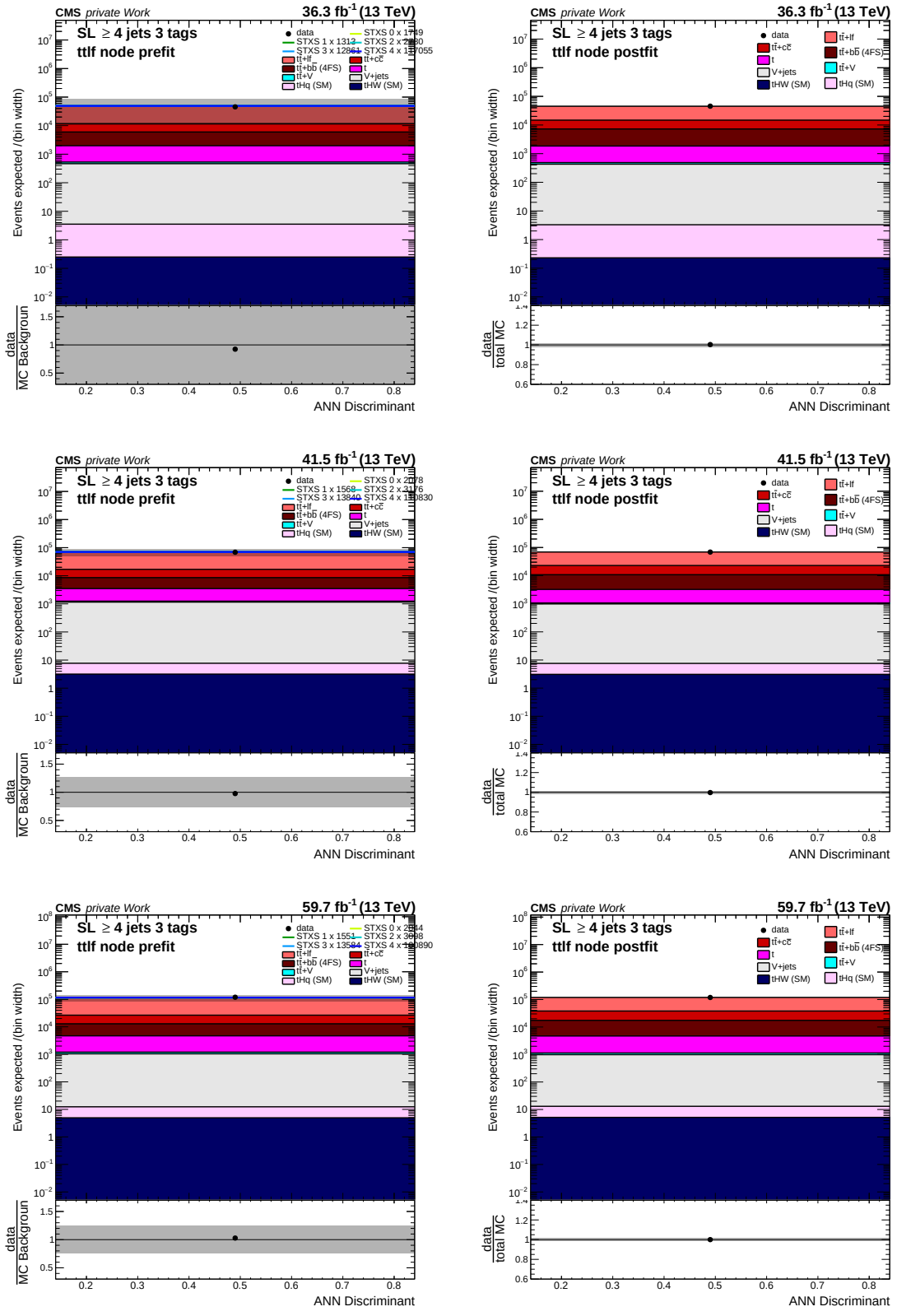


Figure H.37: **Discriminants in the SL channel in the $\geq$ 4j, 3t category $t\bar{t}$ + lf node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

