CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2017/03/20

Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for the production for a standard model Higgs boson in association with a top quark pair ($t\bar{t}H$) is presented. The analysis is based on pp collision data collected by the CMS experiment at a center of mass energy of $\sqrt{s} = 13$ TeV in 2016, and corresponding to an integrated luminosity of 35.9 fb^{-1}. We target final states where the Higgs boson decays to WW^*, ZZ^* or $\tau\tau$, and at least one of the top quarks decays leptonically, by selecting events with two leptons of the same charge, three leptons or at least four leptons, and b-jets. We obtain a best fit $t\bar{t}H$ yield of 1.5 ± 0.5 times the standard model prediction, with an observed (expected) significance of 3.3σ (2.5σ), by the combination of these results with those obtained from the 2015 dataset.

1 Introduction

After the discovery of the Higgs boson by the ATLAS and CMS collaborations [1, 2], its mass has been accurately measured and its couplings to most standard model (SM) particles constrained [3–6]. So far, these results are all in agreement with the SM expectation.

Studies of various final states have been used to constrain the top Yukawa coupling to about 30% accuracy [7–10]. This is a particularly relevant parameter because of the large mass of the top quark. It can be probed either indirectly in studies of Higgs production via gluon fusion, where virtual top quarks provide the dominant SM contribution to the loop amplitude, or directly at tree level in $\mathrm{t\bar{t}H}$ associated production. The combined study of both production channels represents a stringent test of the SM.

The large pp collision dataset being collected during the LHC Run II represents an opportunity to increase the precision in this and other Higgs measured properties, and to study its production and decay in more channels. In particular, the $\mathrm{t\bar{t}H}$ production cross section is enhanced by a factor of about 4 with respect to the LHC Run I center of mass energy. Moreover, any deviation from SM expectations could be an indication of new physics.

This note describes a search for $\mathrm{t\bar{t}H}$ production in leptonic final states with the full dataset collected by the CMS experiment at a center of mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. It represents an extension of a previous result obtained with the first part of this dataset [11]. We target Higgs decays to WW^*, ZZ^* or $\tau\tau$ in events where at least one of the top quarks decays leptonically, by selecting events with two electrons or muons of the same charge, or with more than three electrons or muons, and hadronic jets compatible with the hadronization of b quarks.

The $\mathrm{t\bar{t}H}$ yield is extracted by a simultaneous fit of two multivariate discriminants, aimed at separating the signal from the irreducible background of $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ production, and from the reducible background of events where at least one of the lepton candidates originates from an hadronic jet (mainly from $\mathrm{t\bar{t}}$ +jets), respectively.

With respect to the previous analysis [11], the relative contribution from reducible background processes is reduced by tighter lepton selection requirements. We also complement the kinematic discriminant inputs with dedicated variables aimed at reconstructing hadronic top decays and tagging jets likely to originate from the Higgs decay products. Moreover, we employ an improved method for optimizing the binning of the discriminant outputs and add a four-lepton event category.

2 Object reconstruction and identification

Events are reconstructed using a particle-flow (PF) algorithm that optimally combines the information from all sub-detectors to identify individual particles and provide a global interpretation of the event [12, 13]. PF candidates are classified into charged hadrons, neutral hadrons, photons, muons, and electrons.

Hadronic jets are reconstructed by clustering PF candidates in the anti-k_T algorithm with a distance parameter of 0.4, as implemented in the FASTJET package [14, 15]. Charged hadrons that are not consistent with the selected primary interaction vertex are discarded from the clustering. The jet energy is then corrected for the varying response of the detector as a function of transverse momentum (p_T) and pseudorapidity (η) [16]. Jets are selected for use in the analysis only if they have $p_\mathrm{T} > 25\,\mathrm{GeV}$, $|\eta| < 2.4$, and are separated from any selected lep-

tons by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$. To select jets that are likely to have originated from the hadronization of a b quark, they are analyzed in a multivariate likelihood discriminant using track-based lifetime information and reconstructed secondary vertices ("combined secondary vertex" (CSV) algorithm) [17]. The efficiency to correctly tag b jets and the probability to misidentify jets from light quarks or gluons are measured in data as a function of the jet p_T and η, and are used to correct for differences in the performance of the algorithm in simulated events. Two working points based on the algorithm output are used: "loose", with a b signal tagging efficiency of about 85% and a mistagging rate of about 10%; and "medium", with b efficiency of about 70% and mistagging rate of order 1.5% [18]. Tagging efficiencies for jets from charm quarks are about 40% for the loose, and about 20% for the medium working point.

In each event, the missing transverse momentum (E_T^{miss}) is calculated as the magnitude of the negative vector sum of the transverse momenta of all reconstructed PF candidates. The H_T^{miss} variable is defined as the magnitude of the negative vector sum of the transverse momenta of all selected leptons and jets in the event, i.e. it neglects unclustered energy deposits in the detector, making it robust against the pp collision remnants (underlying events) and additional pp interactions in a bunch crossing (pileup).

Muon candidates are reconstructed by combining information from the silicon tracker and the outer muon spectrometer of CMS in a global fit [19]. The quality of the geometrical matching between the individual measurements in the tracker and the muon system is used to discriminate genuine prompt muons from hadrons punching through the calorimetry and from muons produced in in-flight decays of kaons and pions. In the analysis, muon candidates are considered if they have $p_T > 5\,\text{GeV}$ and $|\eta| < 2.4$. In the same-sign dilepton event categories, the relative uncertainty in the muon p_T from the fit is required to be better than 20% to ensure a high quality charge measurement.

