

## Measurements of $t\bar{t}H$ and $tH$ production at CMS

---

**Cristina Martin Perez**<sup>\*†</sup>

*Laboratoire Leprince-Ringuet (CNRS), France*

*E-mail:* [cmartinp@cern.ch](mailto:cmartinp@cern.ch)

Results of the search for the Higgs boson produced in association with either one top quark ( $tH$ ) or a top quark pair ( $t\bar{t}H$ ) at the CMS experiment are presented. Dedicated searches of the Higgs decaying to a  $b\bar{b}$  pair, multileptons ( $WW$ ,  $ZZ$ ,  $\tau\tau$ ) and a  $\gamma\gamma$  pair are shown in this document. The analyzed data correspond to an integrated luminosity of  $77.4 \text{ fb}^{-1}$  recorded in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$  center-of-mass energy in 2016 and 2017. The sensitivity of the searches is enhanced by using matrix element and machine learning methods to reconstruct the events and to separate the signals from the respective backgrounds. Standalone evidence of the targeted final states of the Standard Model (SM)  $t\bar{t}H$  process has been found and interpretations of the  $tH$  process in beyond the SM scenarios are shown.

*European Physical Society Conference on High Energy Physics - EPS-HEP2019 -  
10-17 July, 2019  
Ghent, Belgium*

---

<sup>\*</sup>Speaker.

<sup>†</sup>On behalf of the CMS Collaboration

## 1. Introduction

The discovery of a Higgs (H) boson by the CMS [1] and ATLAS [2] experiments [3, 4, 5] in 2012 opened a new field of exploration in particle physics. A precise study of the properties of this particle is essential in order to verify that the discovered particle is indeed the one predicted by the Standard Model (SM). Among the different Higgs properties, of particular interest is the Yukawa coupling of the H boson to the top quark (t). In the SM, the coupling  $y_f$  of the H boson to fermions is proportional to the mass  $m_f$  of the fermion, namely  $y_f = m_f/v$ , where  $v = 246$  GeV denotes the vacuum expectation value of the Higgs field. Being the top quark the heaviest fermion known to date ( $m_t = 173.34 \pm 0.76$  GeV), it might have an important role in the electroweak symmetry breaking mechanism, and any deviation of this coupling from the SM prediction would indicate the presence of new physics.

The measurement of the rate at which H bosons are produced in association with top quark pairs ( $t\bar{t}H$ ) provides the most precise model-independent determination of the magnitude of  $y_f$ . The SM cross section for  $t\bar{t}H$  production in proton-proton collisions (pp) at  $\sqrt{s} = 13$  TeV center-of-mass energy amounts to 506.5 fb, computed at next-to-leading (NLO) accuracy in quantum chromodynamics (QCD) with electroweak corrections computed at the same order in perturbation theory. The cross section for the associated production of a H boson with a single top quark (tH) is significantly smaller, amounting to 74.3 fb in the SM, computed at NLO accuracy in QCD with the so-called five-flavour scheme (5FS), where bottom quarks are considered as sea quarks of the proton and may appear in the initial state of the pp scattering process.

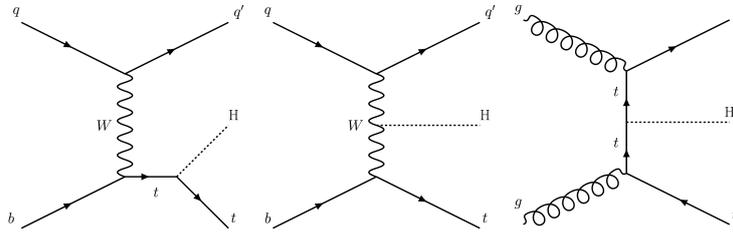


Figure 1: Leading-order Feynman diagrams for tH production for the t-channel in 5FS (two left-most diagrams) and for the  $t\bar{t}H$  process (right diagram).

The leading order (LO) Feynman diagrams for tH and  $t\bar{t}H$  can be found in Fig. 1. The diagrams for the tH process correspond to the t-channel production mode, which is the dominant one. In this case, the Higgs can couple to the top quark or to the W boson. The rate of the tH process is described by the interference between these two diagrams. In the SM, both diagrams interfere destructively and therefore this rate is very small. In case  $y_t$  differs from the value predicted by the SM, the cross section may be enhanced significantly. The largest enhancement is achieved in the case where  $y_t$  is negative, referred to as the inverted top coupling (ITC) scenario, where the interference is constructive and the cross section increases by a factor 10.