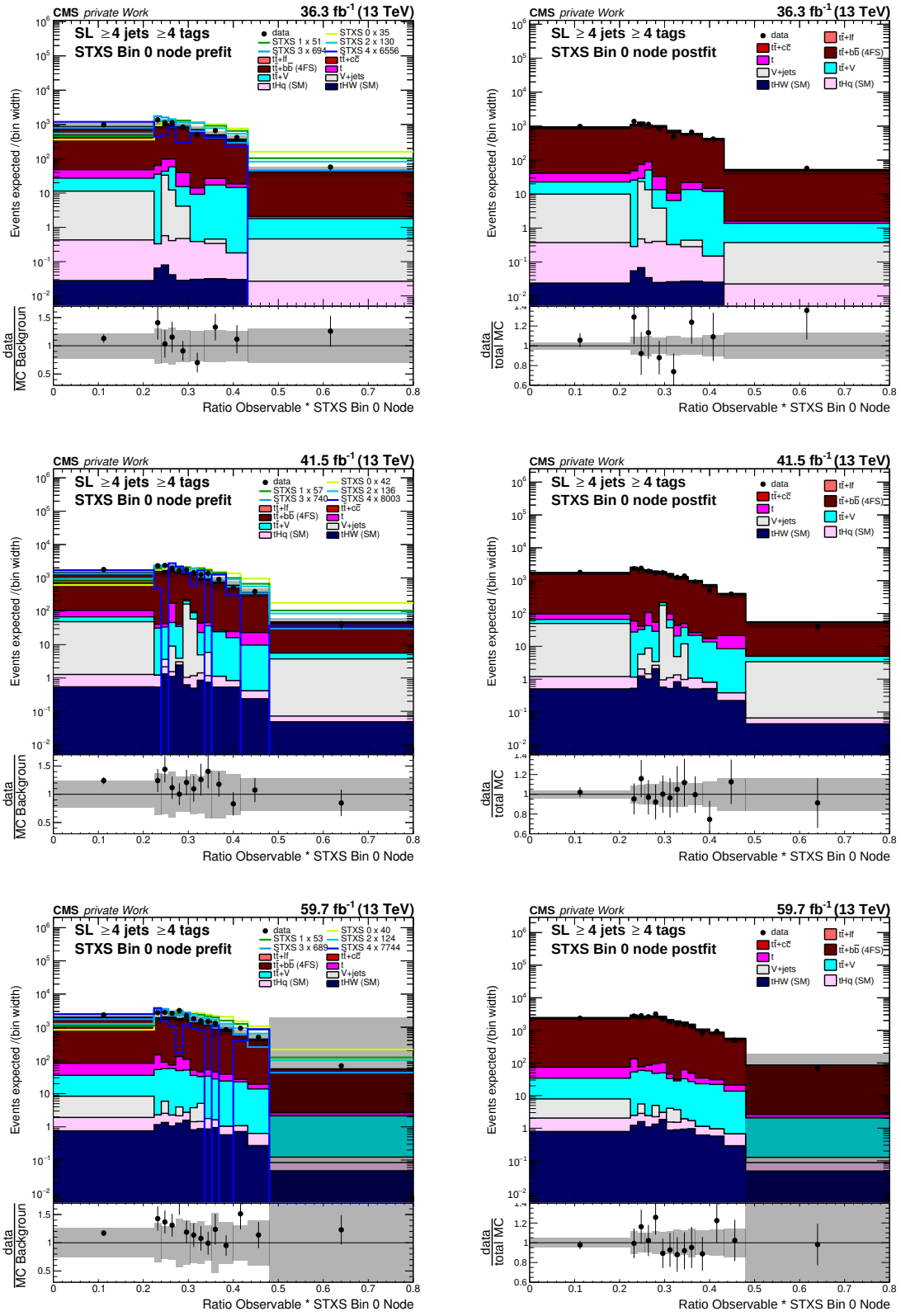


Figure H.38: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ratio observable times the STXS ANN output for the first STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

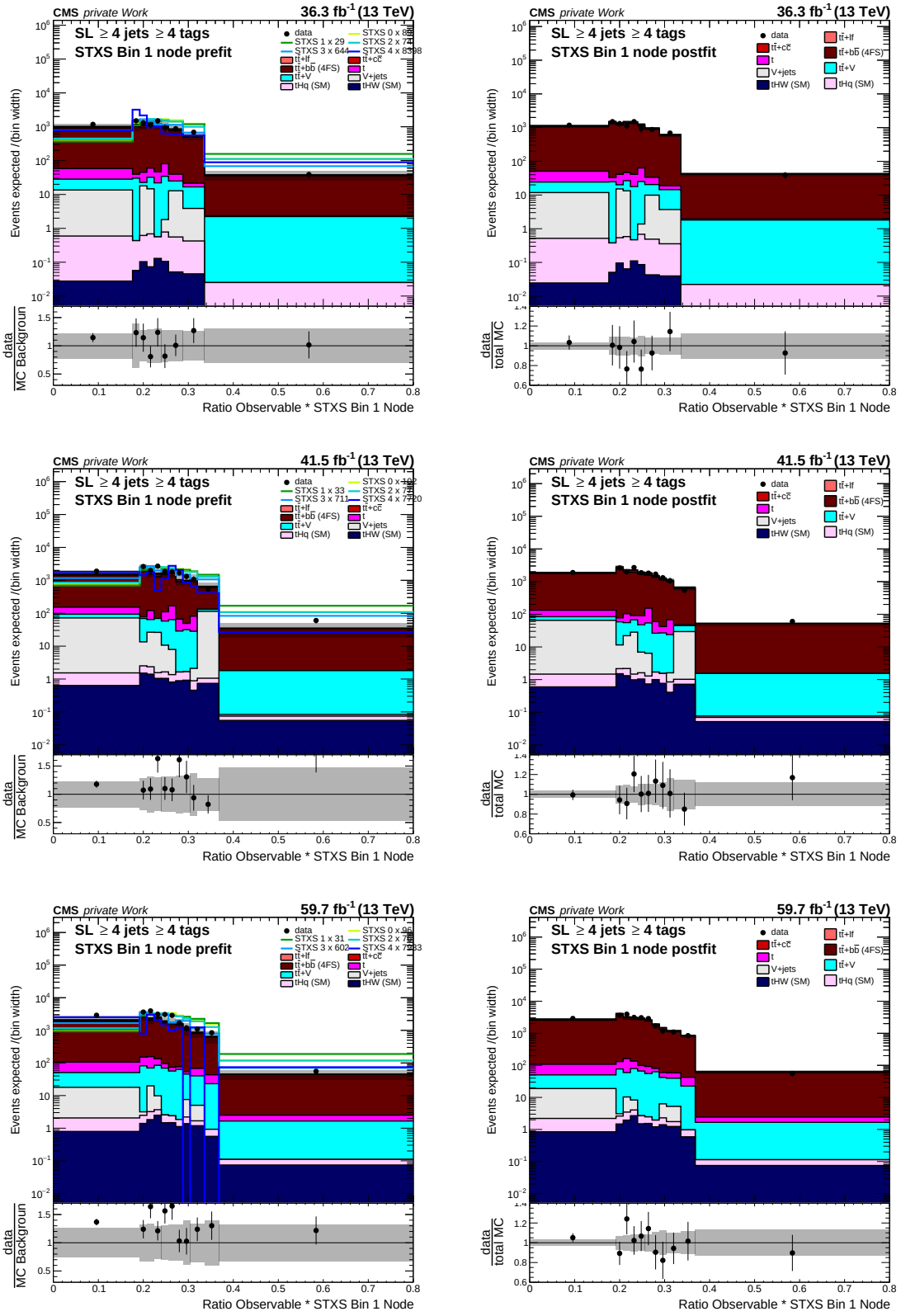


Figure H.39: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ratio observable times the STXS ANN output for the second STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

