Top associated Higgs production in the ttH into multileptons channel with one tau lepton in ATLAS

Frank Seifert[*] on behalf of the ATLAS Collaboration
IEAP CTU in Prague
E-mail: Frank.Seifert@cern.ch

Abstract

The measurement of the Higgs boson and its cross-section in top associated production allows a direct measurement of the top quark Yukawa coupling at tree level. This can be compared with the indirect one in gluon-gluon-fusion production. The decay channel of the ttH system in the multilepton final state gives one of the largest sensitivities. The sub-channel focused on here is that with two same-charged light leptons (electrons or muons) and one τ_{had}: ttH$\rightarrow$2l+1τ_{had}. An overview over the ttH to multileptons analysis is given and the selection and data-driven background estimation of the 2l+1τ_{had} channel is described in more detail. The analysis includes a dataset of 20.3 fb^{-1} which has been recorded in LHC Run 1 at a center-of-mass energy of $\sqrt{s} = 8$ TeV with the ATLAS detector.

With the recorded data set of Run 1 no significant excess of events above the background expectation could be observed and the best fit for the ttH production cross section combining all multilepton channels, assuming a Higgs boson mass of 125 GeV, is $2.1^{+1.4}_{-1.2}$ times the SM expectation. The observed (expected) upper limit at the 95% confidence level is 4.7 (2.4) times the SM rate. The p-value for compatibility with the background-only hypothesis is 1.8σ and the expectation in the presence of a Standard Model signal is 0.9σ.

The European Physical Society Conference on High Energy Physics
22–29 July 2015
Vienna, Austria

[*]Speaker.

1. Introduction

A new particle, H, with a mass of about 125 GeV and compatible with the Standard Model (SM) Higgs boson [1–4] was discovered at the LHC by the ATLAS [5, 6] and CMS [7, 8] experiments in 2012 using proton-proton collisions with center-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV.

The measurement of the process in which the Higgs boson is produced in association with a top quark pair ($t\bar{t}H$) would permit a direct measurement of the top quark–Higgs boson Yukawa coupling in a process that is tree-level at the lowest order, which is otherwise accessible primarily through loop effects. Having both the tree- and loop-level measurements would allow disambiguation of new physics effects that could affect the two differently, such as dimension-six operators contributing to the ggH loop vertex.

Here, a search for the SM Higgs boson in the $t\bar{t}H$ production mode in multilepton final states at the ATLAS experiment of the LHC at a center-of-mass energy of $\sqrt{s} = 8$ TeV is described. The five final states considered are: two same-charge-sign light leptons (e or μ) with no additional hadronically decaying τ lepton (τ_{had}), three light leptons, two same-sign light leptons with one τ_{had}, four light leptons and one light lepton with two τ_{had}. These channels are sensitive to the Higgs decays $H \to WW^*$, $\tau\tau$, and ZZ^* produced in association with a top quark pair.

2. Event selection

The first step of selection is a single-lepton (e or μ) trigger which all events considered in this analysis have to pass. These achieve their maximal plateau efficiency for lepton $p_T > 25$ GeV. The detailed selection of the objects is described in Ref. [9].

Primarily $H \to WW^*$ and $\tau\tau$ decay modes are considered in addition to the final state of the $t\bar{t}$ system. As the largest background, $t\bar{t}$ decay, can produce a maximum of two light leptons which are also opposite-sign, the analysis requires two same-sign leptons or three or more leptons to suppress what would otherwise be the largest single background.

The analysis categories are classified by the number of light leptons and τ_{had} candidates. Events are initially classified by counting the number of light leptons with $p_T > 10$ GeV and τ_{had} candidates. In some cases the lepton selection criteria are tightened further by raising the p_T threshold, tightening isolation selections or restricting the allowed $|\eta|$ range, as explained in more detail in Ref. [9]. It is ensured that an event can only contribute to a single category. The event selection of the category with two light leptons and one τ_{had} is described in more detail here.