This paper is structured as follows. Section 2 describes the searches for the Higgs boson produced in association with two top quarks, and decaying into a  $b\bar{b}$  pair [6], multileptons (arising from H to WW, ZZ or  $\tau\tau$  decays) [7] and a  $\gamma\gamma$  pair [8]. These measurements are based on data recorded by the CMS experiment in pp collisions at  $\sqrt{s} = 13$  TeV center-of-mass energy in 2016

and 2017, corresponding to an integrated luminosity of  $77.4 \text{ fb}^{-1}$ . Section 3 describes both the SM and beyond the SM searches of the Higgs boson produced in association with a single top quark [9] based on the data recorded by the CMS experiment in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  center-of-mass energy in 2016, which amounts to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

## 2. Measurements of $t\bar{t}H$ production

The event topology of the  $t\bar{t}H$  process is that of two top quarks and a Higgs boson. On the one hand, the top quark decays into W bosons and b quarks, which hadronize into what is referred to as b-jets. The decays of these W bosons, which can either be leptonic or hadronic, determine the specific  $t\bar{t}$  signatures recorded in the detector. On the other hand, the Higgs boson can decay into  $b\bar{b}$ , WW, ZZ,  $\tau\tau$  or  $\gamma\gamma$  pairs. These decays have different branching ratios and can produce different jet, lepton and photon multiplicities and kinematics, which are identified with dedicated techniques in each final state.

### 2.1 $t\bar{t}H$ with the H decaying to a $b\bar{b}$ pair

In the SM, the  $H \rightarrow b\bar{b}$  decay has the largest branching fraction (58%) and is thus very attractive experimentally. This analysis covers all three  $t\bar{t}$  decay channels: the fully-hadronic channel, where both W bosons decay into quarks, the semi-leptonic channel, where one W boson decays into a charged lepton and a neutrino and the other one decays into quarks, and the fully-leptonic channel, where both W bosons decay into a charged lepton and a neutrino. These final states contain a large number of high- $p_T$  jets and a few charged leptons, as well as missing transverse energy in the leptonic channels as a result of the presence of neutrinos.

The dominant background contributions arise from the QCD multijet production in the fully-hadronic channel, and from  $t\bar{t}$  + jets production in all channels. The latter includes  $t\bar{t}$ +light-flavour jets where one or more jets is misidentified as a b-jet, as well as the  $t\bar{t}+c\bar{c}$  and  $t\bar{t}+b\bar{b}$  backgrounds, where additional b or c quarks can arise from QCD radiation or loop-induced QCD processes. The  $t\bar{t}+b\bar{b}$  background, in particular, remains almost irreducible with respect to the signal.

The sensitivity of the analysis is limited by a combinatorial background due to multiple b-jets in the final state, with no unambiguous way of reconstructing the invariant mass of the Higgs boson. For that reason the signal is extracted exploiting multivariate analysis techniques in each of the channels separately. The Matrix Element Method (MEM) is used in the fully-hadronic channel, whereas an Artificial Neural Network (ANN) and Boosted-Decision-Trees (BDT) are used in the semi-leptonic and fully-leptonic channels, respectively. The signal strength modifier  $\mu = \sigma/\sigma_{SM}$  is determined in a simultaneous binned profiled likelihood fit of the expected background and signal distributions of the aforementioned discriminant outputs to the data across all the channels. The systematic uncertainties, which arise mainly from the  $t\bar{t}+b\bar{b}$  and  $t\bar{t}H$  modelings, are taken into account using nuisance parameters, which allow for variation in the shape and normalisation of the final discriminant distributions during the fit. The obtained best fit value of the signal strength modifier  $\mu$  is  $\hat{\mu} = 1.15^{+0.32}_{-0.29}$ , which corresponds to an observed (expected) significance of 3.9 (3.5) standard deviations above the background-only hypothesis. The best fit values for each channel separately and in the combination are shown in Fig. 2.

