Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method

CMS Collaboration*
CERN, 1211 Geneva 23, Switzerland

Received: 9 February 2015 / Accepted: 11 May 2015
© CERN for the benefit of the CMS collaboration 2015. This article is published with open access at Springerlink.com

Abstract A search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks is presented. Events with hadronic jets and one or two oppositely charged leptons are selected from a data sample corresponding to an integrated luminosity of 19.5 fb$^{-1}$ collected by the CMS experiment at the LHC in pp collisions at a centre-of-mass energy of 8 TeV. In order to separate the signal from the larger $t\bar{t}$ + jets background, this analysis uses a matrix element method that assigns a probability density value to each reconstructed event under signal or background hypotheses. The ratio between the two values is used in a maximum likelihood fit to extract the signal yield. The results are presented in terms of the measured signal strength modifier, $\mu$, relative to the standard model prediction for a Higgs boson mass of 125 GeV. The observed (expected) exclusion limit at a 95% confidence level is $\mu < 4.2$ ($3.3$), corresponding to a best fit value $\hat{\mu} = 1.2^{+1.6}_{-1.5}$.

1 Introduction

Following the discovery of a new boson with mass around 125 GeV by the ATLAS and CMS Collaborations [1–3] at the CERN LHC, the measurement of its properties has become an important task in particle physics. The precise determination of its quantum numbers and couplings to gauge bosons and fermions will answer the question whether the newly discovered particle is the Higgs boson (H) predicted by the standard model (SM) of particle physics, i.e. the quantum of the field responsible for the spontaneous breaking of the electroweak symmetry [4–9]. Conversely, any deviation from SM predictions will represent evidence of physics beyond our present knowledge, thus opening new horizons in high-energy physics. While the measurements performed with the data collected so far indicate overall consistency with the SM expectations [3,10–13], it is necessary to continue improving on the measurement of all possible observables.

In the SM, the Higgs boson couples to fermions via Yukawa interactions with strength proportional to the fermion mass. Direct measurements of decays into bottom quarks and $\tau$ leptons have provided the first evidence that the 125 GeV Higgs boson couples to down-type fermions with SM-like strength [14]. Evidence of a direct coupling to up-type fermions, in particular to top quarks, is still lacking. Indirect constraints on the top-quark Yukawa coupling can be inferred from measuring either the production or the decay of Higgs bosons through effective couplings generated by top-quark loops. Current measurements of the Higgs boson cross section via gluon fusion and of its branching fraction to photons are consistent with the SM expectation for the top-quark Yukawa coupling [3,10–12].

The first search for $t\bar{t}H$ in association with a top-quark pair ($t\bar{t}H$) provides an example of such an observable. A sample of tree-level Feynman diagrams contributing to the partonic processes $q\bar{q}$, $gg \to t\bar{t}H$ is shown in Fig. 1 (left and centre). The inclusive next-to-leading-order (NLO) $t\bar{t}H$ cross section is about 130 fb in pp collisions at a centre-of-mass energy $\sqrt{s} = 8$ TeV for a Higgs boson mass ($m_H$) of 125 GeV [15–24], which is approximately two orders of magnitude smaller than the cross section for Higgs boson production via gluon fusion [23,24].

The first search for $t\bar{t}H$ events used $p\bar{p}$ collision data at $\sqrt{s} = 1.96$ TeV collected by the CDF experiment at the Tevatron collider [25]. Searches for $t\bar{t}H$ production at the LHC have previously been published for individual decay modes of the Higgs boson [26,27]. The first combination of
tH searches in different final states has been published by the CMS Collaboration based on the full data set collected at √s = 7 and 8 TeV [28]. Assuming SM branching fractions, the results of that analysis set a 95% confidence level (CL) upper limit on the tH signal strength at 4.5 times the SM value, while an upper limit of 1.7 times the SM is expected from the background-only hypothesis. The median expected exclusion limit for tH production in the H → b¯b channel alone is 3.5 in the absence of a signal.

The results of a search for tH production in the decay channel H → b¯b are presented in this paper based on pp collision data at √s = 8 TeV collected with the CMS detector [29] and corresponding to an integrated luminosity of 19.5 fb−1. The analysis described here differs from that of Ref. [28] in the way events are categorized and in its use of an analytical matrix element method (MEM) [30,31] for improving the separation of signal from background. Within the MEM technique, each reconstructed event is assigned a probability density value based on the theoretical differential cross section \( \sigma^{-1}d\sigma/dy \), where \( y \) denotes the four-momenta of the reconstructed particles. Particle-level quantities that are either unknown (e.g. neutrino momenta, jet-parton associations) or poorly measured (e.g. quark energies) are marginalised by integration. The ratio between the probability density values for signal and background provides a discriminating variable suitable for testing the compatibility of an event with either of the two hypotheses [32].

The MEM has already been successfully used at the Tevatron collider in the context of Higgs boson searches [33,34], although for simpler final states. A phenomenological feasibility study for a tH measurement in the H → b¯b decay channel at the LHC using the MEM has been pioneered in Ref. [35] based on the MadWeight package [36] for automatised matrix-element calculations. The present paper makes use of an independent implementation of the MEM, specifically optimized for the final state of interest. This is the first time that the MEM is applied to a search for tH events. The final states typical of tH events with H → b¯b, that are characterised by huge combinatorial background, the presence of nonreconstructed particles, and small signal-to-background ratios, provide an ideal case for the deployment of the MEM. The analysis strategy is designed to maximise the separation between tH and t + b¯b background events,

in order to reduce the systematic uncertainty on the signal extraction related to the modelling of this challenging background.

This paper is organised as follows. Section 2 describes the main features of the CMS detector. Section 3 presents the data and simulation samples, while Sects. 4 and 5 discuss the reconstruction of physics objects and the event selection, respectively. Section 6 describes the signal extraction. The treatment of systematic uncertainties and the statistical interpretation of the results are discussed in Sects. 7 and 8, respectively. Section 9 summarises the results.

2 CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the field volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors. The first level of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a time interval of less than 4 \( \mu \)s. The high-level trigger processor farm further decreases the event rate from around 100 kHz to around 1 kHz, before data storage. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables can be found in Ref. [29].

3 Data and simulated samples

The data sample used in this search was collected with the CMS detector in 2012 from pp collisions at a centre-of-mass energy of 8 TeV, using single-electron, single-muon, or dielectron triggers. The single-electron trigger requires the presence of an isolated electron with transverse momentum...
(\(p_T\)) in excess of 27 GeV. The single-muon trigger requires an isolated muon candidate with \(p_T\) above 24 GeV. The dielectron trigger requires two isolated electrons with \(p_T\) thresholds of 17 and 8 GeV.

Signal and background processes are modelled with Monte Carlo (MC) simulation programs. The CMS detector response is simulated by using the Geant4 software package [37]. Simulated events are required to pass the same trigger selection and offline reconstruction algorithms used on collision data. Correction factors are applied to the simulated samples to account for residual differences in the selection and reconstruction efficiencies with respect to those measured.