Such events are required to include exactly two light leptons, with the same charge and leading (subleading) $p_T > 25$ (15) GeV, and exactly one (τ_{had}) with $p_T > 25$ GeV. The reconstructed charge of the τ_{had} candidate has to be opposite to that of the light leptons. To reduce $t\bar{t}$+jets and $t\bar{t}V$ backgrounds, at least four reconstructed jets are required. To suppress diboson and single-boson backgrounds, at least one jet must be b-tagged. To suppress the $Z \to \ell^+\ell^-$+jets background due to charge-misidentified electrons, events with dielectron invariant mass within 10 GeV around the Z mass are vetoed. A Feynman diagram of production and decay of the ttH signal process targeted by this selection is shown in Fig. 1.
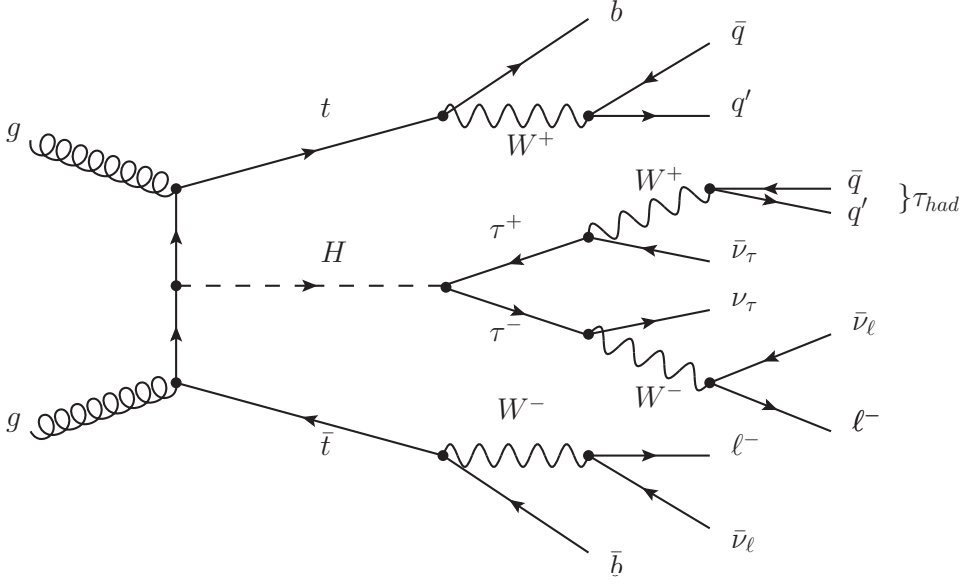
Figure 1: Feynman diagram of the production and decay of a Higgs boson in association with a top quark pair. The Higgs decay into two tau leptons (one decaying leptonically and one hadronically) is shown.

3. Background estimation

The main backgrounds to the signal are $t\bar{t}$ production with additional jets and non-prompt leptons and associated production of a top quark pair and a vector boson W or Z, hence denoted as $t\bar{t}V$. Important irreducible backgrounds include $t\bar{t}V$ and diboson production and are estimated from Monte Carlo (MC) simulation. To check the modelling of data by simulation, validation regions (VR) enriched in these processes are constructed, which is described in more detail in Ref. [9]. Fig. 2 shows data together with the different MC contributions in the validation regions for the most important irreducible backgrounds $t\bar{t}Z$ and $t\bar{t}W$.

Reducible (fake) backgrounds contain non-prompt leptons or mis-identification, and are estimated by data-driven techniques. The focus here is on the channel with 2 same-sign light leptons and one τ_{had}. In the following, a tight lepton is defined as a lepton that passes the nominal selection, a sideband lepton is defined as a lepton candidate which satisfies different criteria than the tight lepton selection (identification selection, isolation, or p_{T}), and (sideband) control regions either require one or more sideband leptons to replace a tight lepton in the signal region selection, or have the same lepton selection as the signal region but different requirements on the jet multiplicity.