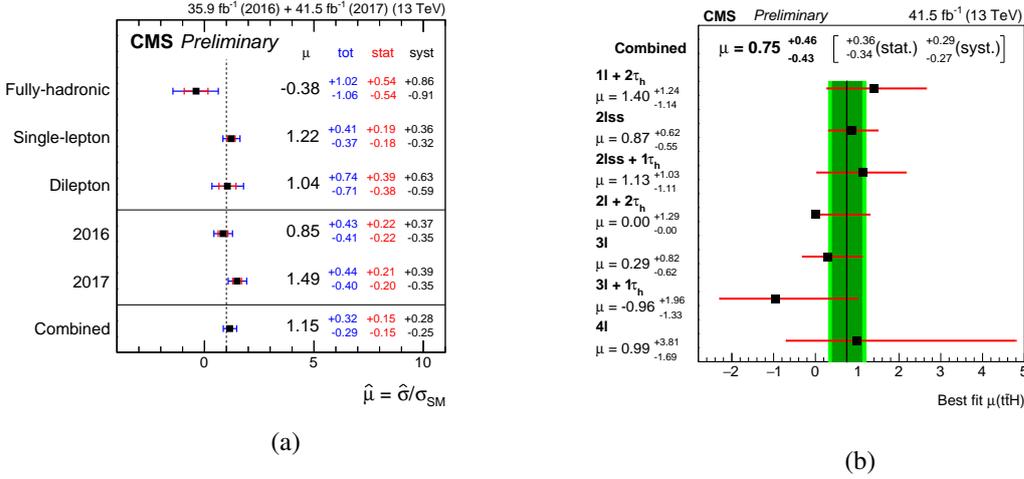


Figure 2: (a) Best fit values of the signal strength modifier  $\mu$  for the  $t\bar{t}H \rightarrow b\bar{b}$  process obtained in the combined fit for the 2016 and 2017 datasets per channel and in the full combination [6]. (b) Best fit values of the signal strength modifier  $\mu$  for the  $t\bar{t}H \rightarrow WW, ZZ, \tau\tau$  processes obtained in the combined fit for the 2017 dataset per channel and in the full combination [7].

## 2.2 $t\bar{t}H$ with the H decaying to multileptons

The H to multileptons channel covers the processes in which the H boson decays to  $WW, ZZ$  or  $\tau\tau$ , which contain multiple electrons, muons and hadronically decaying  $\tau$  leptons ( $\tau_h$ ) in the final state. These processes have intermediate branching ratios (21% for  $WW$ , 3% for  $ZZ$  and 6% for  $\tau\tau$ ), but cleaner final states than in the  $H \rightarrow b\bar{b}$  case. The analysis is divided in seven mutually exclusive channels with different number of light leptons (electrons and muons) and  $\tau_h$  reconstructed in the event. All final states contain multiple isolated high- $p_T$  leptons with different charge and flavor configurations, in addition to two b-jets and further light jets, and missing transverse energy due to the neutrinos present in the W and  $\tau$  decays. Light leptons, which are a key element to this analysis, are identified with multivariate techniques based on BDTs that discriminate those originating from W, Z or  $\tau$ , referred to as prompt leptons, against those arising from B or lighter mesons, referred to as non-prompt leptons.

The dominant irreducible backgrounds in most of the channels are due to the production of top quarks in association with W and Z bosons ( $t\bar{t}W, t\bar{t}Z$ ) and, to a lower extend, diboson+jets production processes ( $WW$ +jets,  $WZ$ +jets,  $ZZ$ +jets). Additionally, there is a considerable amount of the so-called reducible background, which arises mainly from  $t\bar{t}$ +jets processes in which at least one reconstructed electron or muon is due to the misidentification of a non-prompt lepton or hadron, or in which at least one reconstructed  $\tau_h$  arises from the misidentification of a quark or gluon jet. Additional reducible background contributions are due to the charge mis-identification of electrons mainly in the purely leptonic channels.