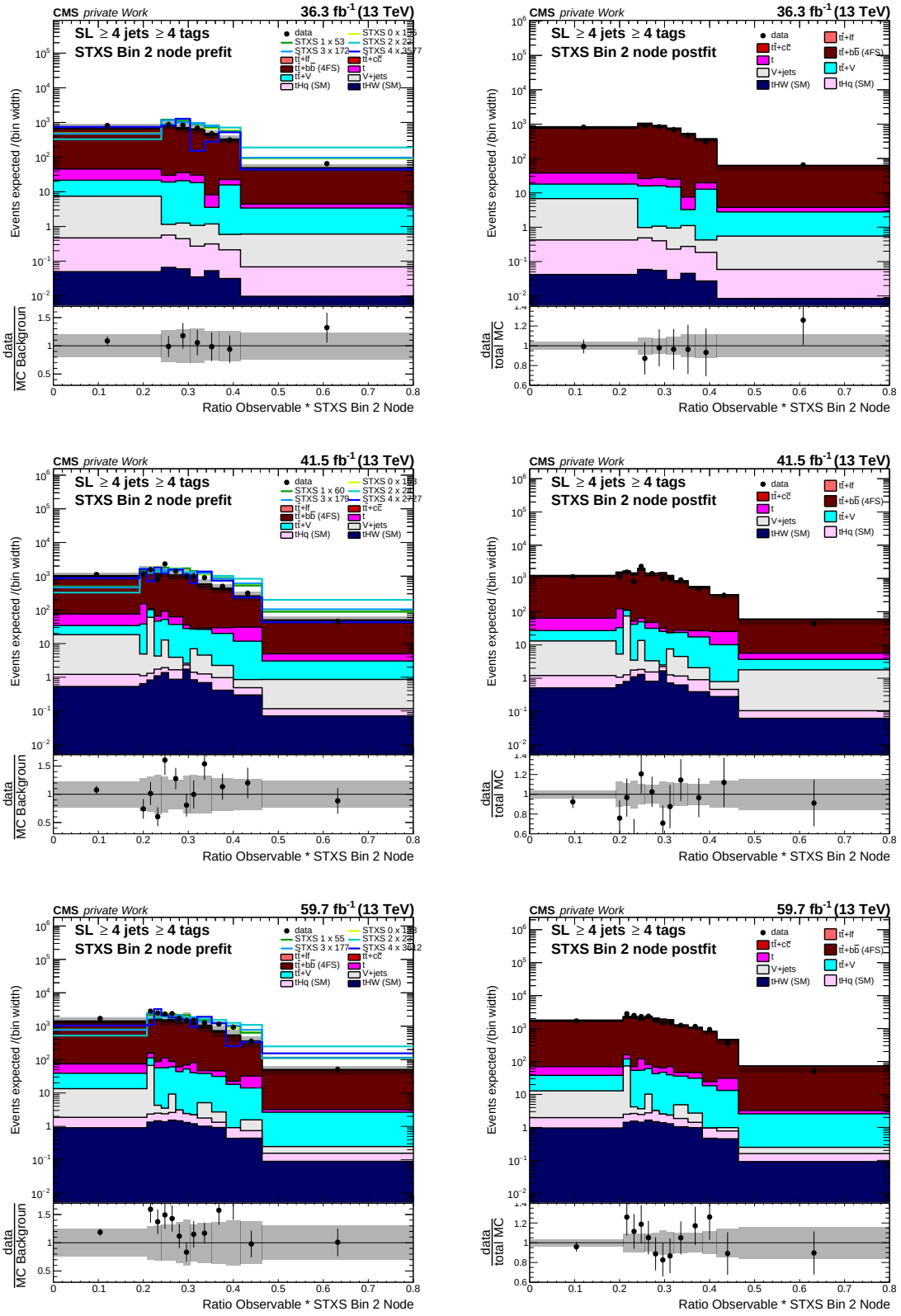


Figure H.40: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ratio observable times the STXS ANN output for the third STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

