PROCEEDINGS OF SCIENCE

Measurements of the $t\bar{t}$H+tH production at CMS

PoS(LHCP2024)133

Matteo Marchegiani* **on behalf of the CMS collaboration**

ETH Zürich,
Otto-Stern-Weg 5, 8093 Zürich, Switzerland
E-mail: matteo.marchegiani@cern.ch

Following the observation of the Higgs boson at the LHC, the high precision measurements of its properties and the search for its production in various final states have become crucial to test the validity of the standard model. The production of the Higgs boson in association with a top quark and antiquark pair ($t\bar{t}$H) is a direct probe of the interaction between the boson and the top quark, providing access to the top-Higgs Yukawa coupling. On the other hand, the production of the Higgs boson in association with a single top quark (tH) is sensitive to the sign of this coupling. The importance of a precise measurement of the top Yukawa coupling stems from its dominant role in the stability of the electroweak vacuum. The latest measurements of the $t\bar{t}$H and tH production cross section performed by the CMS collaboration with proton-proton collisions at $\sqrt{s}$ = 13 TeV in final states with a Higgs boson decaying to a pair of photons, multiple charged leptons or a bottom quark and antiquark pair are presented.

12th Large Hadron Collider Physics Conference (LHCP2024)
3-7 June 2024
Boston, USA

*Speaker

1. Introduction

The observation of the Higgs boson (H) by the CMS [1–3] and ATLAS [4, 5] collaborations at the LHC in 2012 has confirmed the existence of a scalar boson of 125 GeV mass, that was predicted to generate the mass of the electroweak gauge bosons via the electroweak symmetry breaking mechanism. Although this discovery further validates the standard model (SM) of particles, the properties of the Higgs boson still need to be measured with precision in order to unveil the nature of this particle and characterize its interaction with SM fermions and gauge bosons. In the SM, the Higgs boson couples to fermions through a Yukawa-type interaction, with a coupling strength proportional to the fermion mass. The top quark Yukawa coupling (y_t) is the largest coupling of the Higgs boson to fermions and has a critical role in the computation of loop corrections to the Higgs boson mass and its self-coupling [6]. As a consequence, probing the coupling of the Higgs boson to the top quark is essential to test the stability of electroweak vacuum, since the predicted shape of the Higgs potential strongly depends on the top quark mass.

In order to measure the Yukawa coupling of the Higgs boson to fermions, it is possible to measure the branching ratio (BR) of the decay to fermion-antifermion pairs. However, since the Higgs boson is lighter than the top quark, it cannot decay to top quarks. Therefore, the only possible direct measurement of the top quark Yukawa coupling is via the Higgs boson production in association with a top quark and antiquark pair ($t\bar{t}$H) or with a single top quark (tH). While the observation of the $t\bar{t}$H final state is challenging due to its rarity, with a cross section of 507 fb at $\sqrt{s}$ = 13 TeV [7], it was achieved by the CMS collaboration in 2018 with a significance of 5.2 standard deviations (SD) combining measurements in the H $\rightarrow$ $b\bar{b}$, $\gamma\gamma$, $\tau^+\tau^-$, WW* and ZZ* decay modes [8] analyzing data recorded at $\sqrt{s}$ = 7, 8 and 13 TeV corresponding to an integrated luminosity of 5.1, 19.7 and 35.9 fb^{-1}. The tH process, instead, although it has not been observed yet, provides a constraint on the sign of y_t. Furthermore, the combined measurement of the $t\bar{t}$H and tH production provides a probe of the CP-odd admixtures in the top-Higgs coupling, which would indicate the presence of physics beyond the SM.

These proceedings include the latest results obtained by the CMS collaboration in measurements of the $t\bar{t}$H process with proton-proton collisions at $\sqrt{s}$ = 13 TeV in final states with a Higgs boson decaying to a pair of photons (H $\rightarrow$ $\gamma\gamma$) [9, 10], multiple charged leptons [11] or a bottom quark and antiquark pair (H $\rightarrow$ $b\bar{b}$) [12, 13]. In particular, these results feature the first effective field theory (EFT) interpretation measurement by the CMS collaboration in the $t\bar{t}$H (H $\rightarrow$ $b\bar{b}$) final state.