The \(t\bar{t}H, H \rightarrow b\bar{b}\) signal is modelled by using the Pythia 6.426 [38] leading order (LO) event generator normalised to the NLO theoretical cross section [15–24], and assuming the SM Higgs boson with a mass of 125 GeV. The main background in the analysis stems from \(t\bar{t}\) jets production. This process has been simulated with the MadGraph 5.1.3 [39] tree-level matrix element generator matched to Pythia for the parton shower description, and normalised to the inclusive next-to-next-to-leading-order (NNLO) cross section with soft-gluon resummation at next-to-next-to-leading logarithmic accuracy [40]. The \(t\bar{t} +\) jets sample has been generated in a five-flavour scheme with tree-level diagrams for two top quarks plus up to three extra partons, including both charm and bottom quarks. An additional correction factor to the \(t\bar{t} +\) jets samples is applied to account for the differences observed in the top-quark \(p_T\) spectrum when comparing the MadGraph simulation with data [41]. The interference between the \(t\bar{t}H, H \rightarrow b\bar{b}\) diagrams and the \(t\bar{t} + b\bar{b}\) background diagrams is negligible and is not considered in the MC simulation. Minor backgrounds come from the Drell–Yan production of an electroweak boson with additional jets (W + jets, Z + jets), and from the production of a top–quark pair in association with a W\(^\pm\), Z boson (tW, tZ). These processes have been generated by MadGraph matched to the Pythia parton shower description. The Drell–Yan processes have been normalised to the NNLO inclusive cross section from FewZ 3.1 [42], while the NLO calculations from Refs. [43,44] are used to normalise the tW and tZ samples, respectively. Single top quark production is modelled with the NLO generator POWHEG 1.0 [45–50] combined with Pythia. Electroweak diboson processes (WW, WZ, and ZZ) are simulated by using the Pythia generator normalised to the NLO cross section calculated with MCFM 6.6 [51]. Processes that involve top quarks have been generated with a top-quark mass of 172.5 GeV. Samples generated at LO use the CTEQ6L1 parton distribution function (PDF) set [52], while samples generated with NLO programs use the CTEQ6.6M PDF set [53].

Effects from additional pp interactions in the same bunch crossing (pileup) are modelled by adding simulated minimum bias events (generated with PYTHIA) to the generated hard interactions. The pileup multiplicity in the MC simulation is reweighted to reflect the luminosity profile observed in pp collision data.

### 4 Event reconstruction

The global event reconstruction provided by the particle-flow (PF) algorithm [54,55] seeds the reconstruction of the physics objects deployed in the analysis. To minimise the impact of pileup, charged particles are required to originate from the primary vertex, which is identified as the reconstructed vertex with the largest value of \(\sum p_{T,i}^2\), where \(p_{T,i}\) is the transverse momentum of the \(i\)th charged particle associated with the vertex. The missing transverse momentum vector \(\vec{p}_T^{miss}\) is defined as the negative vector sum of the transverse momenta of all neutral particles and of the charged particles coming from the primary vertex. Its magnitude is referred to as \(E_T^{miss}\).

Muons are reconstructed from a combination of measurements in the silicon tracker and in the muon system [56]. Electron reconstruction requires the matching of an energy cluster in the ECAL with a track in the silicon tracker [57]. Additional identification criteria are applied to muon and electron candidates to reduce instrumental backgrounds. An isolation variable is defined starting from the scalar \(p_T\) sum of all particles contained inside a cone around the track direction, excluding the contribution from the lepton itself. The sum of neutral pileup energy is estimated as the average \(p_T\) density calculated from all neutral particles in the event multiplied by an effective area of the isolation cone, and is subtracted from the total sum.

Jets are reconstructed by using the anti-\(k_T\) clustering algorithm [58], as implemented in the FASTJET package [59,60], with a distance parameter of 0.5. Each jet is required to have pseudorapidity (\(\eta\)) in the range \([-2.5, 2.5]\), to have at least two tracks associated with it, and to have electromagnetic and hadronic energy fractions of at least 1 % of the total jet energy. Jet momentum is determined as the vector sum of the momenta of all particles in the jet. An offset correction is applied to take into account the extra energy clustered in jets because of pileup. Jet energy corrections are derived from the simulation, and are confirmed with in situ measurements of the energy balance of dijet and Z/\(\gamma\) + jet events [61]. Additional selection criteria are applied to each event to remove spurious jet-like features originating from isolated noise patterns in few HCAL regions.

The combined secondary vertex (CSV) b-tagging algorithm is used to identify jets originating from the hadronisation of bottom quarks [62]. This algorithm combines the information about track impact parameters and secondary vertices within jets into a likelihood discriminant to provide
separation of b-quark jets from jets that originate from lighter quarks or gluons. The CSV algorithm assigns to each jet a continuous value that can be used as a jet flavour discriminator. Large values of the discriminator correspond preferentially to b-quark jets, so that working points of increasing purity can be defined by requiring higher values of the CSV discriminator. For example, the CSV medium working point (CSVM) is defined in such a way as to provide an efficiency of about 70% (20%) to tag jets originating from a bottom (charm) quark, and of approximately 2% for jets originating from light quarks or gluons. Scale factors are applied to the simulation to match the distribution of the CSV discriminator from light quarks or gluons. Scale factors are applied to the (charm) quark, and of approximately 2% for jets originating of about 70% (20%) to tag jets originating from a bottom

The scale factors have been derived as a function of measured with a tag-and-probe technique [63] in data control simulation to match the distribution of the CSV discriminator from light quarks or gluons. Scale factors are applied to the

5 Event selection

The experimental signature of tH events with H → bĐ is affected by a large multijet background which can be reduced to a negligible level by only considering the semileptonic decays of the top quark. The selection criteria are therefore optimised to accept events compatible with a tH signal where H → bĐ and at least one of the top quarks decays to a bottom quark, a charged lepton, and a neutrino. Events are divided into two exclusive channels depending on the number of charged leptons (electrons or muons), which can be either one or two. Top quark decays in final states with tau leptons are not directly searched for, although they can still satisfy the event selection criteria when the tau lepton decays to an electron or muon, plus neutrinos. Channels of different lepton multiplicities are analysed separately. The single-lepton (SL) channel requires one isolated muon with \( p_T > 30 \text{ GeV} \) and |\( \eta \) | < 2.1, or one isolated electron with \( p_T > 30 \text{ GeV} \) and |\( \eta \) | < 2.5, excluding the 1.44 < |\( \eta \) | < 1.57 transition region between the ECAL barrel and endcap. Events are vetoed if additional electrons or muons with \( p_T \) in excess of 20 GeV, the same |\( \eta \) | requirement, and passing some looser identification and isolation criteria are found. The dilepton (DL) channel collects events with a pair of oppositely charged leptons satisfying the selection criteria used to veto additional leptons in the SL channel. To reduce the contribution from Drell–Yan events in the same-flavour DL channel, the invariant mass of the lepton pair is required to be larger than 15 GeV and at least 8 GeV away from the Z boson mass. Figure 2 (top) shows the jet multiplicity in the SL (left) and DL (right) channels, while the bottom left panel of the same figure shows the multiplicity of jets passing the CSVM working point in the SL channel.