Electrons are reconstructed using information from the tracker and from the electromagnetic calorimeter [20]. Genuine electrons are identified by a multivariate algorithm using the shape of the calorimetric shower and the quality of the reconstructed track. Furthermore, to reject electrons produced in photon conversions, candidates with missing hits in the innermost tracking layers or matched to a conversion secondary vertex are discarded. The residual yield of events with an electron originating from a photon conversion is estimated from simulation. Electrons are selected for the analysis if they have $p_T > 7\,\text{GeV}$ and $|\eta| < 2.5$. To suppress electrons with a mis-assigned electric charge in the same-sign dilepton categories, candidates are required to have consistent charge measurements from three independent observables based on the calorimetry energy deposits and the track curvature.

Finally, hadronically decaying τ leptons (τ_h) are reconstructed using the "hadron-plus-strip" algorithm [21]. We consider candidates with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.3$. They are also required to pass the "decay mode finding" discriminator and to be reconstructed either in the 1- or 3-prong decay modes, with or without additional neutral pions.

Electrons and muons passing the criteria described above are referred to as "loose leptons" in the following. A further discrimination between prompt signal leptons (i.e. from W and Z boson decays and from leptonic τ decays) and non-prompt and spurious leptons from b hadron decays, decays-in-flight, and photon conversions is crucial in light of the overwhelming background from $t\bar{t}$ production. To maximally exploit the available information in each event to that end, a multivariate discriminator based on a boosted decision tree (BDT) [22] algorithm is built, taking as input not just observables related directly to the reconstructed leptons themselves, but also to the clustered energy deposits and charged particles in a cone around the lepton direction. The jet reconstruction and b-tagging algorithms are run on these, and their

output is used to train the algorithm. In particular, the ratio between the lepton p_T and the reconstructed jet p_T, and the transverse momentum of the lepton with respect to the jet axis complement more traditional observables like the relative isolation of the lepton (calculated in a variable cone size depending on the lepton p_T [23, 24]), and the impact parameters of the lepton trajectory. The BDT algorithm is trained on prompt leptons in simulated $t\bar{t}H$ signal and non-prompt leptons in $t\bar{t}$ background events and validated using data in various control regions. Leptons are then selected for the final analysis if they pass a given threshold of the BDT output, referred to as "tight selection" in the following.

3 Event selection

Events are selected using triggers based on the presence of one, two or three reconstructed electrons or muons. The minimum p_T thresholds in single lepton triggers are 24 GeV for muons, 27 GeV for electrons. For double lepton triggers, the p_T thresholds on the leading and subleading leptons are 17 and 8 GeV for muons, 23 and 12 GeV for electrons. Three lepton triggers record events where the third hardest lepton p_T exceeds 5 GeV for muons, 9 GeV for electrons.

Events are categorized as a function of the lepton multiplicity, targeting different $t\bar{t}H$ leptonic final states. In case of a $H \to WW^*$ decay, for instance, two pairs of opposite-sign W bosons and two b quarks from top quark decays are produced (Fig. 1, left). We target events where at least two W bosons decay to leptons. In a similar way, Higgs decays to ZZ^* or $\tau\tau$ can yield leptons in the final state, always produced in association with b-jets (Fig. 1, right).

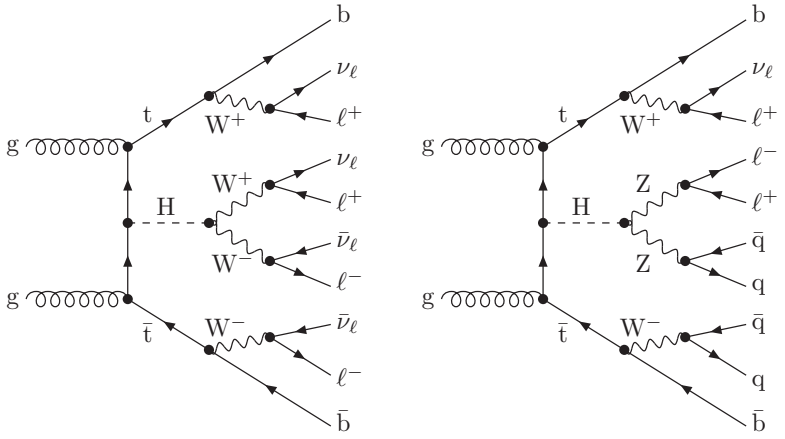

Figure 1: Possible Feynman diagrams for $t\bar{t}H$ production at the LHC, where the Higgs boson decays to WW^* (left) or ZZ^* (right). Subsequent W and Z decays are shown by representing examples of leptonic final states.

We require at least two of the reconstructed hadronic jets to pass the loose working point of the CSV discriminator, or at least one of them to pass the medium working point. Reconstructed τ decays to hadrons are vetoed in all categories, since those events are studied by a different analysis. Events where a pair of leptons passing the loose selection has an invariant mass of less than 12 GeV are rejected, as they are not well modeled by the simulation. Events where at least one of the selected leptons satisfies requirements aimed at tagging products of photon conversions are equally removed.

Same-sign dilepton channel Events with exactly two leptons of the same charge passing the tight selection criteria fall into the same-sign dilepton (2LSS) category of the analysis. We

require the two leptons to have a p_T greater than 25 GeV and 15 GeV respectively, and the final state to contain at least four hadronic jets.

Events where at least one of the selected leptons fails to satisfy requirements aimed at ensuring a reliable electric charge measurement are discarded. Even so, a significant contribution from charge-mismeasured $Z \to ee$ decays remains. Therefore, ee events where the invariant dielectron mass falls within 10 GeV of the Z boson mass are also rejected.

In order to further suppress the contribution from leptonic Z decays, we require $0.6 \times E_T^{miss} + 0.4 \times H_T^{miss}$ to be larger than 30 GeV. This has the effect of rejecting roughly a factor of two more Drell-Yan events than a simple $E_T^{miss} > 25$ GeV requirement, without a significant loss of signal efficiency.

Three lepton channel The three lepton (3L) channel of the analysis consists of events with exactly three leptons passing the tight selection criteria, and at least two hadronic jets. The lepton transverse momenta are required to be greater than 25 GeV, 15 GeV and 15 GeV respectively.