2. Measurements in the $t\bar{t}$H (H $\rightarrow$ $\gamma\gamma$) final state

Despite the branching ratio of the Higgs boson decaying to a pair of photons is only 0.2% [14], the H $\rightarrow$ $\gamma\gamma$ channel is one of the most sensitive to the $t\bar{t}$H production and provides a clear signature to reconstruct the Higgs boson decay and characterize the CP structure of the $t\bar{t}$H coupling.

In the $t\bar{t}$H (H $\rightarrow$ $\gamma\gamma$) analysis [9], the trigger selects photons with a loose selection on calorimetric information and asymmetric photon transverse energy (E_T) thresholds of 30 and 18 (22) GeV for data collected during 2016 (2017-2018). H candidates are selected offline by requiring the presence of two prompt photons reconstructed from energy clusters in the electromagnetic calorimeter (ECAL) passing a loose identification criterion based on a boosted decision tree (BDT) classifier trained to separate photons from jets. The invariant mass of the diphoton pair is required to be within $100 < m_{\gamma\gamma} < 180$ GeV. Events are divided into two exclusive categories: the leptonic and hadronic channels. The leptonic channel is targeting events where at least one top quark decays leptonically, whereas the hadronic channel targets $t\bar{t}$ fully hadronic decays.

A BDT discriminant ("BDT-bkg") trained with kinematic properties of jets, leptons, photons and diphotons as inputs is employed in each channel to distinguish $t\bar{t}$H and background events. Events are either rejected or further divided into eight categories to maximize the expected significance and four categories to maximize the sensitivity to the CP structure of y_t. A simultaneous binned maximum likelihood (ML) fit to the $m_{\gamma\gamma}$ distributions in the eight categories is performed to extract the signal strength $\mu_{t\bar{t}H}$, defined as the ratio of the measured to SM cross section assuming the SM BR for the $H \to \gamma\gamma$ decay. The $t\bar{t}H$ signal distribution of $m_{\gamma\gamma}$ is parametrized using a double-sided Crystal Ball plus Gaussian function, while the background is modeled from data with continuous functions, applying the discrete profiling method to account for the arbitrary choice of the function. The fit yields a signal strength of $\mu_{t\bar{t}H} = 1.38^{+0.36}_{-0.29}$, corresponding to an observed (expected) significance of 6.6 (4.7) SD, resulting in the first observation of $t\bar{t}H$ in a single decay channel.

The CP structure of the $t\bar{t}H$ coupling can be inferred by categorizing events with the optimal observable $\mathcal{D}_{0-}$ which is extracted from the output of a BDT regression trained to distinguish CP-even and CP-odd contributions exploiting several kinematical variables as inputs. Events are split into 12 categories, leptonic or hadronic, two BDT-bkg categories and three $\mathcal{D}_{0-}$ bins. A simultaneous fit to the $m_{\gamma\gamma}$ distribution is performed using the 12 categories to measure f_{CP}^{Htt}, which is the fractional contribution of the CP-odd component of the cross section, resulting in $f_{CP}^{Htt} = 0.00 \pm 0.33$ with the constraint $|f_{CP}^{Htt}| < 0.67$ at 95% confidence level (CL), as shown in Fig. 1 (left). Therefore, the pseudoscalar model of CP structure of the $t\bar{t}H$ coupling ($f_{CP}^{Htt} = 1$) is excluded at 3.2 SD (2.6 SD expected).

The results obtained analyzing the $t\bar{t}H$ ($H \to \gamma\gamma$) final state are also interpreted in the simplified template cross section (STXS) framework to extract a differential cross section as a function of the reconstructed Higgs boson transverse momentum (p_T^H) [10]. The same categorization as in Ref. [9] is applied to define $t\bar{t}H$ enriched categories, which are further split in five bins of the diphoton transverse momentum ($p_T^{\gamma\gamma}$) to target specific STXS bins. Additionally, a category targeting the tH process is defined by selecting events enriched in tHq with a leptonically decaying top quark. The results of the best fit cross sections are shown in Fig. 1 (right), where the cross sections for the $t\bar{t}H$ ($H \to \gamma\gamma$) process are in agreement with the SM prediction. The observed (expected) upper limit at 95% CL on the tH production rate is found to be 14 (8) times the SM prediction.