The optimisation of the selection criteria in terms of signal-to-background ratio requires a stringent demand on the number of jets. At least five (four) jets with \( p_T > 30 \text{ GeV} \) and |\( \eta \) | < 2.5 are requested in the SL (DL) channel. A further event selection is required to reduce the \( t\bar{t} + \text{jets} \) background, which at this stage exceeds the signal rate by more than three orders of magnitude. For this purpose, the CSV discriminator values are calculated for all jets in the event and collectively denoted by \( \xi \). For SL (DL) events with seven or more (five or more) jets, only the six (four) jets with the largest CSV discriminator value are considered. The likelihood to observe \( \xi \) is then evaluated under the alternative hypotheses of \( t\bar{t} + \text{two heavy-flavour jets} \) or \( t\bar{t} + \text{two light-flavour jets} \). For example, for SL events with six jets, and neglecting correlations among different jets in the same event, the likelihood under the \( t\bar{t} + \text{hf} \) hypothesis is estimated as:

\[
f(\xi|t\bar{t}+\text{hf}) = \frac{f(\xi|t\bar{t}+\text{hf})}{f(\xi|t\bar{t}+\text{hf}) + f(\xi|t\bar{t}+\text{lf})}.
\]

(1)

where \( \xi_i \) is the CSV discriminator for the \( i \)th jet, and \( f_{\text{hf}}(\xi) \) is the probability density function (pdf) of \( \xi_i \) when the \( i \)th jet originates from heavy- (light-)flavour partons. The latter include u, d, s quarks and gluons, but not c quarks. For the sake of simplicity, the likelihood in Eq. (1) is rigorous for \( W \to u\bar{d}(s) \) decays, whereas it is only approximate for \( W \to c\bar{s}(d) \) decays, since the CSV discriminator pdf for charm quarks differs with respect to \( f_{\text{lf}} \) [62]. Equation (1) can be extended to the case of SL events with five jets, or DL events with at least four jets, by considering that in both cases four of the jets are associated with heavy-flavour partons, and the remaining jets with light-flavour partons. The likelihood under the alternative hypothesis, \( f(\xi|t\bar{t}+\text{lf}) \), is given by Eq. (1) after swapping \( f_{\text{hf}} \) for \( f_{\text{lf}} \). The variable used to select events is then defined as the likelihood ratio

\[
F(\xi) = \frac{f(\xi|t\bar{t}+\text{hf})}{f(\xi|t\bar{t}+\text{hf}) + f(\xi|t\bar{t}+\text{lf})}.
\]

(2)

The distribution of \( F \) for SL events with six jets is shown in Fig. 2 (bottom right).

In the following, events are retained if \( F \) is larger than a threshold value \( F_L \) ranging between 0.85 and 0.97, depending on the channel and jet multiplicity. The selected events are further classified as high-purity (low-purity) if \( F \) is larger (smaller) than a value \( F_H \), with \( F_L < F_H < 1.0 \). The low-purity categories serve as control regions for \( t\bar{t} + \text{lf} \) jets, providing constraints on several sources of systematic uncertainty. The high-purity categories are enriched in \( t\bar{t} + \text{hf} \) events, and drive the sensitivity of the analysis. The thresholds \( F_L \) and \( F_H \) are optimised separately for each of the analysis categories defined in Sect. 6. The exact values are reported in Table 1.
Fig. 2 Top row distribution of the jet multiplicity in (left) single-lepton and (right) dilepton events, after requiring that at least two jets pass the CSVM working point. Bottom-left distribution of the multiplicity of jets passing the CSVM working point in single-lepton events with at least four jets. Bottom-right distribution of the selection variable $F$ defined in Eq. (2) for single-lepton events with at least six jets after requiring a loose preselection of at least one jet passing the CSVM working point. The plots at the bottom of each panel show the ratio between the observed data and the background expectation predicted by the simulation. The shaded and solid green bands correspond to the total statistical plus systematic uncertainty in the background expectation described in Sect. 7. More details on the background modelling are provided in Sect. 6.3.

After requiring a lower threshold on the selection variable $F$, the background is dominated by $t\bar{t} +$ jets, with minor contributions from the production of a single top quark plus jets, $t\bar{t}$ plus vector bosons, and $W/Z +$ jets; the expected purity for a SM Higgs boson signal is only at the percent level. By construction, the selection criteria based on Eq. (2) enhance the $t\bar{t} + b\bar{b}$ subprocess compared to the otherwise dominant $t\bar{t} + l\bar{f}$ production. The $t\bar{t} + b\bar{b}$ background has the same final state as the signal whenever the two b quarks are resolved as individual jets. Therefore, this background cannot be effectively reduced by means of the $F$ discriminant. The cross section for $t\bar{t} + b\bar{b}$ production with two resolved b-quark jets is larger than that of the signal by about one order of magnitude and is affected by sizable theoretical uncertainties [64], which
Table 1 Expected and observed event yields in the (top) high-purity (H) and (bottom) low-purity (L) categories of the SL and DL channels