Events with a dilepton pair having an invariant mass within 10 GeV of the Z boson mass are rejected, irrespectively of the lepton flavor. We also apply the $0.6 \times E_T^{miss} + 0.4 \times H_T^{miss} > 30$ GeV requirement if the final state contains less than four hadronic jets. The threshold is tightened to 45 GeV if the event contains a pair of opposite-sign leptons of the same flavor.

Four lepton channel This channel is defined exactly as the 3L one, apart from requiring the presence of at least one additional lepton with a p_T greater than 10 GeV.

Table 1 presents the yields for signal and background processes in each channel, together with the observed ones in data. The main backgrounds are $t\bar{t}W$ and $t\bar{t}Z$ production, whose final states are similar to $t\bar{t}H$ in number of prompt leptons and b quarks, and $t\bar{t}$ events that contribute due to non-prompt leptons being selected. The "Rare SM bkg." contribution includes four top, tWZ, tZq and triboson production processes, as well as same-sign W pair production from double parton scattering. MC yields are shown after a fit to data where all processes are constrained to their SM prediction within the experimental and theoretical uncertainties, described in Section 5. The background from non-prompt leptons and charge mis-measurements is estimated from the data, as described in Section 5.4.

4 Signal extraction

Event categorization Before the signal extraction, events in the 2LSS and 3L channels are further categorized. In 2LSS, we split the events in lepton flavor channels: ee, $\mu\mu$, or eμ. With the exception of the ee channel, events are then further split according to whether they contain two jets passing the medium working point of the CSV b-tagging discriminator (events that do are labeled as "b-tight", the remaining ones are labeled as "b-loose"). The same categorization in b-tight and b-loose is done for 3L events, without splitting in lepton flavors. Finally, all resulting 2LSS and 3L categories are split according to the sum of lepton charges, to exploit the charge asymmetry in $t\bar{t}W$ production and other background processes.

Kinematic discrimination Two BDT classifiers are used to improve the separation between the signal and each of such background classes in the 2LSS and 3L channels. In the 4L channel no kinematic discrimination is performed, because of its limited statistical power. The classifiers are trained on simulated events using the TMVA package [25], separately for the two channels, and use the following topological and kinematic variables as input distributions:

Table 1: Yields for expected signal and background processes, and observed yields in data, for the 2LSS (top), 3L and 4l (bottom) channels. The predictions for the non-prompt lepton and charge mis-measurement contributions are extracted from data. Yields are shown after a fit to data, with all processes constrained to the SM expectation.

	$\mu\mu$	eμ	ee
$t\bar{t}W$	51.0 ± 0.6 (stat.) ± 6.9 (syst.)	72.8 ± 0.7 (stat.) ± 10.2 (syst.)	20.5 ± 0.4 (stat.) ± 3.1 (syst.)
$t\bar{t}Z/\gamma^*$	17.7 ± 0.8 (stat.) ± 2.9 (syst.)	47.3 ± 1.6 (stat.) ± 9.0 (syst.)	17.5 ± 1.0 (stat.) ± 3.6 (syst.)
WZ	4.2 ± 0.6 (stat.) ± 4.1 (syst.)	7.0 ± 0.8 (stat.) ± 6.8 (syst.)	1.8 ± 0.4 (stat.) ± 1.7 (syst.)
Rare SM bkg.	4.2 ± 1.5 (stat.) ± 3.0 (syst.)	13.3 ± 1.9 (stat.) ± 9.3 (syst.)	4.8 ± 1.1 (stat.) ± 3.6 (syst.)
WWss	3.5 ± 0.6 (stat.) ± 2.5 (syst.)	4.1 ± 0.6 (stat.) ± 3.2 (syst.)	1.4 ± 0.3 (stat.) ± 1.2 (syst.)
Conversions		7.8 ± 2.5 (stat.) ± 2.3 (syst.)	3.6 ± 3.5 (stat.) ± 1.7 (syst.)
Charge mis-meas.		16.4 ± 0.2 (stat.) ± 9.1 (syst.)	10.5 ± 0.2 (stat.) ± 5.9 (syst.)
Non-prompt leptons	38.7 ± 1.6 (stat.) ± 20.5 (syst.)	61.8 ± 2.0 (stat.) ± 13.0 (syst.)	17.7 ± 1.1 (stat.) ± 5.4 (syst.)
All backgrounds	120.3 ± 2.5 (stat.) ± 11.7 (syst.)	231.2 ± 4.3 (stat.) ± 13.3 (syst.)	77.9 ± 4.0 (stat.) ± 9.0 (syst.)
$t\bar{t}H$ signal	20.1 ± 0.5 (stat.) ± 2.1 (syst.)	27.9 ± 0.5 (stat.) ± 3.0 (syst.)	8.0 ± 0.3 (stat.) ± 1.1 (syst.)
Data	150	268	89

	3L	4L
$t\bar{t}W$	32.8 ± 1.0 (stat.) ± 4.9 (syst.)	
$t\bar{t}Z/\gamma^*$	49.8 ± 3.9 (stat.) ± 11.1 (syst.)	2.15 ± 0.24 (stat.) ± 0.44 (syst.)
WZ	9.1 ± 0.9 (stat.) ± 4.0 (syst.)	
Rare SM bkg.	8.8 ± 4.3 (stat.) ± 5.9 (syst.)	0.27 ± 0.16 (stat.) ± 0.19 (syst.)
WWss		
Conversions	5.3 ± 1.2 (stat.) ± 4.0 (syst.)	
Charge mis-meas.		
Non-prompt leptons	30.8 ± 1.5 (stat.) ± 10.9 (syst.)	
All backgrounds	137.3 ± 6.2 (stat.) ± 12.4 (syst.)	2.42 ± 0.28 (stat.) ± 0.56 (syst.)
$t\bar{t}H$ signal	19.5 ± 1.0 (stat.) ± 3.0 (syst.)	1.00 ± 0.09 (stat.) ± 0.11 (syst.)
Data	148	3