PoS(LHCP2024)133

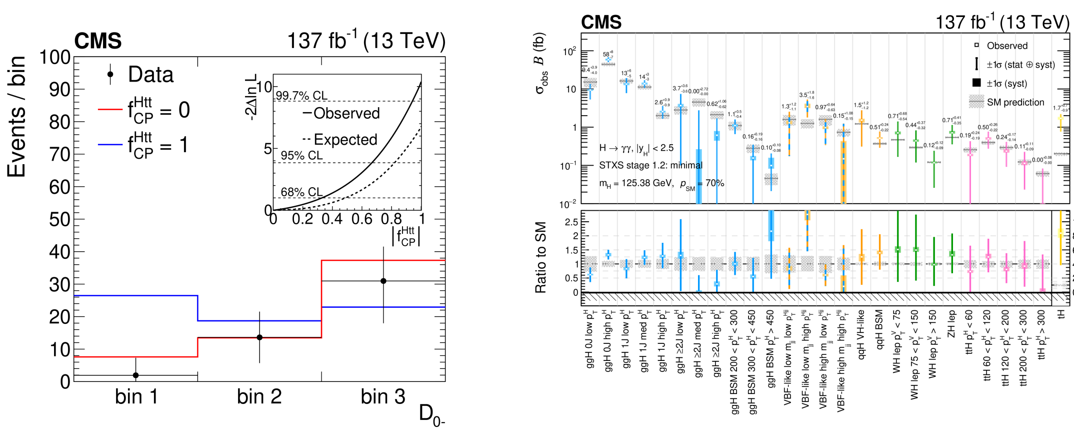

Figure 1: Left: Distribution of events weighed by $S/(S+B)$, in three bins of the $\mathcal{D}_{0-}$ discriminant. The inner panel shows the likelihood scan for $|f_{CP}^{Htt}|$ [9].
Right: Observed results of the STXS fit. The best fit cross sections are plotted together with the respective 68% CL intervals. The lower panel shows the ratio of the fitted values to the SM predictions [10].

3. Measurements in the $t\bar{t}$H multilepton final state

The search for $t\bar{t}$H production in final states containing multiple charged leptons (e and μ) and hadronically decaying τ leptons (τ_h) [11] targets the H $\rightarrow$WW, H $\rightarrow \tau\tau$ and H $\rightarrow$ZZ decays in association with a top quark pair decaying semi-leptonically or hadronically. Although these decay channels have a larger branching ratio than the H $\rightarrow \gamma\gamma$ decay channel, they are affected by larger background contamination. In order to enhance the sensitivity to the $t\bar{t}$H and tH signal strengths, events are categorized based on lepton flavour, multiplicity and electric charge, for a total of ten mutually exclusive categories of which three enriched in the tH final state. In order to separate the $t\bar{t}$H and tH signals from background events, artificial neural networks (ANNs) are used in the tH enriched categories, and events are analyzed in subcategories based on the lepton flavour and on the b-tagged jets (b jets) multiplicity. In the remaining categories, BDTs are used to separate the $t\bar{t}$H signal from background events. Among the reducible backgrounds, the dominant one is the contribution coming from misidentified nonprompt leptons or hadrons and misidentified τ_h, which are estimated with a data-driven method. The irreducible backgrounds include the $t\bar{t}$W, $t\bar{t}$Z/$t\bar{t}\gamma^*$, Drell-Yan and diboson backgrounds. Two dedicated control regions (CR) enriched in the $t\bar{t}$W and $t\bar{t}$Z backgrounds are defined and are used to constrain the corresponding background yields in the signal region (SR).