<table>
<thead>
<tr>
<th>b-Tagging selection</th>
<th>SL Cat-1 (H)</th>
<th>SL Cat-2 (H)</th>
<th>SL Cat-3 (H)</th>
<th>DL (H)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.995 ≤ F ≤ 1.000</td>
<td>0.993 ≤ F ≤ 1.000</td>
<td>0.995 ≤ F ≤ 1.000</td>
<td>0.925 ≤ F ≤ 1.000</td>
</tr>
<tr>
<td>( t\overline{t}H )</td>
<td>3.5 ± 0.7</td>
<td>5.9 ± 1.1</td>
<td>7.5 ± 1.4</td>
<td>4.5 ± 0.7</td>
</tr>
<tr>
<td>( t\overline{t} + l\overline{f} )</td>
<td>17 ± 3</td>
<td>70 ± 13</td>
<td>152 ± 21</td>
<td>84 ± 11</td>
</tr>
<tr>
<td>( t\overline{t} + \ell \ell )</td>
<td>22 ± 8</td>
<td>66 ± 20</td>
<td>81 ± 24</td>
<td>85 ± 24</td>
</tr>
<tr>
<td>( t\overline{t} + b\ell )</td>
<td>16 ± 8</td>
<td>44 ± 23</td>
<td>70 ± 32</td>
<td>47 ± 23</td>
</tr>
<tr>
<td>( t\overline{t} + b\overline{b} )</td>
<td>43 ± 11</td>
<td>75 ± 17</td>
<td>69 ± 18</td>
<td>50 ± 13</td>
</tr>
<tr>
<td>( t\overline{t} + W/Z )</td>
<td>3.2 ± 0.8</td>
<td>4.4 ± 1.1</td>
<td>4.2 ± 1.0</td>
<td>5.1 ± 1.1</td>
</tr>
<tr>
<td>Single t</td>
<td>3.1 ± 1.4</td>
<td>5.3 ± 2.2</td>
<td>14 ± 4</td>
<td>5.9 ± 1.7</td>
</tr>
<tr>
<td>( W/Z + \text{jets} )</td>
<td>...</td>
<td>0.3 ± 2.2</td>
<td>...</td>
<td>6 ± 5</td>
</tr>
<tr>
<td>Total background</td>
<td>103 ± 11</td>
<td>265 ± 24</td>
<td>390 ± 28</td>
<td>283 ± 24</td>
</tr>
<tr>
<td>Data</td>
<td>107</td>
<td>272</td>
<td>401</td>
<td>279</td>
</tr>
<tr>
<td>( S/B (S/\sqrt{B}) )</td>
<td>3.4% (0.34)</td>
<td>2.2% (0.36)</td>
<td>1.9% (0.38)</td>
<td>1.6% (0.27)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b-Tagging selection</th>
<th>SL Cat-1 (L)</th>
<th>SL Cat-2 (L)</th>
<th>SL Cat-3 (L)</th>
<th>DL (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.960 ≤ F &lt; 0.995</td>
<td>0.960 ≤ F &lt; 0.995</td>
<td>0.970 ≤ F &lt; 0.995</td>
<td>0.85 ≤ F &lt; 0.925</td>
</tr>
<tr>
<td>( t\overline{t}H )</td>
<td>3.8 ± 0.7</td>
<td>5.2 ± 0.8</td>
<td>7.9 ± 1.3</td>
<td>0.8 ± 0.1</td>
</tr>
<tr>
<td>( t\overline{t} + l\overline{f} )</td>
<td>111 ± 13</td>
<td>268 ± 32</td>
<td>737 ± 62</td>
<td>69 ± 8</td>
</tr>
<tr>
<td>( t\overline{t} + \ell \ell )</td>
<td>94 ± 25</td>
<td>161 ± 45</td>
<td>268 ± 74</td>
<td>40 ± 11</td>
</tr>
<tr>
<td>( t\overline{t} + b\ell )</td>
<td>45 ± 24</td>
<td>80 ± 42</td>
<td>162 ± 77</td>
<td>14 ± 7</td>
</tr>
<tr>
<td>( t\overline{t} + b\overline{b} )</td>
<td>48 ± 13</td>
<td>69 ± 17</td>
<td>84 ± 22</td>
<td>7.6 ± 2.1</td>
</tr>
<tr>
<td>( t\overline{t} + W/Z )</td>
<td>4.0 ± 1.0</td>
<td>6.9 ± 1.5</td>
<td>7.8 ± 1.6</td>
<td>2.3 ± 0.5</td>
</tr>
<tr>
<td>Single t</td>
<td>8.9 ± 2.4</td>
<td>13 ± 3</td>
<td>32 ± 6</td>
<td>3.1 ± 1.1</td>
</tr>
<tr>
<td>( W/Z + \text{jets} )</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>5 ± 3</td>
</tr>
<tr>
<td>Total background</td>
<td>311 ± 22</td>
<td>598 ± 38</td>
<td>1291 ± 60</td>
<td>142 ± 10</td>
</tr>
<tr>
<td>Data</td>
<td>310</td>
<td>603</td>
<td>1310</td>
<td>140</td>
</tr>
<tr>
<td>( S/B (S/\sqrt{B}) )</td>
<td>1.2% (0.21)</td>
<td>0.9% (0.21)</td>
<td>0.6% (0.22)</td>
<td>0.5% (0.07)</td>
</tr>
</tbody>
</table>

The expected event yields with their uncertainties are obtained from a signal-plus-background fit as described in Sect. 8. In the last row of each table, the symbol S (B) denotes the signal (total background) yield.

Hampers the possibility of extracting the signal via a counting experiment. A more refined approach, which thoroughly uses the kinematic properties of the reconstructed event, is therefore required to improve the separation between the signal and the background.

6 Signal extraction

As in other resonance searches, the invariant mass reconstructed from the \( H \to b\overline{b} \) decay provides a natural discriminating variable to separate the narrow Higgs boson dijet resonance from the continuum mass spectrum expected from the \( t\overline{t} + \text{jets} \) background. However, in the presence of additional \( b \) quarks from the decay of the top quarks, an ambiguity in the Higgs boson reconstruction is introduced, leading to a combinatorial background. The distribution of the experimental mass estimator built from a randomly selected jet pair is much broader compared to the detector resolution, since wrongly chosen jet pairs are only mildly or not at all correlated with \( m_H \). Unless a selection rule is introduced to filter out the wrong combinations, the existence of such a combinatorial background results in a suppression of the statistical power of the mass estimator, which grows as the factorial of the jet multiplicity. Multivariate techniques that exploit the correlation between several observables in the same event are naturally suited to deal with signal extraction in such complex final states.

In this paper, a likelihood technique based on the theoretical matrix elements for the \( t\overline{t}H \) process and the \( t\overline{t} + b\overline{b} \) background is applied for signal extraction. This method utilises the kinematics and dynamics of the event, providing a powerful discriminant between the signal and background. The \( t\overline{t} + b\overline{b} \) matrix elements are considered as the prototype to model all background processes. This choice guarantees optimal separation between the signal and the \( t\overline{t} + b\overline{b} \) background, which is a desirable property given the large rate and theoretical uncertainty in the latter. The performance on the
other $t\bar{t} + \text{jets}$ subprocesses might not be necessarily optimal, even though some separation power is still preserved; indeed, the $t\bar{t} + b\bar{b}$ matrix elements describe these processes better than the signal matrix elements do, as it has been verified a posteriori with the simulation. More specifically, the shapes of the matrix element discriminant predicted by the simulation for the different $t\bar{t} + \text{jets}$ subprocesses are found to be similar to each other, with a slightly better separation power for the $t\bar{t} + b\bar{b}$ background. The approximate degeneracy in shape between several processes can be ascribed to a smearing effect of the combinatorial background, as well as to the impact of the Higgs boson mass constraint on the calculation of the event likelihood under the signal hypothesis. The latter provides a similar discrimination against all $t\bar{t} + \text{jets}$ subprocesses. A slightly worse separation power is instead observed for minor backgrounds, such as single top events, for which neither of the two matrix elements tested really applies. However, all of the background processes analysed are found to yield discriminant shapes that can be well distinguished from that for the signal. Also, it is found that most of the statistical power attained by this method in separating $t\bar{t}H$, $H \rightarrow b\bar{b}$ from $t\bar{t} + b\bar{b}$ events relies on the different correlation and kinematic distributions of the two $b$-quark jets not associated with the top quark decays.