- the maximum $|\eta|$ of the two leading leptons;
- the jet multiplicity;
- the minimum ΔR separation between each of the two leading leptons and a jet;
- the transverse mass of the leading lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$;
- the maximum score among jet permutations of a BDT discriminator that aims at reconstructing hadronic top decays (used in the training against $t\bar{t}$ in the 2LSS channel only);
- the maximum score among jet permutations of a BDT discriminator that aims at tagging jets from Higgs decay products (used in the training against $t\bar{t}W/t\bar{t}Z$ in the 2LSS channel only);
- the highest and lowest p_{T} of the selected leptons (used in the trainings against $t\bar{t}W/t\bar{t}Z$ only);
- matrix element weights for signal and irreducible backgrounds, combined in one likelihood ratio variable (used in the 3L category only).

The hadronic top decay tagger aims at an event reconstruction by matching each selected jet and lepton to a final state particle in 2LSS $t\bar{t}H$ events, and testing the compatibility of jets with the hadronic decay of a top quark. Hypotheses with up to two non-reconstructed jets are tested, in order to keep efficiency for cases of only partial event reconstruction. The training is performed against a background consisting of all incorrectly matched jet and lepton permutations in $t\bar{t}H$ events. The following variables are used in the training: the mass and p_{T} of the reconstructed hadronic top; the mass of the W boson and the CSV discriminator of the b-jet originating from the hadronic top decay; the p_{T} ratio of the lepton from the Higgs to the lepton from the top, and the ΔR separation between leptons and b-jets from the top quark pair system.

The Hj discriminator is designed to identify jets originating from the Higgs decay products. The classifier is trained against a background of jets in $t\bar{t}W/t\bar{t}Z$ events in the 2LSS category, and uses jet identification (CSV discriminator, quark-gluon jet likelihood) and geometric (ΔR with respect to the leptons) properties as input variables. It is not evaluated on jets compatible with top decay products according to the previous discriminator.

In the 3L event category only, the kinematic variables listed above are complemented by matrix element weights. A weight $w_{i,\alpha}$ is computed for each hypothesis α (where α is either $t\bar{t}H$, $t\bar{t}W$, or $t\bar{t}Z$) and for the event i as follows:

$$w_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha} \int d\Phi_\alpha \cdot \delta^4\left(p_1^\mu + p_2^\mu - \sum_{k\geq 2} p_k^\mu\right) \cdot \frac{f(x_1,\mu_F) f(x_2,\mu_F)}{x_1 x_2 s} \cdot \left|\mathcal{M}_\alpha(p_k^\mu)\right|^2 \cdot W(\Phi'|\Phi_\alpha),$$

where σ_α is the cross section; Φ' are the 4-momenta of the reconstructed particles; $d\Phi_\alpha$ is the element of phase space corresponding to unmeasured quantities with momentum conservation enforced; $f(x,\mu_F)$ are the parton density functions, computed using NNPDF3.0 LO [26]; $|\mathcal{M}_\alpha|^2$ is the squared matrix element, computed with MADGRAPH 5_AMC@NLO standalone [27] at LO in the narrow-width approximation for t, $\bar{t}$ and H; and W are the transfer functions for jet energy and E_T^{miss}, relating parton to reconstructed quantities, estimated from simulated $t\bar{t}H$ events.

The two jets with the highest CSV tagging output are assigned to the two b quarks in the matrix element. Among the remaining jets, the pair with dijet mass closest to m_W is selected. In $t\bar{t}H$, for semileptonic decays of the Higgs daughters, the pair with lowest dijet mass is selected. If one or two jets needed to evaluate $|\mathcal{M}_\alpha|^2$ fail to be reconstructed, the weight is recovered by extending the integration phase space for the missing jets.

The final weight for each hypothesis α is taken as the average of the weights computed for each lepton and jet permutation. The MEM weights of signal and backgrounds are combined in a likelihood ratio that is used as an input variable to the BDT. Including the MEM weights in the BDT training against $t\bar{t}W/t\bar{t}Z$ improves the background rejection power by about 10% for the three lepton category.

The plane spanned by the outputs of the two BDT classifiers is binned using a method based on the likelihood ratio between signal and background. Starting from a fine binning allowed by considerations on the signal and background statistical uncertainties in each bin, the joint likelihood is approximated by the signal-to-background ratio in each bin, and then smoothed using gaussian kernels. Each background event is associated to the value of the likelihood ratio in the bin the event belongs to; the cumulative distribution of the likelihood ratio for background events is then partitioned, based on its quantiles, in a certain number of regions of equal background content. The number of regions is chosen using a recursive application of the k-means clustering algorithm [28]. The resulting regions are finally interpreted as bins of a one-dimensional distribution, which features in a natural way a roughly constant number of background events and an increasing number of signal events.

The distributions obtained in this way for each category are simulaneously fit to extract the signal normalization. Figures 2, 3, 4, and 5 show distributions of event observables and BDT classifier outputs in data, compared to the predicted background processes.

5 Signal and background modeling and systematic uncertainties

Signal $t\bar{t}H$ events are generated using the MADGRAPH 5_AMC@NLO package (version 5.222) [27], which includes up to one additional hadronic jet at next-to-leading order (NLO) QCD accuracy.