The $t\bar{t}$H and tH production rates are determined through a binned simultaneous maximum likelihood fit to the ANN and BDT distributions in the SR and distributions of discriminating observables in the CRs, in each subcategory. The measured signal strength for the $t\bar{t}$H production rate is $\mu_{t\bar{t}H} = 0.92 \pm 0.19$ (stat)$^{+0.17}_{-0.13}$ (syst) and $\mu_{tH} = 5.7 \pm 2.7$ (stat) $\pm$ 3.0 (syst) for the tH process. The corresponding observed (expected) significance of the $t\bar{t}$H signal amounts to 4.7 (5.2) SD and that of the tH signal to 1.4 (0.3) SD, and are in agreement with their SM expectation, as shown in Fig. 2 (left). The event yields of background processes obtained from the ML fit agree reasonably well with their expected production rate, with the $t\bar{t}$Z and $t\bar{t}$W production rates determined to be $\mu_{t\bar{t}Z}$ = 1.03 $\pm$ 0.14 (stat+syst) and $\mu_{t\bar{t}W}$ = 1.43 $\pm$ 0.21 (stat+syst) times their SM expectation. The measured signal yields are also interpreted as a function of the coupling modifier of the Higgs boson to the top quark (κ_t) and to vector bosons (κ_V). The likelihood function is profiled to test the compatibility of data with different values of (κ_t, κ_V). The 95% CL region on κ_t corresponds to the intervals $-0.9 < \kappa_t < -0.7$ and $0.7 < \kappa_t < 1.1$, which is compatible with the SM expectation of $\kappa_t = 1$, as shown in Fig. 2 (right).

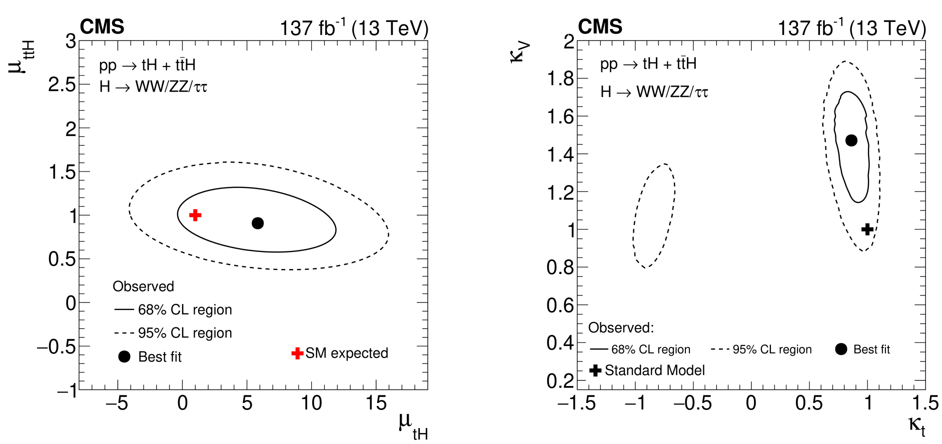

Figure 2: Left: Two-dimensional contours of the likelihood function, as a function of the production rates of the $t\bar{t}$H and tH signals.
Right: Two-dimentsional contours of the likelihood function, as a function of the coupling modifiers κ_t and κ_V [11].

4. Measurements in the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) final state

The $t\bar{t}H$ production with $H \rightarrow b\bar{b}$ decays [12] features the Higgs decay channel with the largest branching ratio of 58% [14]. Furthermore, the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) process involves Higgs boson couplings only to fermions and thus exclusively probes the Higgs Yukawa sector. Despite attractive for its larger cross section compared to other channels, the $H \rightarrow b\bar{b}$ final state is challenging due to considerable SM backgrounds and the combinatorial ambiguity in assigning jets to the Higgs boson candidate. In particular, the irreducible $t\bar{t}$+$b\bar{b}$ background, consisting in the production of a $t\bar{t}$ pair in association with additional bottom quarks, has identical final state as the $t\bar{t}H$ signal and constitutes the critical background to the analysis. The simulation of the $t\bar{t}$+$b\bar{b}$ background is affected by large systematic uncertainties impacting the analysis. To better constrain the background processes, $t\bar{t}$ events are categorized into $t\bar{t}B$, $t\bar{t}C$, and $t\bar{t}LF$ events, depending on whether the flavour of the hadrons in the additional jets is B, C, or light flavour (LF).