6.1 Construction of the MEM probability density functions

The MEM probability density functions under the signal and background hypothesis are constructed at LO assuming for simplicity that in both cases the reactions proceed via gluon fusion. At $\sqrt{s} = 8$ TeV, the fraction of the gluon-gluon initiated subprocesses is about 55% (65%) of the inclusive LO (NLO) cross section, and it grows with the centre-of-mass energy [21]. Examples of diagrams entering the calculation are shown in the middle and right panels of Fig. 1. All possible jet-quark associations in the reconstruction of the final state are considered. For each event, the MEM probability density function $w(y|H)$ under the hypothesis $H = t\bar{t}H$ or $t\bar{t} + b\bar{b}$ is calculated as:

$$w(y|H) = \frac{N_o}{\prod_{i=1}^{8} p_k} \int \frac{d\rho_T}{(2\pi)^2} \prod_{i=1}^{8} \left( \frac{d^3p_k}{2E_k} \right) \times (2\pi)^4 g(E,c) \left( p_a + p_b - \sum_{k=1}^{8} p_k \right) \times R^{(x,y)}(\rho_T, \sum_{k=1}^{8} p_k) \times g(x_a, \mu_F) g(x_b, \mu_F) \times |\mathcal{M}_H(p_a, p_b, p_1, \ldots, p_8)|^2 W(y, p),$$

where $y$ denotes the set of observables for which the matrix element pdf is constructed, i.e. the momenta of jets and leptons. The sum extends over the $N_o$ possibilities of associating the jets with the final-state quarks. The integration on the right-hand side of Eq. (3) is performed over the phase space of the final-state particles and over the gluon energy fractions $x_{a,b}$ by using the VEGAS [65] algorithm. The four-momenta of the initial-state gluons $p_{a,b}$ are related to the four-momenta of the colliding protons $P_{a,b}$ by the relation $p_{a,b} = x_{a,b} P_{a,b}$. The delta function enforces the conservation of longitudinal momentum and energy between the incoming gluons and the $k = 1, \ldots, 8$ outgoing particles with four-momenta $p_k$. To account for the possibility of initial/final state radiation, the total transverse momentum of the final-state particles, which should be identically zero at LO, is instead loosely constrained by the resolution function $R^{(x,y)}$ to the measured transverse recoil $\rho_T$, defined as the negative of the total transverse momentum of jets and leptons, plus the missing transverse momentum.

The remaining part of the integrand in Eq. (3) contains the product of the gluon PDFs in the protons ($g$), the square of the scattering amplitude ($\mathcal{M}$), and the transfer function ($W$). For $H = t\bar{t}H$, the factorisation scale $\mu_F$ entering the PDF is taken as half of the sum of twice the top-quark mass and the Higgs boson mass [20], while for $H = t\bar{t} + b\bar{b}$ a dynamic scale is used equal to the quadratic sum of the transverse masses for all coloured partons [66]. The scattering amplitude for the hard process is evaluated numerically at LO accuracy by the program OPENLOOPS [67]; all resonances are treated in the narrow-width approximation [68], and spin correlations are neglected. The transfer function $W(y, p)$ provides a mapping between the measured set of observables $y$ and the final-state particles momenta $p = (p_1, \ldots, p_8)$. Given the good angular resolution of jets, the direction of quarks is assumed to be perfectly measured by the direction of the associated jets. Also, since energies of leptons are measured more precisely than for jets, their momenta are considered perfectly measured. Under these assumptions, the total transfer function reduces to the product of the quark energy transfer function times the probability for the quarks that are not reconstructed as jets to fail the acceptance criteria. The quark energy transfer function is modelled by a single Gaussian function for jets associated with light-flavour partons, and by a double Gaussian function for jets associated with bottom quarks; the latter are constructed by superimposing two Gaussian functions with different mean and standard deviation. Such an asymmetric parametrisation provides a good description of both the core of the detector energy response and the low-energy tail arising from semileptonic B hadron decays. The parametrisation of the transfer functions has been derived from MC simulated samples.
6.2 Event categorisation

To aid the evaluation of the MEM probability density functions at LO, events are classified into mutually exclusive categories based on different parton-level interpretations. Firstly, the set of jets yielding the largest contribution to the sum defined by Eq. (1), determines the four (tagged) jets associated with bottom quarks; the remaining \( N_{\text{untag}} \) (untagged) jets are assumed to originate either from \( W \rightarrow q\bar{q}' \) decays (SL channel) or from initial- or final-state gluon radiation (SL and DL channels). There still remains a twelve-fold ambiguity in the determination of the parton matched to each jet, which is reflected by the sum in Eq. (3). Indeed, without distinguishing between \( b \) and \( \bar{b} \) quarks, there exist \( 4!/(2!2!) = 6 \) combinations for assigning two jets out of four with the Higgs boson decay (\( \mathcal{H} = \text{ttH} \), or with the bottom quark-pair radiation (\( \mathcal{H} = \text{tt} + \text{bb} \); for each of these possibilities, there are two more ways of assigning the remaining tagged jets to either the \( t \) or \( \bar{t} \) quark, thus giving a total of twelve associations. In the SL channel, an event can be classified in one of three possible categories. The first category (Cat-1) is defined by requiring at least six jets; if there are exactly six jets, the mass of the two untagged jets is required to be in the range \([60, 100]\) GeV, i.e. compatible with the mass of the \( W \) boson. If the number of jets is larger than six, the mass range is tightened to compensate for the increased ambiguity in selecting the correct \( W \) boson decay products. In the event interpretation, the \( W \rightarrow q\bar{q}' \) decay is assumed to be fully reconstructed, with the two quarks identified with the jet pair satisfying the mass constraint. The definition of the second category (Cat-2) differs from that of Cat-1 by the inversion of the dijet mass constraint. This time, the event interpretation assumes that one of the quarks from the \( W \) boson decay has failed the reconstruction. The integration on the right-hand side of Eq. (3) is extended to include the phase space of the nonreconstructed quark. The other untagged jet(s) is (are) interpreted as gluon radiation, and do not enter the calculation of \( w(y|\mathcal{H}) \). The total number of associations considered is twelve times the multiplicity of untagged jets eligible to originate from the \( W \) boson decay: \( N_a = 12N_{\text{untag}} \). In the third category (Cat-3), exactly five jets are required, and an incomplete \( W \) boson reconstruction is again assumed. In the DL channel, only one event interpretation is considered, namely that each of the four bottom quarks in the decay is associated with one of the four tagged jets.

Finally, two event discriminants, denoted by \( P_{s/b} \) and \( P_{h/l} \), are defined. The former encodes only information from the event kinematics and dynamics via Eq. (3), and is therefore suited to separate the signal from the background; the latter contains only information related to \( b \) tagging, thus providing a handle to distinguish between the heavy- and the light-flavour components of the \( \text{tt} + \text{jets background} \). They are defined as follows:

\[
P_{s/b} = \frac{w(y|\text{ttH})}{w(y|\text{ttH}) + k_{s/b} w(y|\text{tt} + \text{bb})}
\]

and

\[
P_{h/l} = \frac{f(\xi|\text{tt} + \text{hf})}{f(\xi|\text{tt} + \text{hf}) + k_{h/l} f(\xi|\text{tt} + \text{lf})},
\]

where the functions \( f(\xi|\text{tt} + \text{hf}) \) and \( f(\xi|\text{tt} + \text{lf}) \) are defined as in Eq. (1) but restricting the sum only to the jet-quark associations considered in the calculation of \( w(y) \); the coefficients \( k_{s/b} \) and \( k_{h/l} \) in the denominators are positive constants that can differ among the categories and will be treated as optimisation parameters, as described below.