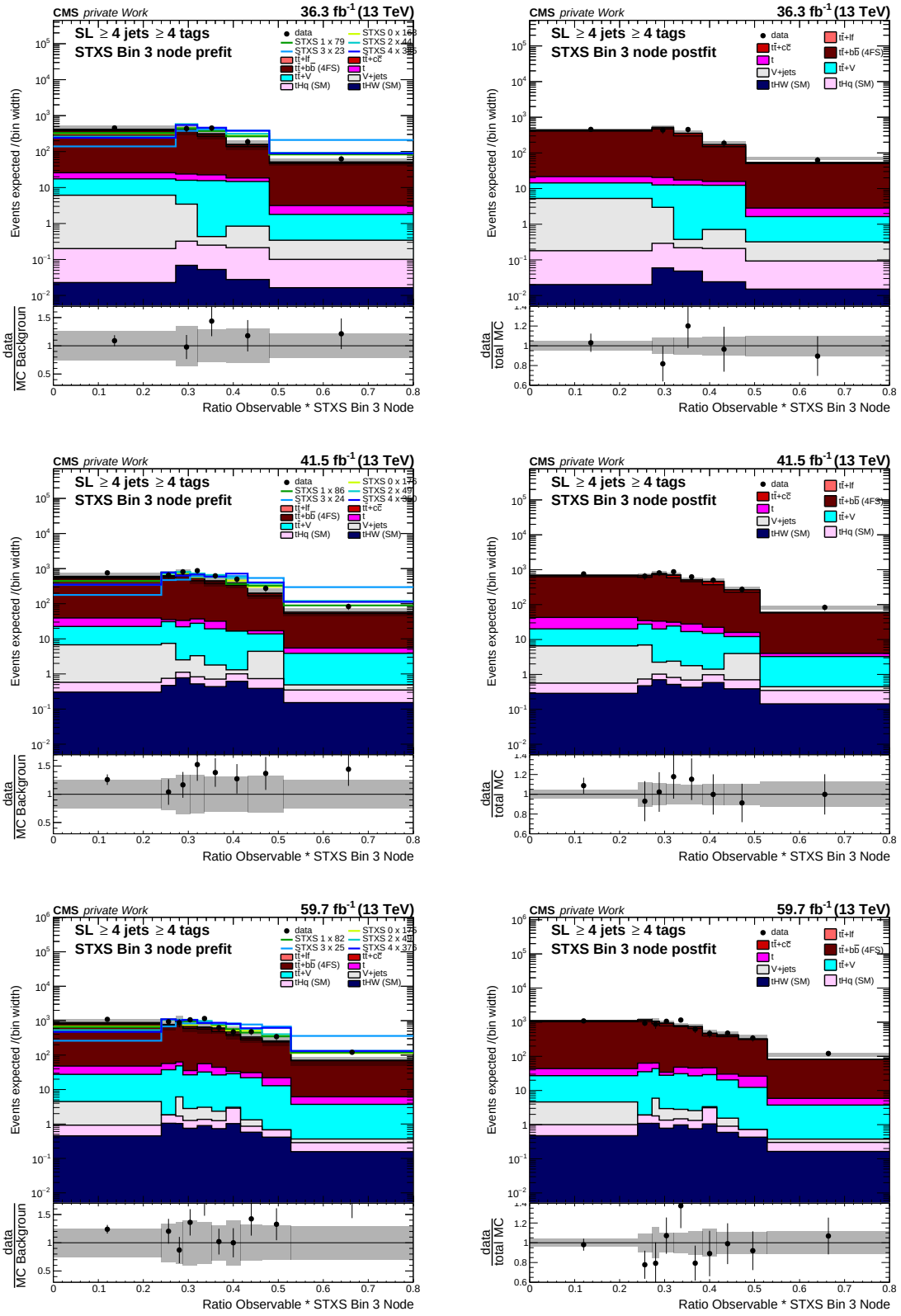


Figure H.41: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ratio observable times the STXS ANN output for the forth STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

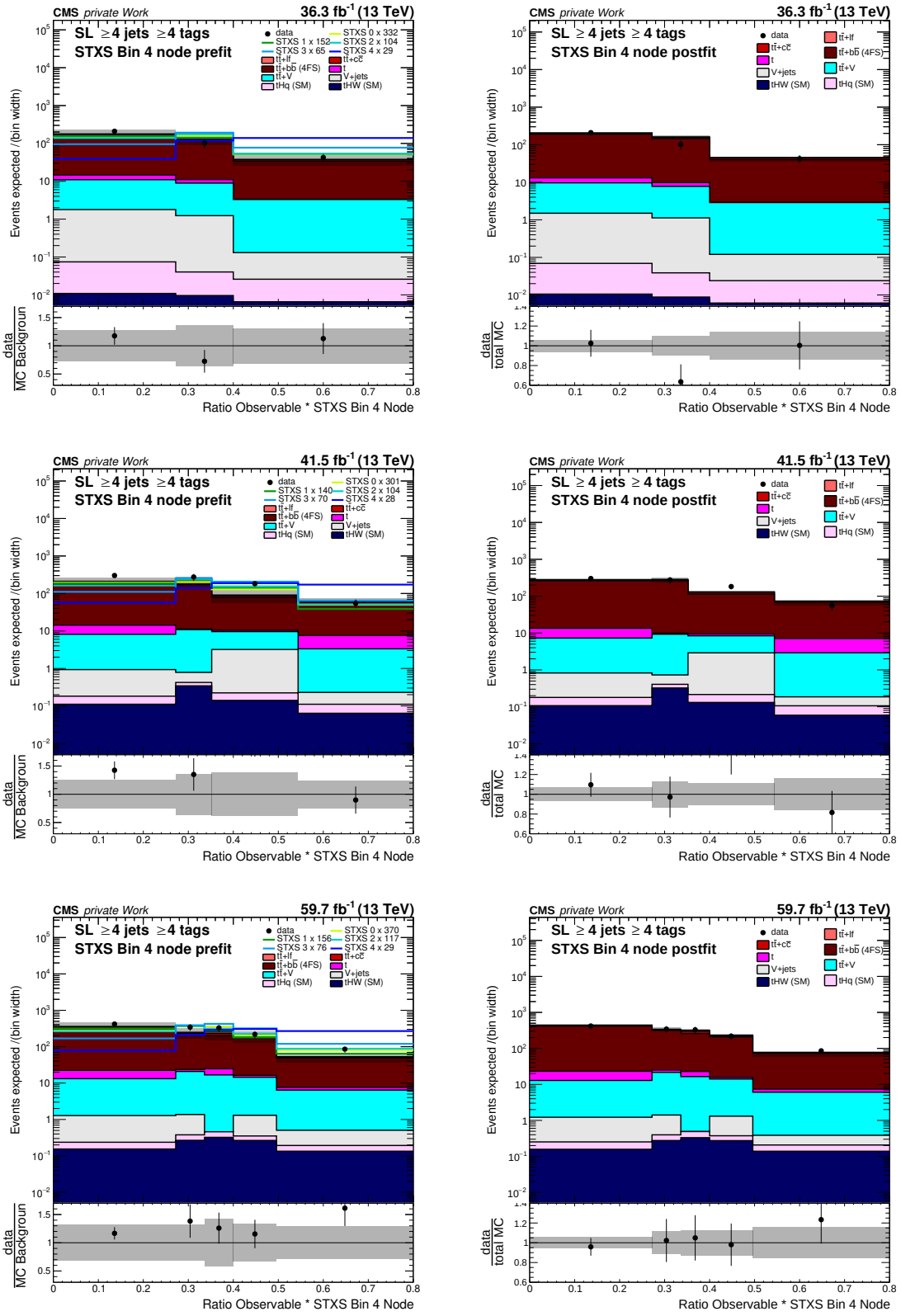


Figure H.42: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category ratio observable times the STXS ANN output for the last STXS bin category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