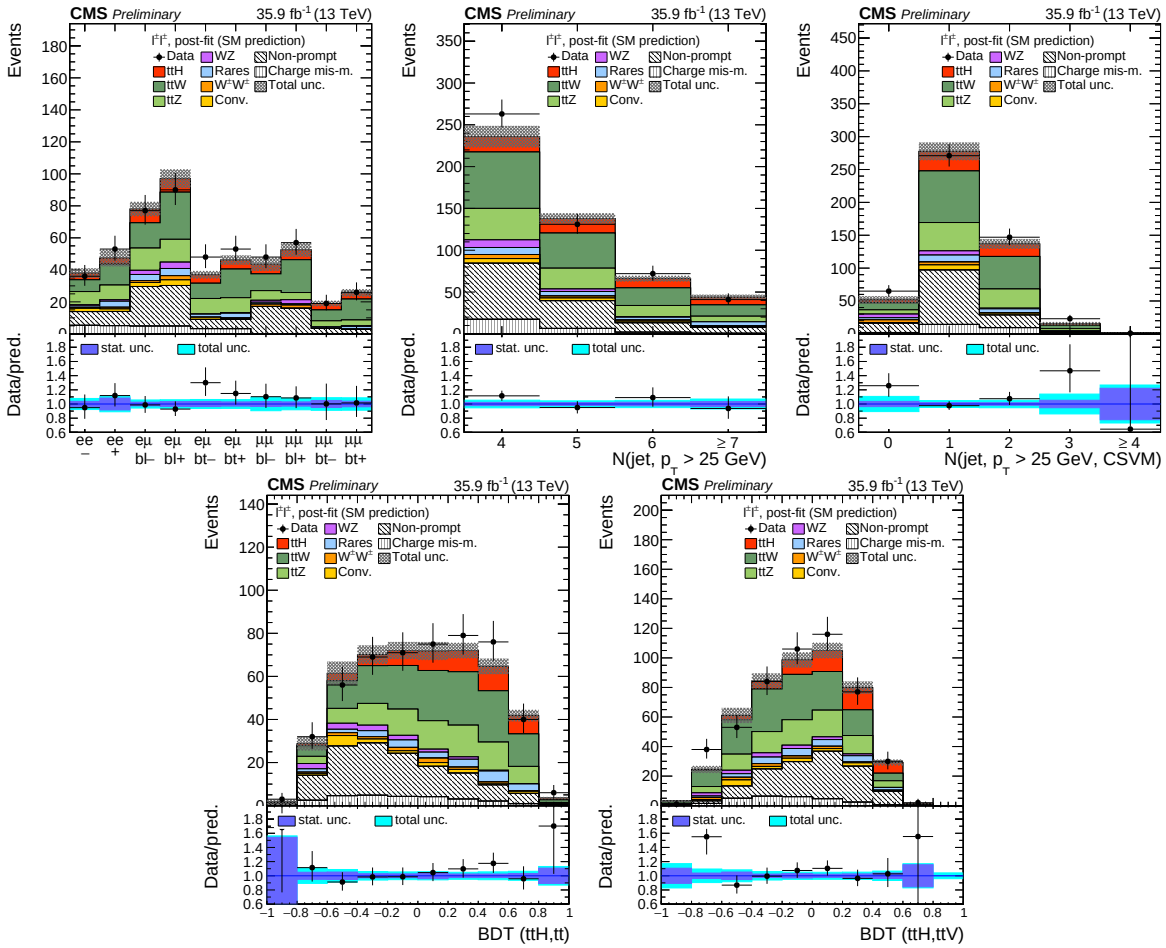

Figure 2: Same-sign dilepton channel; top row: distribution in the categories used for the fit, number of jets, number of jets passing the medium working point of the CSV tagger; bottom row: distributions of the BDT classifier outputs. The distributions are shown after the fit to the data, with all processes constrained to the SM expectation.

The same generator is used for the $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}\gamma$+jets processes. The diboson (VV, V=W,Z) processes are generated with POWHEG [29–34]. Other minor backgrounds are simulated using POWHEG or MADGRAPH as generators at leading order QCD accuracy. All generated events are interfaced to PYTHIA8 (v8.205) [35] for the parton shower and hadronization steps. Pileup interactions are simulated to reflect the observed multiplicity in data. All events are finally passed through a full simulation of the CMS detector based on GEANT4 [36], and reconstructed using the same algorithms as used for the data.

5.1 Theoretical uncertainties

The NLO calculation for the inclusive production cross section of the $t\bar{t}H$ signal has uncertainties from unknown higher orders in the perturbative series of $+5.8/-9.2\%$ and 3.6% from the uncertainty in the parton distribution functions (PDFs) and in α_S [37]. These are propagated to the final normalization of the signal yields. Uncertainties in the shape of the final classifier output are estimated by varying the renormalization and factorization scales in generated events.

$t\bar{t}W$ and $t\bar{t}Z$ production cross sections are calculated at NLO QCD and EWK, with theoretical uncertainties from unknown higher orders of about 12% and 10%, respectively [37]. As for