The joint distribution of the \( (P_{s/b}, P_{h/l}) \) discriminants is used in a two-dimensional maximum likelihood fit to search for events resulting from Higgs boson production. By construction, the two discriminants satisfy the constraint \( 0 \leq P_{s/b}, P_{h/l} \leq 1 \). Because of the limited size of the simulated samples, the distributions of \( P_{s/b} \) and \( P_{h/l} \) are binned. A finer binning is used for the former, which carries the largest sensitivity to the signal, while the latter is divided into two equal-sized bins. The coefficient \( k_{s/b} \) appearing in the definition of \( P_{s/b} \) is introduced to adjust the relative normalisation between \( w(y|\text{ttH}) \) and \( w(y|\text{tt} + \text{bb}) \); likewise for \( k_{h/l} \). A redefinition of any of the two coefficients would change the corresponding discriminant monotonically, thus with no impact on its separation power. However, since both variables are analysed in bins with fixed size, an optimisation procedure, based on minimising the expected exclusion limit on the signal strength as described in Sect. 8, is carried out to choose the values that maximise the sensitivity of the analysis. More specifically, the coefficients \( k_{s/b} \) are first set to the values that remove any local maximum for the \( \text{tt} + \text{bb} \) distribution around \( P_{s/b} \sim 1 \), a condition that is found to provide already close to optimal coefficients. Then, starting from this initial point, several values of \( k_{s/b} \) are scanned and the \( P_{s/b} \) distributions are recomputed accordingly. An expected upper limit on the signal strength is then evaluated for each choice of \( k_{s/b} \) using the simulated samples. This procedure is repeated until a minimum in the expected limit is obtained. A similar procedure is applied for choosing the optimal \( k_{h/l} \) coefficients.

6.3 Background modelling

The background normalisation and the distributions of the event discriminants are derived by using the MC simulated samples described in Sect. 3. In light of the large theoretical uncertainty that affects the prediction of \( \text{tt} + \text{heavy-flavour} \) [64,69], the MadGraph sample is further divided into subsamples based on the quark flavour associated with the jets generated in the acceptance region \( p_T > 20 \text{ GeV}, \)
$|\eta| < 2.5$. Events are labelled as $t\bar{t} + b\bar{b}$ if at least two jets are matched within $\sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.5$ to bottom quarks not originating from the decay of a top quark. If only one jet is matched to a bottom quark, the event is labelled as $t\bar{t} + b$. These cases typically arise when the second extra b quark in the event is either too far forward or too soft to be reconstructed as a jet, or because the two extra b quarks are emitted almost collinearly and end up in a single jet. Similarly, if at least one reconstructed jet is matched to a c quark, the event is labelled as $t\bar{t} + c$. In the latter case, single- and double-matched events are treated as one background. If none of the above conditions is satisfied, the event is classified as $t\bar{t}$ plus light-flavour. Table 1 reports the number of events observed in the various categories, together with the expected signal and background yields. The latter are obtained from the signal-plus-background fit described in Sect. 8.

### 7 Systematic uncertainties

There are a number of systematic uncertainties of experimental and theoretical origin that affect the signal and the background expectations. Each source of systematic uncertainty is associated with a nuisance parameter that modifies the likelihood function used to extract the signal yield, as described in Sect. 8. The prior knowledge on the nuisance parameter is incorporated into the likelihood in a frequentist manner by interpreting it as a posterior arising from a pseudo-measurement [70]. Nuisance parameters can affect either the yield of a process (normalisation uncertainty), or the shape of the $P_{/V}$ and $P_{h/l}$ discriminants (shape uncertainty), or both. Multiple processes across several categories can be affected by the same source of uncertainty. In that case the related nuisance parameters are treated as fully correlated.

The uncertainty in the integrated luminosity is estimated to be $2.6\%$ [71]. The lepton trigger, reconstruction, and identification efficiencies are determined from control regions by using a tag-and-probe procedure. The total uncertainty is evaluated from the statistical uncertainty of the tag-and-probe measurement, plus a systematic uncertainty in the method, and is estimated to be $1.6\%$ per muon and $1.5\%$ per electron. It is conservatively approximated to a constant $2\%$ per charged lepton. The uncertainty on the jet energy scale (JES) ranges from $1\%$ up to about $8\%$ of the expected energy scale depending on the jet $p_T$ and $|\eta|$ [61]. For each simulated sample, two alternative distributions of the $P_{/V}$ and $P_{h/l}$ discriminants are obtained by varying the energy scale of all simulated jets up or down by their uncertainty, and the fit is allowed to interpolate between the nominal and the alternative distributions with a Gaussian prior [70]. A similar procedure is applied to account for the uncertainty related to the jet energy resolution (JER), which ranges between about $5$ and $10\%$ of the expected energy resolution depending on the jet direction. Since the analysis categories are defined in terms of the multiplicity and kinematic properties of the jets, a variation of either the scale or the resolution of the simulated jets can induce a migration of events in or out of the analysis categories, as well as migrations among different categories. The fractional change in the event yield induced by a shift of the JES (JER) ranges between $4–13\%$ (0.5–2\%).

#### Table 2 Summary of the systematic uncertainties affecting the signal and background expectation

<table>
<thead>
<tr>
<th>Source</th>
<th>Rate uncertainty (%)</th>
<th>Shape</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental uncertainties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated luminosity</td>
<td>2.6</td>
<td>No</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Trigger and lepton identification</td>
<td>2–4</td>
<td>No</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>JES</td>
<td>4–13</td>
<td>Yes</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>JER</td>
<td>0.5–2</td>
<td>Yes</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>b tagging</td>
<td>2–17</td>
<td>Yes</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Theoretical uncertainties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top $p_T$ modelling</td>
<td>3–8</td>
<td>Yes</td>
<td>✓</td>
</tr>
<tr>
<td>$\mu_T / \mu_F$ variations</td>
<td>2–25</td>
<td>Yes</td>
<td>✓</td>
</tr>
<tr>
<td>$t\bar{t} + b\bar{b}$ normalisation</td>
<td>50</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td>$t\bar{t} + b$ normalisation</td>
<td>50</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$ normalisation</td>
<td>50</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td>Signal cross section</td>
<td>7</td>
<td>No</td>
<td>✓</td>
</tr>
<tr>
<td>Background cross sections</td>
<td>2–20</td>
<td>No</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>PDF</td>
<td>3–9</td>
<td>No</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Statistical uncertainty (bin-by-bin)</td>
<td>4–30</td>
<td>Yes</td>
<td>✓ ✓ ✓</td>
</tr>
</tbody>
</table>
Fig. 3 Distribution of the $P_{sh}$ discriminant in the two $P_{sh}$ bins for the high-purity (H) categories. The signal and background yields have been obtained from a combined fit of all nuisance parameters with the constraint $\mu = 1$. The bottom panel of each plot shows the ratio between the observed and the overall background yields. The solid blue line indicates the ratio between the signal-plus-background and the background-only distributions. The shaded and solid green bands correspond to the $\pm 1\sigma$ uncertainty in the background prediction after the fit depending on the process type and on the category. When the JES and JER are varied from their nominal values, the $p_{T}^{\text{miss}}$ vector is recomputed accordingly. The scale factors applied to correct the CSV discriminator, as described in Sect. 4, are affected by several sources of systematic uncertainty. In the statistical interpretation, the fit can interpolate between the nominal and the two alternative distributions constructed by varying each scale factor up or down by its uncertainty.