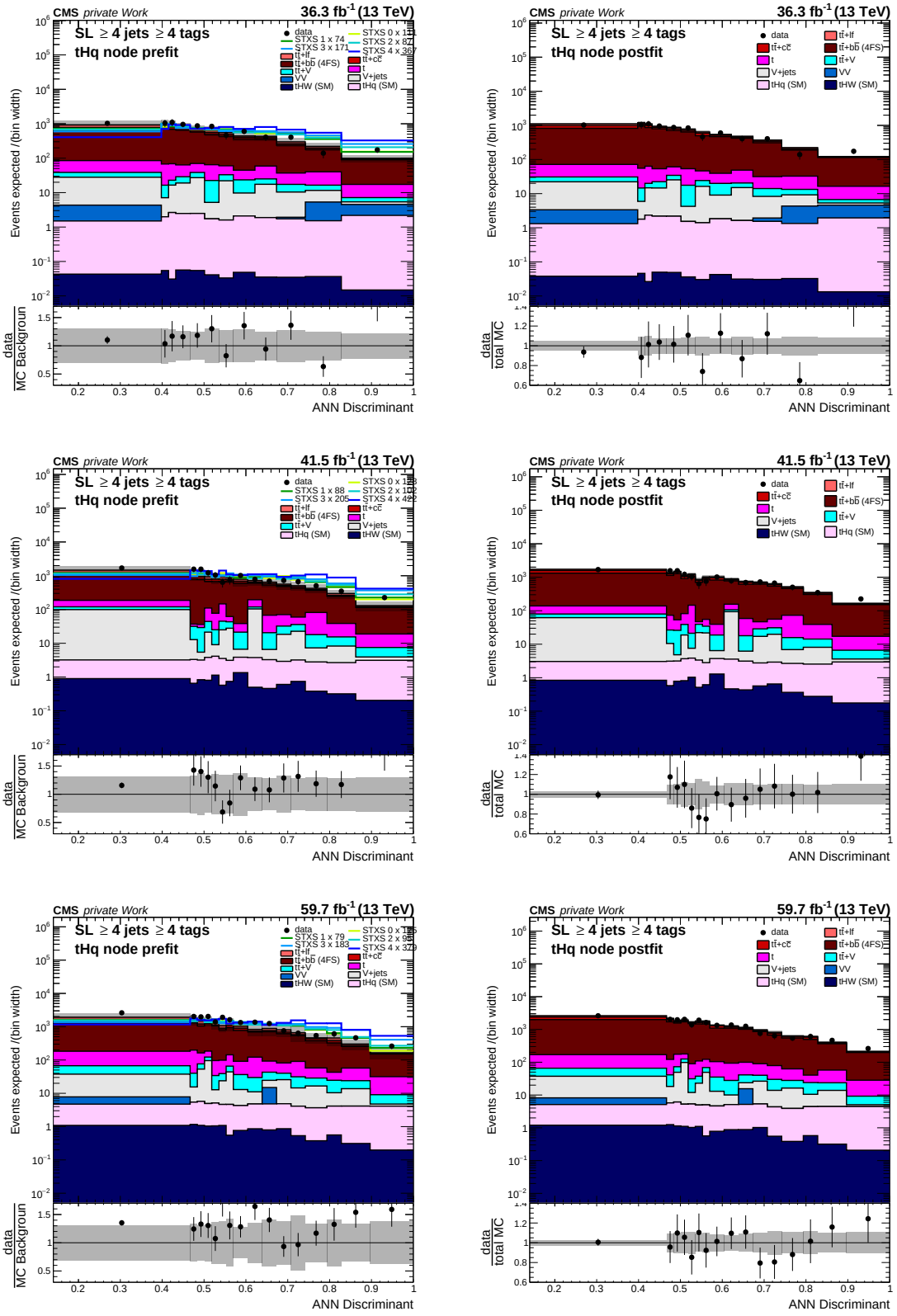


Figure H.43: **Discriminants in the SL channel in the ≥ 4j, 3t category tHq node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