Due to the large multiplicity of objects in the final state and the different background contributions in each of the categories, the signal is extracted exploiting multivariate techniques based on BDTs separately for each of the seven channels. Where applicable, additional BDTs are used

to identify tops that decay hadronically and jets that arise from the Higgs boson decays to  $WW$ . The signal strength modifier  $\mu$  is determined in a simultaneous binned profiled likelihood fit of the expected background and signal discriminant output to the data across all the channels. The systematic uncertainties, mainly arising from the  $t\bar{t}H$  and  $t\bar{t}V$  modeling, as well as the normalization of the reducible backgrounds, are taken into account using nuisance parameters. The obtained best fit value of the signal strength modifier  $\mu$  is  $\hat{\mu} = 0.96^{+0.34}_{-0.31}$ , which corresponds to an observed (expected) significance of 3.2 (4.0) standard deviations above the background-only hypothesis. The best fit values for each channel separately and in the combination are shown in Fig. 2 for the data collected during 2017 data-taking period.

### 2.3 $t\bar{t}H$ with the H decaying to a $\gamma\gamma$ pair

The diphoton decay channel has a low decay branching ratio (0.2%), but plays a primary role in the study of the Higgs boson properties thanks to its clean final state, since the diphoton invariant mass ( $m_{\gamma\gamma}$ ) can be reconstructed with high precision in the CMS detector. The  $t\bar{t}H$  production is identified by the decay products of the top-antitop quark pair in the final state, making use of three separate categories depending on whether it decays fully hadronically, semi-leptonically or fully-leptonically. The signature left in the detector is that of two central high- $p_T$  isolated photons, as well as additional b-jets and lighter jets. The photons, which drive the sensitivity in this statistically-limited channel, are identified with a dedicated BDT which is trained to discriminate against jets.

The main irreducible backgrounds arise from photon pairs produced in association with jets or leptons as well as  $t\bar{t}$  pairs. The reducible component comes from  $\gamma$ +jet or jet+jet processes in which jets are misidentified as photons, together with additional jets, leptons or a  $t\bar{t}$  pair. Dedicated multivariate discriminants, based on variables related to the photons and to the additional objects, are used to discriminate the signal from the backgrounds in each category separately, as well as to further sub-categorize within each of the channels to enhance the sensitivity.

The signal is extracted based on the diphoton invariant mass distributions, which shows a narrow peak on a slowly falling background. The signal shape is modelled from simulation with a sum of Gaussian functions, while the background model is extracted from data with a discrete profiling method. The signal strength modifier  $\mu = \sigma/\sigma_{SM}$  is determined in a simultaneous binned profiled likelihood fit across all the channels. The systematic uncertainties, mainly arising from the QCD scale uncertainty and the experimental uncertainty on the photon identification, are taken into account using nuisance parameters. The obtained best fit value of the signal strength modifier  $\mu$  is  $\hat{\mu} = 1.7^{+0.6}_{-0.5}$ , which corresponds to an observed (expected) significance of 4.1 (2.7) standard deviations above the background-only hypothesis. The measurement is still largely dominated by the statistical uncertainty, but is a golden channel for the current and future studies of the  $t\bar{t}H$  process.

## 3. Measurements of $tH$ production

The event topology of the  $tH$  production is that of two heavy objects (the top quark and the Higgs boson) in the central portion of the detector recoiling against each other, while a light-flavor quark and a soft b quark escape in the forward-backward regions of the detector. This paper includes dedicated searches for the final states in which the H decays to a  $b\bar{b}$  pair or to

multileptons, as well as a reinterpretation of the  $t\bar{t}H$   $H \rightarrow \gamma\gamma$  measurements beyond the SM. The analysis techniques in terms of particle reconstruction, background rejection and signal extraction are very similar to those of the  $t\bar{t}H$  production presented here. The  $tH$  process, nonetheless, has lower particle multiplicity than the  $t\bar{t}H$  one, but profits from a more distinct signal-background separation thanks to the presence of the forward activity.