Theoretical uncertainties are treated as process-specific if they impact the prediction of one simulated sample at the time. They are instead treated as correlated across several samples if they are related to common aspects of the simulation (e.g. PDF, scale variations). The modelling of the $t\bar{t}$+jets background is affected by a variety of systematic uncertainties. The uncertainty due to the top-quark $p_T$ modelling is evaluated by varying the reweighting function $r_t(p_T)$, where
$p_T^t$ is the transverse momentum of the generated top quark, between one (no correction at all) and 2$\sqrt{r_1}$ − 1 (the relative correction is doubled). This results in both a shape and a normalisation uncertainty. The latter can be as large as 20% for a top quark $p_T$ around 300 GeV, and corresponds to an overall normalisation uncertainty of about 3–8% depending on the category. To account for uncertainties in the $t\bar{t}$ jets acceptance, the factorisation and renormalisation scales used in the simulation are varied in a correlated way by factors of 1/2 and 2 around their central value. The scale variation is assumed uncorrelated among $t\bar{t} + b\bar{b}$, $t\bar{t} + b$, and $t\bar{t} + c\bar{c}$. In a similar way, independent scale variations are introduced for events with exactly one, two, or three extra partons in the matrix element. To account for possibly large K-factors due to the usage of a LO MC generator, the $t\bar{t} + b\bar{b}$, $t\bar{t} + b$, and $t\bar{t} + c\bar{c}$ normalisations predicted by the MADGRAPH simulation are assigned a 50% uncertainty each. This value can be seen as a conservative upper limit to the theoretical uncertainty in the observed and the overall background yields. The solid blue line indicates the ratio between the signal-plus-background and the background-only distributions. The shaded and solid green bands correspond to the ±1σ uncertainty in the background prediction after the fit.
8 Results

The statistical interpretation of the results is performed by using the same methodology employed for other CMS Higgs boson analyses and extensively documented in Ref. [2]. The measured signal rate is characterised by a strength modifier \( \mu = \sigma/\sigma_{SM} \) that scales the Higgs boson production cross section times branching fraction with respect to its SM expectation for \( m_H = 125 \text{ GeV} \). The nuisance parameters, \( \theta \), are incorporated into the likelihood as described in Sect. 7. The total likelihood function \( \mathcal{L}(\mu, \theta) \) is the product of a Poissonian likelihood spanning all bins of the \( (P_{b/b}, P_{b/1}) \) distributions for all the eight categories, times a likelihood function for the nuisance parameters. Based on the asymptotic properties of the profile likelihood ratio test statistic \( q(\mu) = -2 \ln [\mathcal{L}(\mu, \hat{\theta}_\mu)/\mathcal{L}(\hat{\mu}, \hat{\theta})] \), confidence intervals on \( \mu \) are set, where \( \hat{\theta} \) and \( \hat{\theta}_\mu \) indicate the best-fit value for \( \theta \) obtained when \( \mu \) is floating in the fit or fixed at a hypothesised value, respectively.

Figures 3 and 4 show the binned distributions of \( (P_{b/b}, P_{b/1}) \) in the various categories and for the two channels. For visualisation purposes, the two-dimensional histograms are projected onto one dimension by showing first the distribution of \( P_{a/b} \) for events with \( P_{b/1} < 0.5 \) and then for \( P_{b/1} \geq 0.5 \). The observed distributions are compared to the signal-plus-background expectation obtained from a combined fit to all categories with the constraint \( \mu = 1 \). No evidence of a \( t\bar{t}H \) signal over the background is observed. The statistical interpretation is performed both in terms of exclusion upper limits (UL) at a 95% CL, where the modified CLs prescription [73, 74] is adopted to quote confidence intervals, and in terms of the maximum likelihood estimator of the strength modifier \( \hat{\mu} \).

Figure 5 (top) shows the observed 95% CL UL on \( \mu \), compared to the signal-plus-background and to the background-
The best-fit values of the signal strength modifier obtained from the SL and DL channels alone, and from their combination

<table>
<thead>
<tr>
<th>Channel</th>
<th>Best-fit $\mu$</th>
<th>Observed UL (signal injected)</th>
<th>Median exp. UL (background only)</th>
<th>Median exp. UL (background-only)</th>
<th>$\pm 1\sigma$ CL interval</th>
<th>$\pm 2\sigma$ CL interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL</td>
<td>+1.7$^{+2.0}_{-1.8}$</td>
<td>5.5</td>
<td>5.0</td>
<td>4.2</td>
<td>[2.9, 6.2]</td>
<td>[2.1, 9.1]</td>
</tr>
<tr>
<td>DL</td>
<td>+1.0$^{+3.3}_{-3.0}$</td>
<td>7.7</td>
<td>7.8</td>
<td>6.9</td>
<td>[4.7, 10.6]</td>
<td>[3.4, 15.8]</td>
</tr>
<tr>
<td>Combined</td>
<td>+1.2$^{+1.6}_{-1.5}$</td>
<td>4.2</td>
<td>4.1</td>
<td>3.3</td>
<td>[2.3, 4.9]</td>
<td>[1.7, 7.0]</td>
</tr>
</tbody>
</table>

The observed 95% CL UL on $\mu$ are given in the third column, and are compared to the median expected limits for both the signal-plus-background and for the background-only hypotheses. For the latter, the $\pm 1\sigma$ and $\pm 2\sigma$ CL intervals are also given.

9 Summary

A search for Higgs boson production in association with a top-quark pair with $H \rightarrow b\bar{b}$ has been presented. A total of 19.5 fb$^{-1}$ of pp collision data collected by the CMS experiment at $\sqrt{s} = 8$ TeV has been analysed. Events with one lepton and at least five jets or two opposite-sign leptons and at least four jets have been considered. Jet b-tagging information is exploited to suppress the $t\bar{t}$ plus light-flavour background. A probability density value under either the $t\bar{t}H$ or the $t\bar{t} + b\bar{b}$ background hypothesis is calculated for each event using an analytical matrix element method. The ratio of probability densities under these two competing hypotheses allows a one-dimensional discriminant to be defined, which is then used together with b-tagging information in a likelihood analysis to set constraints on the signal strength modifier $\mu = \sigma/\sigma_{SM}$.