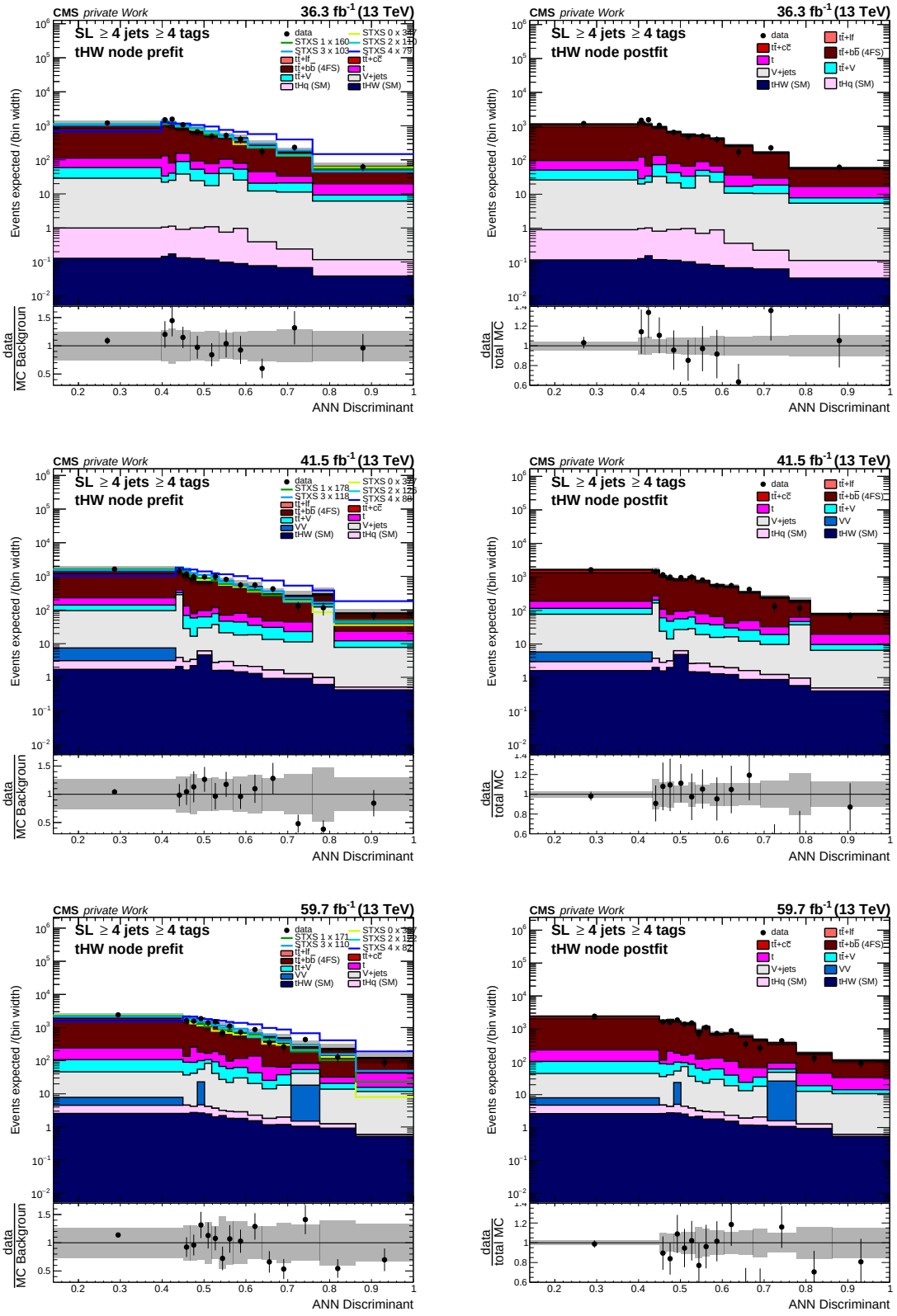


Figure H.44: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category tHW node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

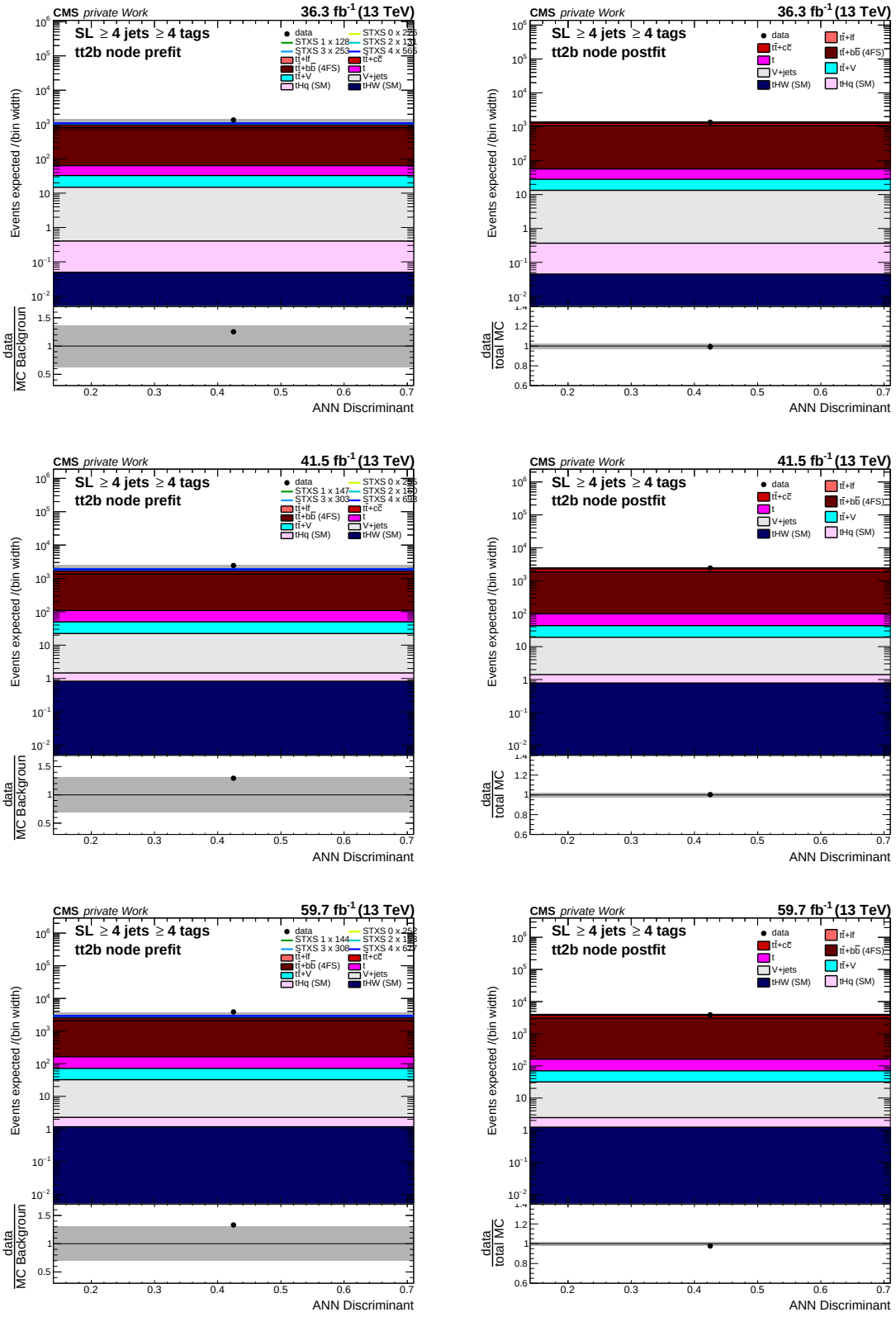


Figure H.45: **Discriminants in the SL channel in the ≥ 4j, ≥ 4t category $t\bar{t}$ + 2b node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