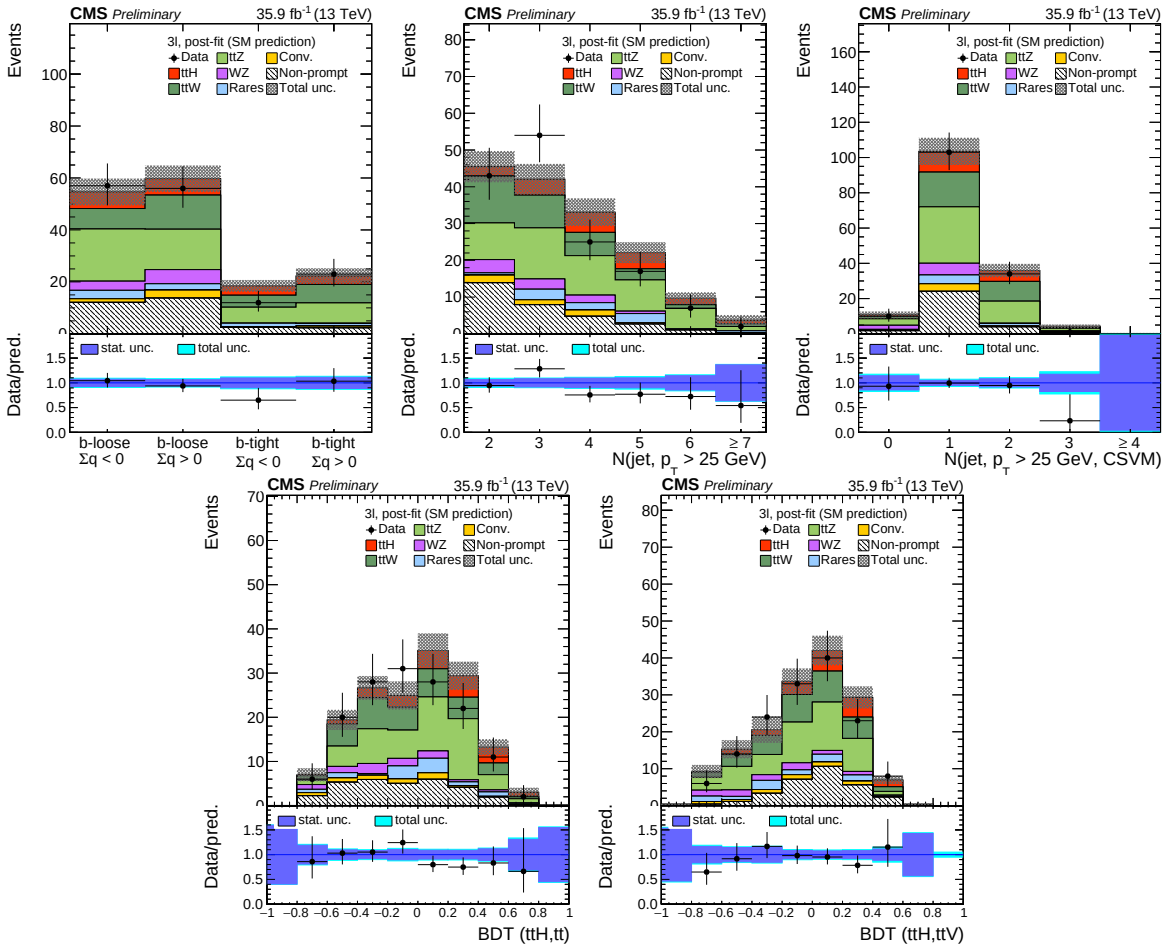

Figure 3: Three lepton channel; top row: distribution in the categories used for the fit, number of jets, number of jets passing the medium working point of the CSV tagger; bottom row: distributions of the BDT classifier outputs. The distributions are shown after the fit to the data, with all processes constrained to the SM expectation.

the signal, a further uncertainty arises from the knowledge of PDFs and α_S of 2 to 4%, and an uncertainty in the shape of the classifier outputs is estimated by varying renormalization and factorization scales, leading to a variation of the classifier output shape of about 2 to 4% in amplitude.

5.2 Efficiency scale factors

Scale factors are applied to correct for differences between data and simulation in the trigger and lepton selection efficiency. Trigger efficiency scale factors are compatible with unity and are measured with an uncertainty ranging from 1% to 3%. Lepton efficiency scale factors are measured separately for electrons and muons in $Z \to \ell^+\ell^-$ events as a function of the lepton p_T, with a typical uncertainty of 4%.

The impact of the uncertainty in the signal selection efficiency from jet energy corrections is evaluated by varying the correction factors within their uncertainty and propagating the effect to the final result by recalculating all kinematic quantities. Both effects on the overall normalization of event yields and on the shape of kinematic properties are taken into account. Jet energy resolution uncertainties have a negligible impact on the analysis. Correction factors for data/MC differences in the b-tagging performance are applied depending on the p_T, η and fla-

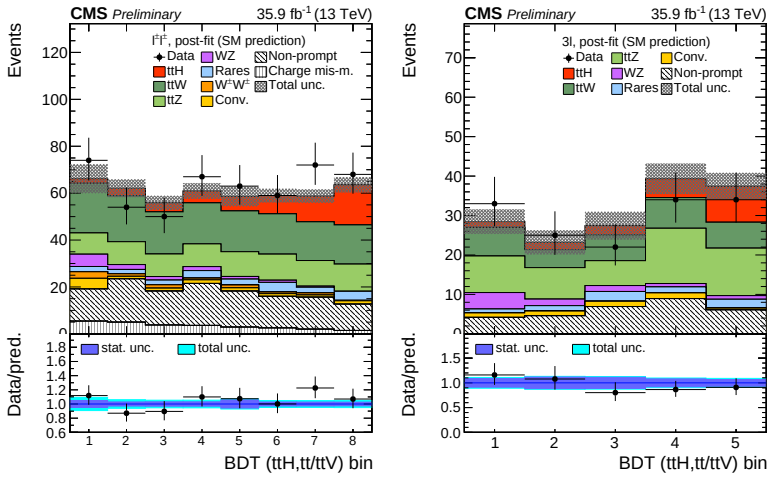

Figure 4: Combination of the BDT classifier outputs in the bins used for signal extraction, for the same-sign dilepton (left) and three-lepton (right) channels. The distributions are shown after the fit to the data, with all processes constrained to the SM expectation.

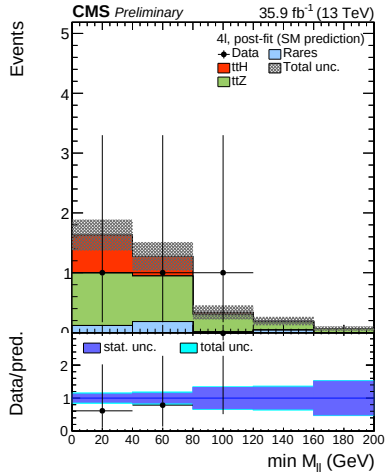

Figure 5: Minimum mass of opposite-sign (OS) lepton pairs in the 4L category, without any selection on lepton flavor (all-flavor, AF). The distribution is shown after the fit to the data, with all processes constrained to the SM expectation.

vor of the jet, and their effect on the signal efficiency is evaluated by varying the factors within their measured uncertainty and recalculating the overall event scale factors.