No evidence of a signal is found. The expected upper limit at a 95% CL is $\mu < 3.3$ under the background-only hypothesis. The observed limit is $\mu < 4.2$, corresponding to a best-fit value $\hat{\mu} = 1.2^{+1.6}_{-1.5}$. Within the present statistics, the analysis documented in this paper yields competitive results compared to those obtained on the same data set and for the same final state by using non-analytical multivariate techniques [28]. However, the matrix element method applied for a maximal separation between the signal and the dominant $t\bar{t} + b\bar{b}$ background allows for a better control of the systematic uncertainty due to this challenging background. This method represents a promising strategy towards a precise determination of the top quark Yukawa coupling. Once the statistical uncertainty will be reduced by the inclusion of the upcoming 13 TeV collision data, systematic uncertainties will start to play a more important role. By incorporating experimental and theoretical model parameters into an event likelihood, the matrix element method offers a natural handle to minimise the impact of systematic uncertainties on the extraction of the signal.
Acknowledgments We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l’Énergie Atomique et aux Énergies Alternatives / CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innovation Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (RFN), Republic of Korea; the Lithuanian Academy of Sciences; the Ministry of Education and University of Malaya (Malaysia); the Mexican Funding Agencies (CINVESTAV, CONACYT, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Science and Higher Education and the National Science Centre, Poland; the Fundação para a Ciência e a Tecnologia, Portugal; JINR, Dubna; the Ministry of Education and Science of the Russian Federation, the Federal Agency of Atomic Energy, the Federal Agency of Atomic Energy of the Russian Federation, Russian Academy of Sciences, and the Russian Foundation for Basic Research; the Ministry of Education, Science and Technological Development of Serbia; the Secretaria de Estado de Investigacion, Desarrollo e Innovacion and Programa Consolider-Ingenio 2010, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, UniZH, Canton Zurich, and SER); the Ministry of Science and Technology, Taipei; the Thailand Center of Excellence in Physics, the Institute for the Promotion of Teaching Science and Technology of Thailand, Special Task Force for Activating and Research and the National Science and Technology Development Agency of Thailand; the Scientific and Technical Research Council of Turkey, and Turkish Atomic Energy Authority; the National Academy of Sciences of Ukraine, and State Fund for Fundamental Researches, Ukraine; the Science and Technology Facilities Council, UK; the US Department of Energy, and the US National Science Foundation. Individuals have received support from the Marie-Curie programme and the European Research Council and EPLANET (European Union); the Levitsky Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS programme of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund; the Compagnia di San Paolo (Torino); the Consorzio per la Fisica (Trieste); MIUR project 2010RT4XTM (Italy); the Thalis and Aristeia programmes cofinanced by EU-ESF and the Greek NSF; and the National Priorities Research Program by Qatar National Research Fund.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

References


University of Canterbury, Christchurch, New Zealand
P. H. Butler, S. Reucroft

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
A. Ahmad, M. Ahmad, Q. Hassan, H. R. Hoorani, W. A. Khan, T. Khurshid, M. Shoaib

National Centre for Nuclear Research, Swierk, Poland

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal

Joint Institute for Nuclear Research, Dubna, Russia

Petersburg Nuclear Physics Institute, Gatchina (St. Petersburg), Russia

Institute for Nuclear Research, Moscow, Russia

Institute for Theoretical and Experimental Physics, Moscow, Russia
V. Epshteyn, V. Gavrilov, N. Lychkovskaya, V. Popov, I. Pozdnyakov, G. Safronov, S. Semenov, A. Spiridonov, V. Stolin, E. Vlasov, A. Zhokin

P. N. Lebedev Physical Institute, Moscow, Russia

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
A. Belyaev, E. Boos, V. Bunichev, M. Dubinin, L. Dudko, A. Ershov, A. Gribushin, V. Klyukhin, O. Kodolova, I. Lokhtin, S. Obraztsov, S. Petrushanko, V. Savrin

State Research Center of Russian Federation, Institute for High Energy Physics, Protvino, Russia

University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
P. Adzic, E. Ekmedzic, J. Milosevic, V. Rekovic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Universidad Autónoma de Madrid, Madrid, Spain
C. Albajar, J. F. de Trocóniz, M. Missiroli, D. Moran

Universidad de Oviedo, Oviedo, Spain
H. Brun, J. Cuevas, J. Fernandez Menendez, S. Folgueras, I. Gonzalez Caballero

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain
J. A. Brochero Cifuentes, I. J. Cabrillo, A. Calderon, J. Duarte Campderrlos, M. Fernandez, G. Gomez, A. Graziano,
University of Colorado at Boulder, Boulder, USA

Cornell University, Ithaca, USA

Fairfield University, Fairfield, USA
D. Winn

Fermi National Accelerator Laboratory, Batavia, USA

University of Florida, Gainesville, USA

Florida International University, Miami, USA
S. Hewamanage, S. Linn, P. Markowitz, G. Martinez, J. L. Rodriguez

Florida State University, Tallahassee, USA

Florida Institute of Technology, Melbourne, USA
M. M. Baarmand, M. Hohlmann, H. Kalakhety, F. Yumiceva

University of Illinois at Chicago (UIC), Chicago, USA

The University of Iowa, Iowa City, USA

Johns Hopkins University, Baltimore, USA

The University of Kansas, Lawrence, USA

Kansas State University, Manhattan, USA
I. Chakaberia, A. Ivanov, K. Kaadze, S. Khalil, M. Makouski, Y. Maravin, L. K. Saini, N. Skhirtladze, I. Svintradze

Lawrence Livermore National Laboratory, Livermore, USA
J. Gronberg, D. Lange, F. Rebassoo, D. Wright

University of Maryland, College Park, USA
24: Also at University of Tehran, Department of Engineering Science, Tehran, Iran
25: Also at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran
26: Also at Università degli Studi di Siena, Siena, Italy
27: Also at Centre National de la Recherche Scientifique (CNRS)-IN2P3, Paris, France
28: Also at Purdue University, West Lafayette, USA
29: Also at International Islamic University of Malaysia, Kuala Lumpur, Malaysia
30: Also at Institute for Nuclear Research, Moscow, Russia
31: Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia
32: Also at National Research Nuclear University, “Moscow Engineering Physics Institute” (MEPhI), Moscow, Russia
33: Also at California Institute of Technology, Pasadena, USA
34: Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia
35: Also at Facoltà Ingegneria, Università di Roma, Rome, Italy
36: Also at Scuola Normale e Sezione dell’INFN, Pisa, Italy
37: Also at University of Athens, Athens, Greece
38: Also at Paul Scherrer Institut, Villigen, Switzerland
39: Also at Institute for Theoretical and Experimental Physics, Moscow, Russia
40: Also at Albert Einstein Center for Fundamental Physics, Bern, Switzerland
41: Also at Gaziosmanpasa University, Tokat, Turkey
42: Also at Adiyaman University, Adiyaman, Turkey
43: Also at Mersin University, Mersin, Turkey
44: Also at Cag University, Mersin, Turkey
45: Also at Piri Reis University, Istanbul, Turkey
46: Also at Anadolu University, Eskisehir, Turkey
47: Also at Ozyegin University, Istanbul, Turkey
48: Also at Yildiz Technical University, Istanbul, Turkey
49: Also at Necmettin Erbakan University, Konya, Turkey
50: Also at Mimar Sinan University, Istanbul, Istanbul, Turkey
51: Also at Marmara University, Istanbul, Turkey
52: Also at Kafkas University, Kars, Turkey
53: Also at Yildiz Technical University, Istanbul, Turkey
54: Also at Rutherford Appleton Laboratory, Didcot, UK
55: Also at School of Physics and Astronomy, University of Southampton, Southampton, UK
56: Also at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
57: Also at Argonne National Laboratory, Argonne, USA
58: Also at Erzincan University, Erzincan, Turkey
59: Also at Texas A&M University at Qatar, Doha, Qatar
60: Also at Kyungpook National University, Taegu, Korea