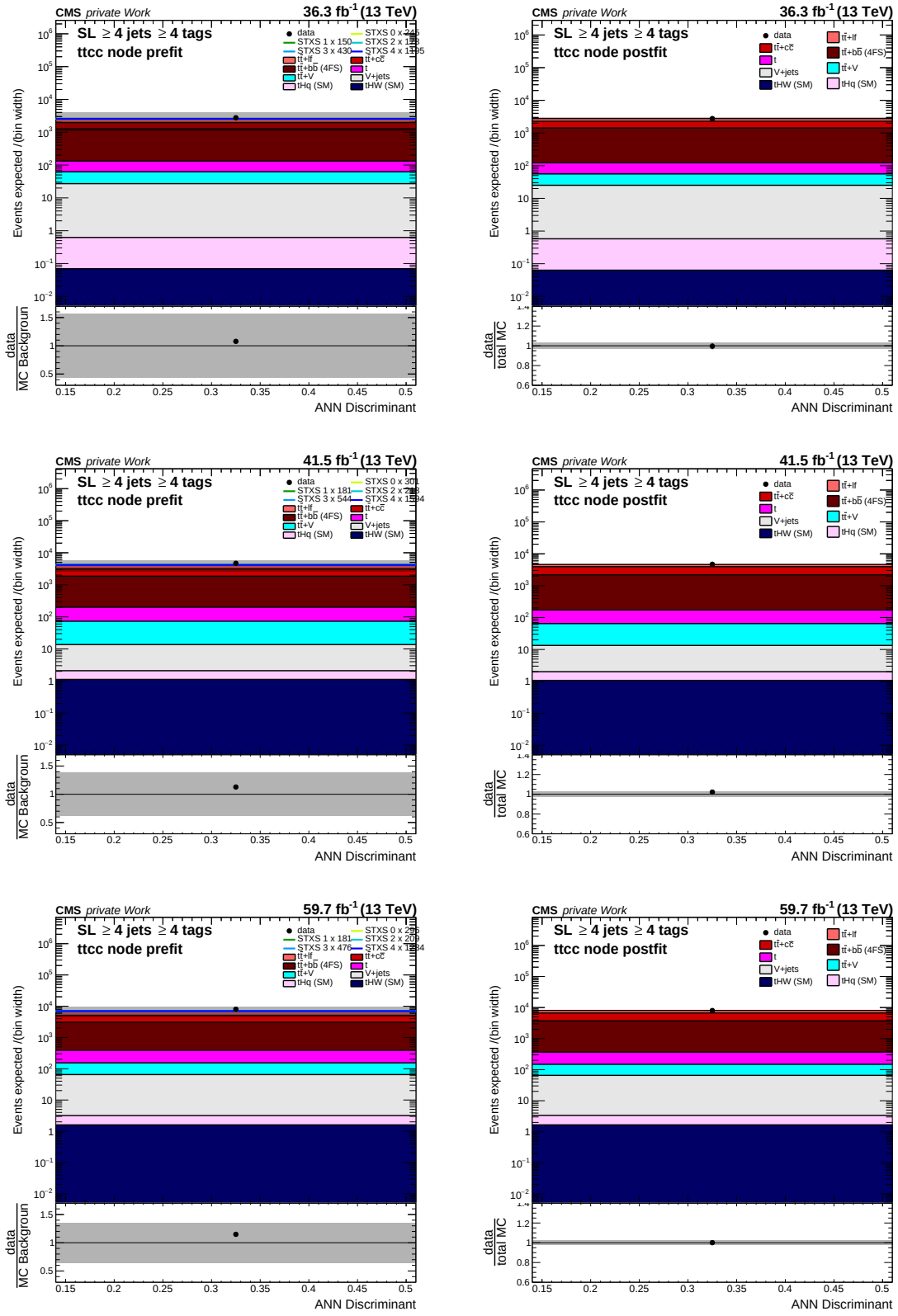


Figure H.46: **Discriminants in the SL channel in the $\geq$ 4j, $\geq$ 4t category $\mathrm{t\bar{t}} + \mathrm{c\bar{c}}$ node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