5.3 WZ and ZZ backgrounds

Diboson production with leptonic Z decays and additional jet radiation in the final state can lead to signatures very similar to that of the $t\bar{t}H$ signal. Due to the larger cross section, the main contribution arises from WZ production. Inclusive production cross sections for both WZ and ZZ have been measured at the LHC and agree well with the NLO calculations. However, the good agreement in the inclusive phase space of the cross section measurements does not necessarily hold in the signal region of this analysis which requires the presence of hadronic jets, including b jets. Therefore, a dedicated control region dominated by WZ production is used to constrain the overall normalization of this process. It is defined by the presence of at least three leptons, of which one opposite-sign pair is to be compatible with a Z boson decay. Furthermore, at least two jets are required, with a veto on jets that pass the loose b tag selection to ensure exclusivity with the signal selection. A scale factor is then extracted from the predicted

distribution of WZ events in the control region, and the observed data, keeping other processes fixed. The measured scale factor is compatible with unity.

The majority (about 70%) of diboson events passing the signal selection contains jets from light quarks and gluons that are wrongly tagged as b jets, making this estimate mainly sensitive to the experimental uncertainty in the mis-tag rate rather than the theoretical uncertainty in the jet flavor composition. The overall uncertainty assigned to the diboson prediction is estimated from the statistical uncertainty due to the limited sample size in the control region (30%), the residual background in the control region (20%), the uncertainties on the b-tagging rate (between 10–40%), and from the knowledge of PDFs and the theoretical uncertainties of the extrapolation (up to 10%).

5.4 Non-prompt and charge mis-identified leptons

Events from processes with comparatively large cross sections, in which one of the leptons is produced inside a jet (i.e. it is non-prompt), give a sizeable contribution even after all selection requirements are applied. They mostly contain real leptons from b hadron decay, but also hadronic jets misidentified as leptons. The yield of such events is estimated from a loose-to-tight extrapolation, in which a looser lepton selection is defined and the rate at which such leptons enter the tighter selection is measured in a control region and then used to extrapolate from a sideband with looser leptons to the signal selection with tight leptons. The definition of the looser lepton selection is based on the one described in Section 2, complemented by further requirements on isolation and b-tagging observables.

The probability for a non-prompt lepton candidate passing the loose selection to also pass the tight requirements is measured in a sample dominated by QCD multijet events, collected using single lepton triggers. At low muon p_{T}, a jet with p_{T} greater than 40 GeV is also required in the trigger. Contributions from prompt leptons, mainly from W and Z +jet or from WZ and ZZ events, respectively, are first suppressed by vetoing additional leptons in the selection, and the residual contamination is then subtracted using the transverse mass as a discriminating variable. The tight-to-loose ratio is then extracted as a function of p_{T} and $|\eta|$ and separately for muons and electrons.

A sideband control region is defined by relaxing the lepton selection criteria to the loose set, without modifying any other requirement. By weighting events in this expanded selection with a factor dependent on the measured tight-to-loose ratios, a fully data-driven estimation for the contribution of non-prompt leptons to the signal selection is obtained. In events where just one of the two leptons fails the tight criteria, the applied event weight is $f/(1-f)$ (where f is the tight-to-loose ratio measured as described above), while events where both leptons fail the tight criteria are weighted by $-f_1 f_2/[(1-f_1)(1-f_2)]$. The resulting prediction of the event yield in the signal selection carries a typical uncertainty of 20–40%, arising from the statistical uncertainty in the measurement of the tight-to-loose ratios, and from a systematical uncertainty derived by comparing alternative methods of subtracting prompt lepton backgrounds and from testing the closure of the method in simulated background events.

Similarly, the background from events where the charge of one of the leptons is wrongly measured, relevant only for the 2LSS channel, is determined by measuring the charge mis-assignment probability in a sample of same-sign dilepton events compatible with a Z boson decay, and weighting events with opposite-sign leptons in the signal selection accordingly. The charge mis-assignment probability is found to be negligible for this analysis for muons, whereas for electrons it ranges from about 0.02% in the barrel section up to about 0.4% in the detector endcaps. It is measured as a function of the electron p_{T} and $|\eta|$. A systematic uncer-

tainty of 30% is assigned to the prediction from the statistical uncertainty of the measurement and closure tests on simulated events.

6 Results

The fitted signal yields are compared with the expectation for a SM Higgs boson of 125 GeV. We introduce a signal strength parameter $\mu = \sigma/\sigma_{\mathrm{SM}}$, acting as a pure scaling of the $\mathrm{t\bar{t}H}$ yields.

The observed (expected) best fit signal strength amounts to $1.5^{+0.5}_{-0.5}$ ($1.0^{+0.5}_{-0.4}$), with an observed (expected) significance of 3.3σ (2.4σ). The contributions of statistical, theoretical and experimental uncertainties to the total uncertainty on μ amount to 0.29, 0.24 and 0.32 respectively. The observed (expected) 95% CL exclusion limit [38–41] on μ, in the context of the background-only hypothesis, is 2.5 (0.8).

Category	Observed limit	Expected limit $\pm1\sigma$
Same-sign di-lepton	2.8	$0.9\,(-0.3)\,(+0.4)$
Three lepton	2.5	$1.4\,(-0.4)\,(+0.7)$
Four lepton	5.9	$4.9\,(-1.7)\,(+3.1)$
Combined	2.5	$0.8\,(-0.2)\,(+0.3)$

Table 2: Asymptotic 95% CL upper limits on μ under the background-only hypothesis.