### 3.1 $tH$ with the $H$ decaying to a $b\bar{b}$ pair

Events from a  $tH$  signal where the Higgs boson decays to  $b\bar{b}$  produce final states with at least three central b-jets and a hard lepton from the top quark decay. The analysis is divided into two separate categories, grouping events with either three or four jets. As in the  $t\bar{t}H$  case, these final states have challenging backgrounds arising from  $t\bar{t}$ +jets production, which can be produced in QCD radiation or loop-induced QCD processes. A multivariate classification algorithm is trained to discriminate different  $t\bar{t}$ +jets background components in an opposite-sign leptons control region. Moreover, additional multivariate algorithms are used to optimize the jet-parton assignment and solve the combinatorics problem. For the signal extraction, the analysis makes use of separate BDTs in each of the two categories which are designed in a similar manner as in the  $t\bar{t}H$  case.

### 3.2 $tH$ with the $H$ decaying to multileptons

As in the  $t\bar{t}H$  analysis,  $tH$  events where the top quark decay produces leptons and the Higgs boson decays to  $WW$ ,  $ZZ$  or  $\tau\tau$  lead to final states containing multiple light leptons. Of particular interest among these are those with three or more charged leptons or with two leptons of the same charge, as they appear with comparatively low backgrounds. The background contributions are a mixture of  $t\bar{t}$ +jets with non-prompt leptons and events from associated production of  $t\bar{t}$  and  $W$  or  $Z$  bosons. This analysis separates the  $tH$  signal from these two backgrounds by training two multivariate classifiers and exploiting the features of the forward light jet, amongst others. The two classifier outputs are combined into a single binned distribution, which is then fitted to the data to extract the signal yield.

### 3.3 Combination of $tH$ final states

The dedicated  $tH$  searches in the  $H$  to  $b\bar{b}$  and  $H$  to multileptons final states, together with the  $t\bar{t}H$  category of the diphoton channel shown in Section 2.3, are used in combination to put constraints on the coupling modifier  $\kappa_t$ , defined as  $y_t/y_{t,SM}$ , and on the production cross section of the  $tH$  events. The discriminator output distributions of the  $b\bar{b}$  and multilepton channels, as well as the  $\gamma\gamma$  invariant mass distributions, are compared to the data in a combined maximum likelihood fit for various assumptions of the signal kinematics and normalizations depending on the  $\kappa_t$  values, and are used to derive constraints on the respective signal yields.

A profile likelihood scan is performed as a function of the coupling modifier  $\kappa_t$ , which affects the production cross sections of  $tH$  and  $t\bar{t}H$ , as well as the Higgs boson branching fractions. To derive the constraints on  $\kappa_t$ , a scan of the likelihood ratio  $\mathcal{L}(\kappa_t)/\mathcal{L}(\hat{\kappa}_t)$  is performed, where  $\hat{\kappa}_t$  is the best fit. Fig. 3 shows the negative of twice the logarithm of this likelihood ratio, for scans on the data and for an Asimov data set with SM expectations for  $t\bar{t}H$  and  $tH$ . The data slightly favors a positive value of  $\kappa_t$  over a negative one by about 1.5 standard deviations, while excluding values outside the ranges of about  $[0.9, 0.5]$  and  $[1.0, 2.1]$  at 95% confidence level (CL).