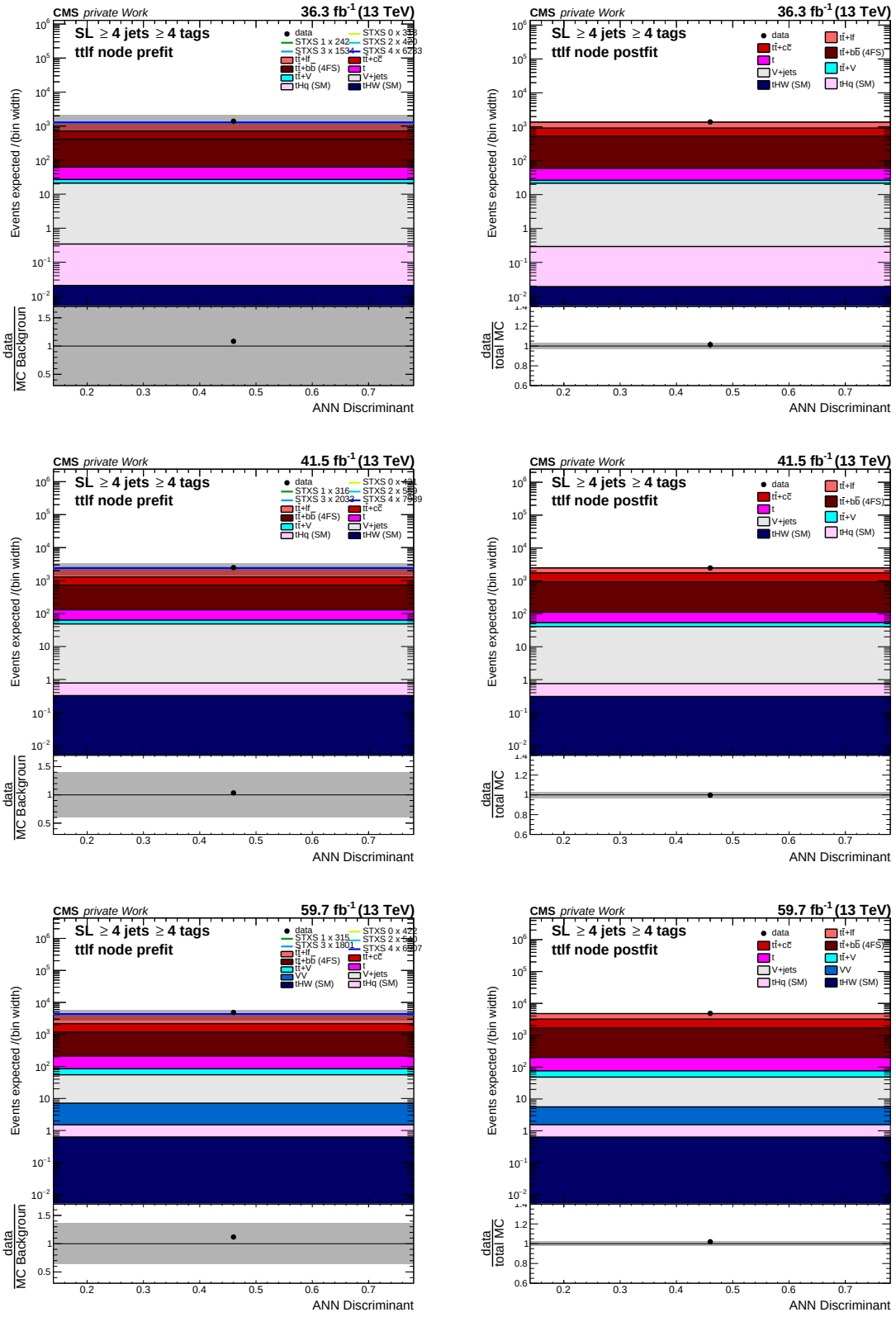


Figure H.47: **Discriminants in the SL channel in the $\geq$ 4j, $\geq$ 4t category $t\bar{t}+$lf node category**. Shown are the pre-fit (left) and post-fit (right) distributions for 2016 (top), 2017 (middle), and 2018 (bottom). The black points represent the observed data. The stacked distributions correspond to the simulations for the respective processes. The uncertainty band corresponds to the uncertainties in the full Run-II SL combination. For better visibility, the signal templates are scaled to the total background contribution in the pre-fit figures (colored lines).

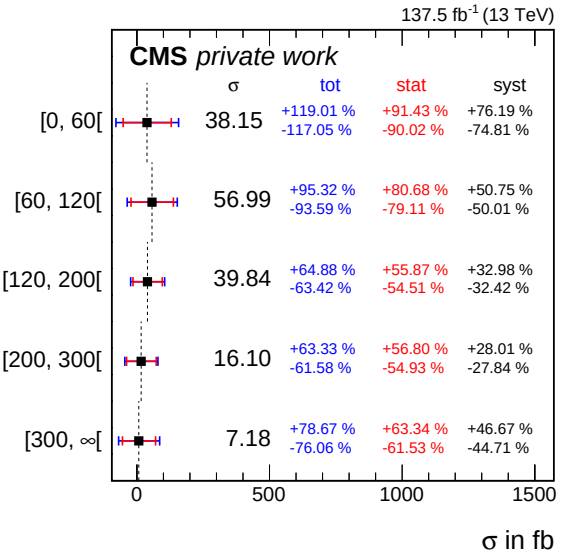


Figure H.48: **Expected sensitivity for the cross section measurement in the combination of the DL+SL channels**. The uncertainties are extracted in a fit to pseudo data corresponding to the full Run-II data set.

Acknowledgements

Zunächst möchte ich Prof. Dr. Ulrich Husemann für die Betreuung im Rahmen meiner Master- und Doktorarbeit und die gute Zusammenarbeit danken, die mich die letzten Jahre geprägt hat. Im Rahmen meiner Promotion hatte ich zahlreiche Gelegenheiten mich sowohl fachlich als auch persönlichen weiterzubilden, was nur durch seine Unterstützung möglich war. Dafür ein herzliches Dankeschön!

Des Weiteren möchte ich Prof. Dr. Thomas Müller für die Übernahme des Korreferats bedanken. Durch seine Unterstützung während des Studiums und meiner Bachelorarbeit bin ich auf diesen Weg gekommen, der nun in meiner Promotion endet. Auch dafür herzlichen Dank.

Außerdem will ich mich an dieser Stelle bei Dr. Matthias Schröder bedanken, der mich während meiner Zeit sowohl fachlich als auch persönlich immer unterstützt hat. Die Zusammenarbeit hat immer viel Spaß gemacht und mich stark in meinem Werdegang als Wissenschaftler geprägt.

Ich bedanke mich bei der ganzen Datenanalysegruppe von Prof. Husemann am ETP für die schöne Zeit am Institut und darüber hinaus. Ein spezieller Dank gilt Dr. Michael Waßmer, Sebastian Wieland und Jan van der Linden. Unsere gemeinsame Zeit in R418 werde ich nie vergessen. Darüber hinaus will ich mich bei Sebastian Wieland, Jan van der Linden, Nikita Shadskiy, Dr. Michael Waßmer, Dr. Karim El Morabit und Dr. Matthias Schröder für das Korrekturlesen dieser Arbeit bedanken. An dieser Stelle möchte ich mich auch beim ganzen ETP, dem Graduiertenkolleg GRK 1694 und KSETA für die persönliche, fachliche und finanzielle Unterstützung bedanken.

Ich möchte mich besonders herzlich bei Jan van der Linden und Sebastian Wieland für die sehr gute Zusammenarbeit im Rahmen der $t\bar{t}H(b\bar{b})$ Analyse bedanken. Gerade Sebastians Mitarbeit am Ende meiner Promotion war essenziell und eine enorme Unterstützung. Auf dich kann man sich immer verlassen! Außerdem möchte ich Dr. Matthias Schröder und Dr. Michael Waßmer für ihre überragende Unterstützung gerade am Ende der Promotion danken, die auch des Öfteren meine flatternden Nerven wieder beruhigen konnte.

Additionally, I would like to thank the whole $t\bar{t}H(b\bar{b})$ team at the CMS experiment for the good collaboration during my studies, which was essential in many aspects to produce the results presented in this thesis.

Zu guter Letzt will ich mich bei meiner Familie und meinen Freunden bedanken. Danke an meine Eltern Gaby und Ralf und meinen Bruder Felix für die Unterstützung und dafür, dass ihr immer ein offenes Ohr für mich habt. Besonders bedanken möchte ich mich bei meiner Freundin Rebecca, die eine wirklich sehr große Stütze gerade am Ende meiner Promotion war und mir auch durch ihre überragenden Theorie-Kenntnisse immer weiterhelfen konnte. Ich freue mich sehr auf unsere gemeinsame Zeit in Hamburg! Außerdem möchte ich mich bei meinen Freunden Matthias, Daniel, Kevin, Sabrina, Martin, Fabian, Simon, Ann-Kathrin und Laura für schöne Stunden auch in nicht immer schönen Zeiten bedanken.