Three mutually exclusive channels are considered in the analysis, depending on the number of leptons (e or μ) in the $t\bar{t}$ decay modes: the fully hadronic (FH) channel, in which both top quarks decay hadronically, the single-lepton (SL) channel in which one top quark decays hadronically and the other leptonically and the dilepton (DL) channel in which both top quarks decay leptonically. The SL channel also includes dedicated event categories targeting tH production. Events are further categorized based on the jet and b jet multiplicities to enhance the sensitivity to the $t\bar{t}H$ and tH signals. The signal extraction is performed by training dedicated ANNs in each analysis channel with several discriminating observables as inputs to distinguish the $t\bar{t}H$ and tH signals from background events.

The $t\bar{t}H$ and tH production rates are measured through a binned profile likelihood fit to the final discriminants, that include ANN output scores, ratios of ANN discriminants or the event yields depending on the category. The measured $t\bar{t}H$ signal strength obtained combining all channels is $\mu_{t\bar{t}H} = 0.33^{+0.17}_{-0.16}$ (stat)$^{+0.20}_{-0.21}$ (syst), corresponding to an observed (expected) significance of 1.3 (4.1) SD. The fitted values of the $t\bar{t}B$ and $t\bar{t}C$ background normalizations are $1.19^{+0.13}_{-0.12}$ and $1.07^{+0.20}_{-0.19}$, with an anticorrelation of 48% between $\mu_{t\bar{t}H}$ and the $t\bar{t}B$ background normalisation, which is visible in Fig. 3 (left). An upper limit at 95% CL on the tH signal strength μ_{tH} of 14.6 is obtained. The results are also interpreted in the STXS framework by splitting events in different regions of p_T^H and are compatible with the inclusive measurement of $\mu_{t\bar{t}H}$. The CP structure of the top-Higgs coupling is probed by profiling the likelihood ratio as a function of the CP-even and CP-odd coupling modifiers, κ_t and $\widetilde{\kappa}_t$, with the best fit values of (+0.53, 0.00) for (κ_t, $\widetilde{\kappa}_t$), compatible with the SM at the level of 2 SD.

An EFT interpretation measurement is performed in a dedicated analysis targeting the $t\bar{t}H$ and $t\bar{t}Z$ final states [13]. Events containing a large-radius jet identified as a boosted Z or Higgs boson decaying to bottom quarks are selected using multivariate techniques and effects of new physics are probed by adding eight possible dimension-six operators to the SM Lagrangian. The observed 68% and 95% CL intervals for the Wilson coefficients (WC) are shown in Fig. 3 (right), providing the first constraints of EFT effects in the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) process. No indications of new physics are observed in data.