Category	Observed μ fit $\pm1\sigma$	Expected μ fit $\pm1\sigma$
Same-sign di-lepton	$1.7\,(-0.5)\,(+0.6)$	$1.0\,(-0.5)\,(+0.5)$
Three lepton	$1.0\,(-0.7)\,(+0.8)$	$1.0\,(-0.7)\,(+0.8)$
Four lepton	$0.9\,(-1.6)\,(+2.3)$	$1.0\,(-1.6)\,(+2.4)$
Combined (2016 data)	$1.5\,(-0.5)\,(+0.5)$	$1.0\,(-0.4)\,(+0.5)$
Combined (2015 data) [42]	$0.6\,(-1.1)\,(+1.4)$	$1.0\,(-1.1)\,(+1.3)$
Combined (2015+2016 data)	$1.5\,(-0.5)\,(+0.5)$	$1.0\,(-0.4)\,(+0.5)$

Table 3: Best fit of the signal strength parameter.

We perform a cross-check fit where cross section modifiers are introduced for the $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ processes, and left free to float together with μ. In order to improve the constraining power for these parameters, we also include in this fit two control regions: 2LSS + 3 jets: obtained from the nominal 2LSS + 4 jets selection by requiring exactly 3 jets in the final state, enriched in $\mathrm{t\bar{t}W}$ and reducible backgrounds; 3L on-Z: obtained from the nominal 3L selection by inverting the Z mass veto requirement, enriched in $\mathrm{t\bar{t}Z}$. The sensitivity obtained on μ is only slightly worse than with the nominal fit. The observed (expected) best fit μ in the cross-check fit is $1.3^{+0.5}_{-0.5}$ ($1.0^{+0.5}_{-0.4}$), with an observed (expected) significance of 3.0σ (2.4σ). The fitted $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ cross section modifiers are compatible with unity within 1σ.

Finally, these results are combined with those obtained using data collected in 2015 [42]. The combined fit yields an observed (expected) best fit signal strength of $1.5^{+0.5}_{-0.5}$ ($1.0^{+0.5}_{-0.4}$) times the SM expectation, corresponding to an observed (expected) significance of 3.3σ (2.5σ).

Results, including fits separately performed in the different analysis channels, are summarised in Tab. 2 and Tab. 3.

7 Conclusions

A search for the production of a SM Higgs boson in association with a top anti-top quark pair has been presented. The analysis is based on pp collision data collected by the CMS experiment

in 2016, and corresponding to an integrated luminosity of 35.9 fb^{-1}. We target events where the Higgs boson decays to WW^*, ZZ^* or $\tau\tau$ by requiring with two same-sign leptons or at least three leptons in the final state. Results are combined with those obtained from the 2015 dataset. The observed (expected) best fit $\text{t}\bar{\text{t}}\text{H}$ yield is $1.5^{+0.5}_{-0.5}$ ($1.0^{+0.5}_{-0.4}$) times the SM prediction, corresponding to a significance of 3.3σ (2.5σ).

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] ATLAS and CMS Collaborations, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.

[4] ATLAS Collaboration, "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment", *Eur. Phys. J.* **C76** (2016), no. 1, 6, doi:10.1140/epjc/s10052-015-3769-y, arXiv:1507.04548.

[5] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J.* **C75** (2015), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[6] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV", ATLAS-CONF-2015-044, CMS-PAS-HIG-15-002, 2015.

[7] CMS Collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106, arXiv:1408.1682. [Erratum: JHEP10,106(2014)].

[8] ATLAS Collaboration, "Search for H $\to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector", *Phys. Lett.* **B740** (2015) 222–242, doi:10.1016/j.physletb.2014.11.049, arXiv:1409.3122.

[9] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Eur. Phys. J.* **C75** (2015), no. 7, 349, doi:10.1140/epjc/s10052-015-3543-1, arXiv:1503.05066.

[10] ATLAS Collaboration, "Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector", *Phys. Lett.* **B749** (2015) 519–541, doi:10.1016/j.physletb.2015.07.079, arXiv:1506.05988.

[11] CMS Collaboration, "Search for associated production of Higgs bosons and top quarks in multilepton final states at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-022, 2016.

[12] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[13] CMS Collaboration, "Commissioning of the Particle–Flow reconstruction in Minimum–Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[14] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[15] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett.* **B641** (2006) 57–61, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[16] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[17] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[18] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[19] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$ = 7 TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[20] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[21] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS", *JINST* **11** (2016), no. 01, P01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

[22] B. P. Roe et al., "Boosted decision trees as an alternative to artificial neural networks for particle identification", *Nucl. Instrum. Meth. A* **543** (2005) 577, doi:10.1016/j.nima.2004.12.018, arXiv:0408124.

[23] K. Rehermann and B. Tweedie, "Efficient Identification of Boosted Semileptonic Top Quarks at the LHC", *JHEP* **03** (2011) 059, doi:10.1007/JHEP03(2011)059, arXiv:1007.2221.

[24] CMS Collaboration, "Search for new physics in same-sign dilepton events in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J.* **C76** (2016), no. 8, 439, doi:10.1140/epjc/s10052-016-4261-z, arXiv:1605.03171.

[25] A. Hoecker et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS ACAT* **040** (2007) arXiv:physics/0703039.

[26] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[27] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[28] J. MacQueen, "Some methods for classification and analysis of multivariate observations", in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, pp. 281–297. University of California Press, Berkeley, Calif., 1967.

[29] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[30] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[31] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[32] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J.* **C71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[33] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076.

[34] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX", *JHEP* **11** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.

[35] T. Sjstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[36] J. Allison et al., "Geant4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

[37] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.

[38] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys.* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.

[39] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth.* **A434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[40] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727. [Erratum: Eur. Phys. J.C73,2501(2013)].

[41] ATLAS and CMS Collaborations, LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in summer 2011", ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, 2011.

[42] CMS Collaboration, "Search for $t\bar{t}H$ production in multilepton final states at $\sqrt{s} = 13\,\text{TeV}$", CMS Physics Analysis Summary CMS-PAS-HIG-15-008, 2016.