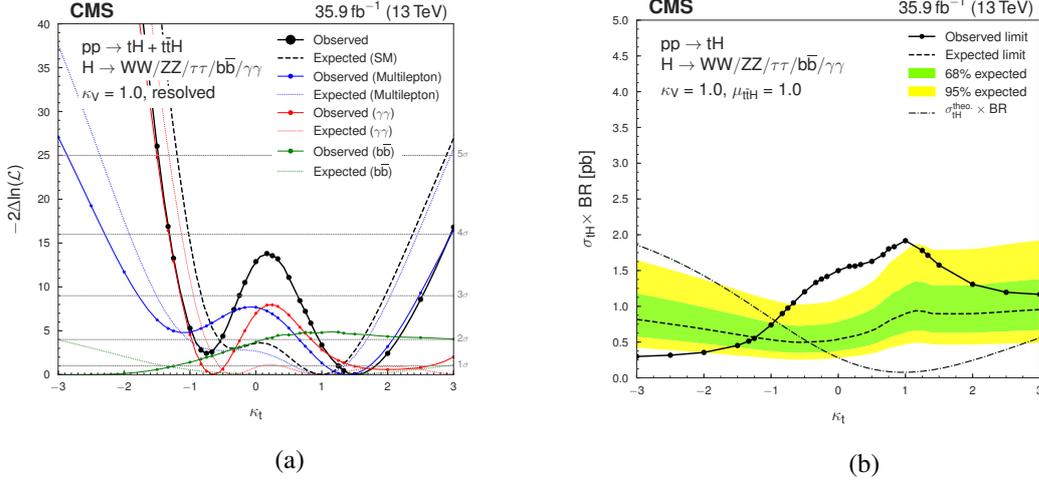


Figure 3: (a) Scan of  $-2\Delta\ln(\mathcal{L})$  versus  $\kappa_t$  for the data (black line) and the individual channels (blue, red and green), compared to Asimov datasets corresponding to the SM predictions (dashed lines) [9]. (b) Observed and expected 95% CL upper limit on the  $tH$  cross-section times the branching ratio for different values of the coupling ratio  $\kappa_t$ . The expected limit is calculated on a background-only dataset (without  $tH$  contribution) but including a  $\kappa_t$ -dependent contribution from  $t\bar{t}H$ . The  $t\bar{t}H$  normalization is kept fixed in the fit, while the  $tH$  signal strength is allowed to float [9].

To establish limits on  $tH$  production, a different signal strength parameter is introduced, in this case not including  $t\bar{t}H$ . A maximum likelihood fit for this  $tH$  signal strength is performed based on the profile likelihood at fixed points of  $\kappa_t$ , while the  $t\bar{t}H$  contribution is kept fixed to its  $\kappa_t$ -dependent expectation. The upper limits on this signal strength are shown in Fig. 3. The fiducial cross section for SM-like  $tH$  production is limited to about 1.9 pb, with an expected limit of 0.9 pb, corresponding, respectively, to about 25 and 12 times the expected SM cross section times branching fraction in the combination of the channels included.

#### 4. Conclusions

The searches for the Higgs boson produced in association with a single top ( $tH$ ) and a top quark pair ( $t\bar{t}H$ ) in the  $b\bar{b}$ , multileptons and  $\gamma\gamma$  final states have been presented. The search for  $t\bar{t}H$  has been performed in  $77.4 \text{ fb}^{-1}$  of pp collision data recorded with the CMS detector at a center-of-mass energy of 13 TeV in 2016 and 2017, whereas the search for  $tH$  has been performed in  $41.5 \text{ fb}^{-1}$  collected in 2016. Candidate events are selected in mutually exclusive categories depending on the top and Higgs decays. Multivariate discriminants are used to enhance the sensitivity and extract the signal. The best fit values of the  $t\bar{t}H$  signal cross section are all in agreement with the SM expectations and evidence for each of the final states alone has been found. The measurements of the  $tH$  process show that data favor positive over negative values of the top quark Yukawa coupling, and upper limits on the SM-like  $tH$  process have been established.

## References

- [1] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 S08004 (2008).
- [2] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 S08003 (2008).
- [3] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, JHEP 06 (2013) 081, doi:10.1007/JHEP06(2013)081.
- [4] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) 30, doi:10.1016/j.physletb.2012.08.021.
- [5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B 716 (2012) 1, doi:10.1016/j.physletb.2012.08.020.
- [6] CMS Collaboration, *Measurement of  $t\bar{t}H$  production in the  $H \rightarrow b\bar{b}$  decay channel in 41.5 fb<sup>1</sup> of proton-proton collision data at  $\sqrt{s} = 13$  TeV*, CMS PAS HIG-18-030 [<https://cds.cern.ch/record/2675023>] (2019).
- [7] CMS Collaboration, *Measurement of the associated production of a Higgs boson with a top quark pair in final states with electrons, muons and hadronically decaying  $\tau$  leptons in data recorded in 2017 at  $\sqrt{s} = 13$  TeV*, CMS PAS HIG-18-019 [<https://cds.cern.ch/record/2649199>] (2018).
- [8] CMS Collaboration, *Measurement of the associated production of a Higgs boson and a pair of top-antitop quarks with the Higgs boson decaying to two photons in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, CMS PAS HIG-18-018 [<https://cds.cern.ch/record/2649208>] (2018).
- [9] CMS Collaboration, *Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Phys. Rev. D 99, 092005 (2019), doi:10.1103/PhysRevD.99.092005.