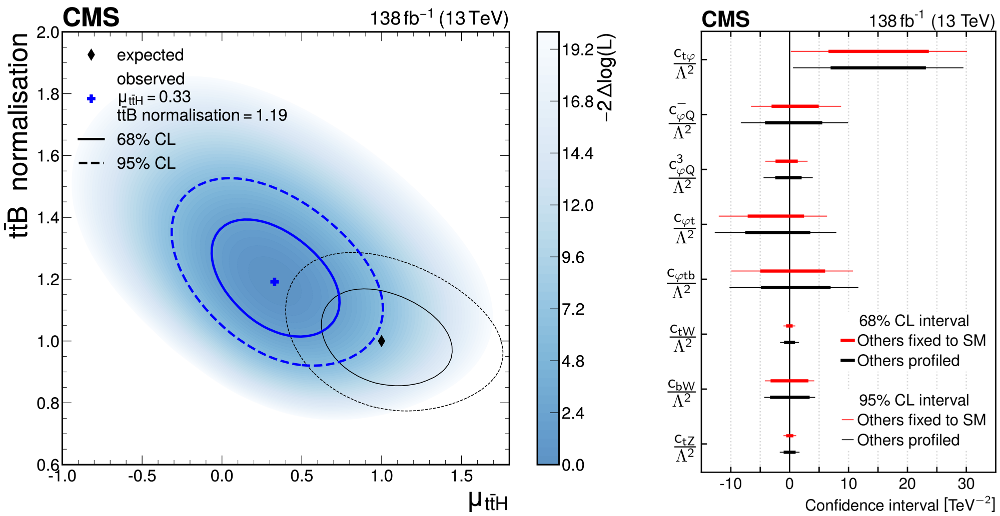

Figure 3: Left: Observed likelihood-ratio test statistic (blue shading) as a function of the $t\bar{t}H$ signal-strength modifier $\mu_{t\bar{t}H}$ and the $t\bar{t}B$ background normalisation, together with the observed (blue) and SM expected (black) best fit points [12]. **Right:** The observed 68% and 95% CL intervals for the WCs are shown by the thick and thin bars, respectively. The intervals are found by scanning over a single WC while either profiling the other seven WCs (black) or fixing them to the SM value of zero (red) [13].

5. Conclusions

The top quark Yukawa coupling is a free parameter of the SM that has a crucial role in the stability of the electroweak vacuum and is sensitive to new physics. The $t\bar{t}H$ and tH processes are direct probes of the top-Higgs interaction and have been explored with several measurements using proton-proton collisions at $\sqrt{s}$ = 13 TeV recorded by the CMS experiment during its second data-taking campaign. The $t\bar{t}H$ and tH processes pose challenges due to their rarity and the complexity of their final states, therefore the finest multivariate techniques are essential to extract relevant information from data.

The first observation of $t\bar{t}H$ in a single Higgs boson decay channel is achieved in the $t\bar{t}H$ ($H \to \gamma\gamma$) measurement and the $t\bar{t}H$ signal strength measured in the multilepton and $t\bar{t}H$ ($H \to b\bar{b}$) channels are compatible with the SM. The CP structure of the top quark Yukawa coupling is probed with dedicated measurements that exclude a pure CP-odd coupling at 95% CL. The first constraints on Wilson coefficients accounting for EFT effects in the $t\bar{t}H$ ($H \to b\bar{b}$) final state are presented and show no sign of new physics.

All presented results are in agreement with the SM expectations.

References

[1] CMS collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08004.

[2] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012) 30.

[3] CMS collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV*, *Journal of High Energy Physics* **2013** (2013) .

[4] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) S08003.

[5] CMS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012) 1.

[6] S. Alekhin, A. Djouadi and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, *Physics Letters B* **716** (2012) 214.

[7] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, J. Mazzitelli and C. Savoini, *Higgs Boson Production in Association with a Top-Antitop Quark Pair in Next-to-Next-to-Leading Order QCD*, *Phys. Rev. Lett.* **130** (2023) 111902.

[8] CMS collaboration, *Observation of $t\bar{t}H$ Production*, *Physical Review Letters* **120** (2018) 231801.

[9] CMS collaboration, *Measurements of $t\bar{t}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel*, *Phys. Rev. Lett.* **125** (2020) 061801.

[10] CMS collaboration, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{\mathrm{s}}$ = 13 TeV*, *Journal of High Energy Physics* **2021** (2021) .

[11] CMS collaboration, *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **81** (2021) .

[12] CMS collaboration, *Measurement of the* $\mathrm{t\bar{t}H}$ *and* tH *production rates in the* $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ *decay channel with* $138\,\mathrm{fb}^{-1}$ *of proton-proton collision data at* $\sqrt{s} = 13\,\mathrm{TeV}$, *CMS-PAS-HIG-19-011* (2023) .

[13] CMS collaboration, *Search for new physics using effective field theory in 13 TeV pp collision events that contain a top quark pair and a boosted Z or Higgs boson*, *Phys. Rev. D* **108** (2023) 032008.

[14] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, *CERN-2017-002-M* (2017) .