Due to the same-charge sign requirement of the two light leptons, $t\bar{t}$ events can only enter the signal region if at least one light lepton is non-prompt or has its charge misidentified. It is obtained from MC simulation studies that the non-prompt lepton arises nearly always from B-meson decays and that around half of the τ_{had} candidates in these events come from $W \rightarrow \tau\nu$ decays, and the remaining fake ones from misidentified light-quark or gluon jets. For the data-driven fake estimation, sidebands are constructed by requiring a sideband lepton instead of a tight lepton and by using a low jet multiplicity region: $2 \leq n_{\mathrm{jet}} \leq 3$. For the non-prompt sideband
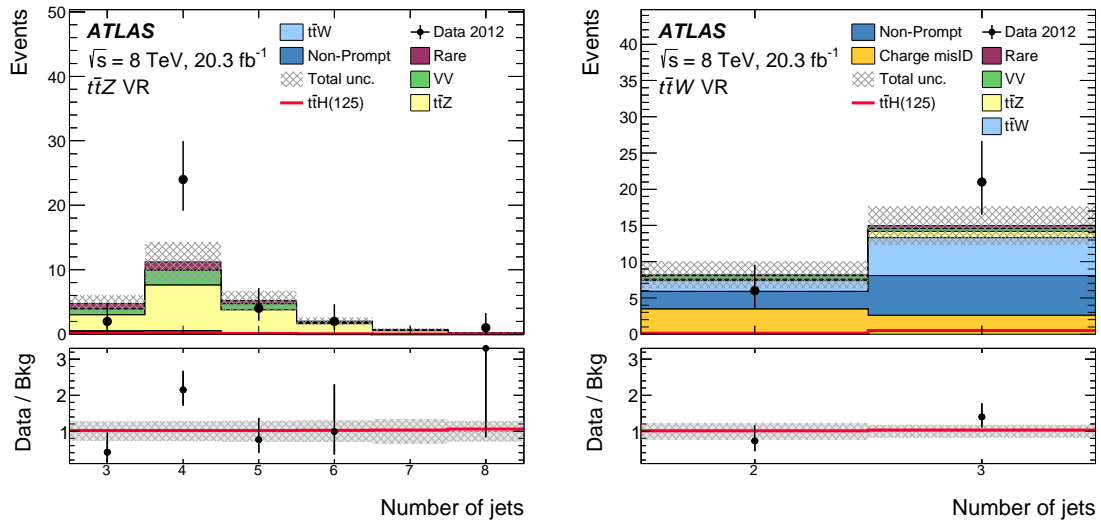
Figure 2: The spectrum of the number of jets expected and observed in the $t\bar{t}Z$ (left) and $t\bar{t}W$ (right) validation regions (VR). The hatched band represents the total uncertainty on the background prediction in each bin. The "non-prompt" backgrounds are those with a lepton arising from a hadron decay or from a photon conversion in detector material. Rare processes include tZ, $t\bar{t}WW$, triboson, $t\bar{t}t\bar{t}$, and tH production. The overlaid red line corresponds to the $t\bar{t}H$ signal predicted by the SM [9].

leptons, the E_T isolation requirement is inverted, the p_T isolation requirement is relaxed, and the electron identification is relaxed. The low jet multiplicity region is used to determine a transfer factor from sideband to tight lepton selections. The expected non-prompt lepton yield in the signal region is obtained by using this transfer factor to extrapolate from a control region with the same jet selection as the signal region but with one tight and one sideband light lepton. Fake events with charge mis-identified electrons are included in the transfer factor and the method. Since the ratio of real to fake τ_{had} candidates is similar in the signal and all control regions, the method is expected to correctly reproduce the fake/real τ_{had} composition in the signal region. A small correction factor, $k_{corr} = 0.96$, due to the minor variations in the ratio in the control regions is calculated and applied to the total estimate in the signal region. The overall uncertainty on the non-prompt background prediction is dominated by the statistical uncertainty of the high jet multiplicity sideband region. The overall obtained yields and uncertainties for all sub-channels are summarized in Table 1. For most backgrounds the uncertainties from Monte Carlo simulation sample sizes are negligible. For the diboson backgrounds, however, these can reach 50% of the total diboson yield uncertainties. Further systematic uncertainties of the analysis can be found in Ref. [9].

4. Results

The expected yields for backgrounds, $t\bar{t}H$ signal and data observation are shown in Table 1. The distributions of the number of jets in the signal region are shown in Fig. 3. The best-fit value of the signal strength $\mu = \sigma_{t\bar{t}H,\text{obs}}/\sigma_{t\bar{t}H,\text{SM}}$ is determined using a maximum likelihood fit to the data yields of the categories listed in Table 1, which are treated as independent Poisson terms in
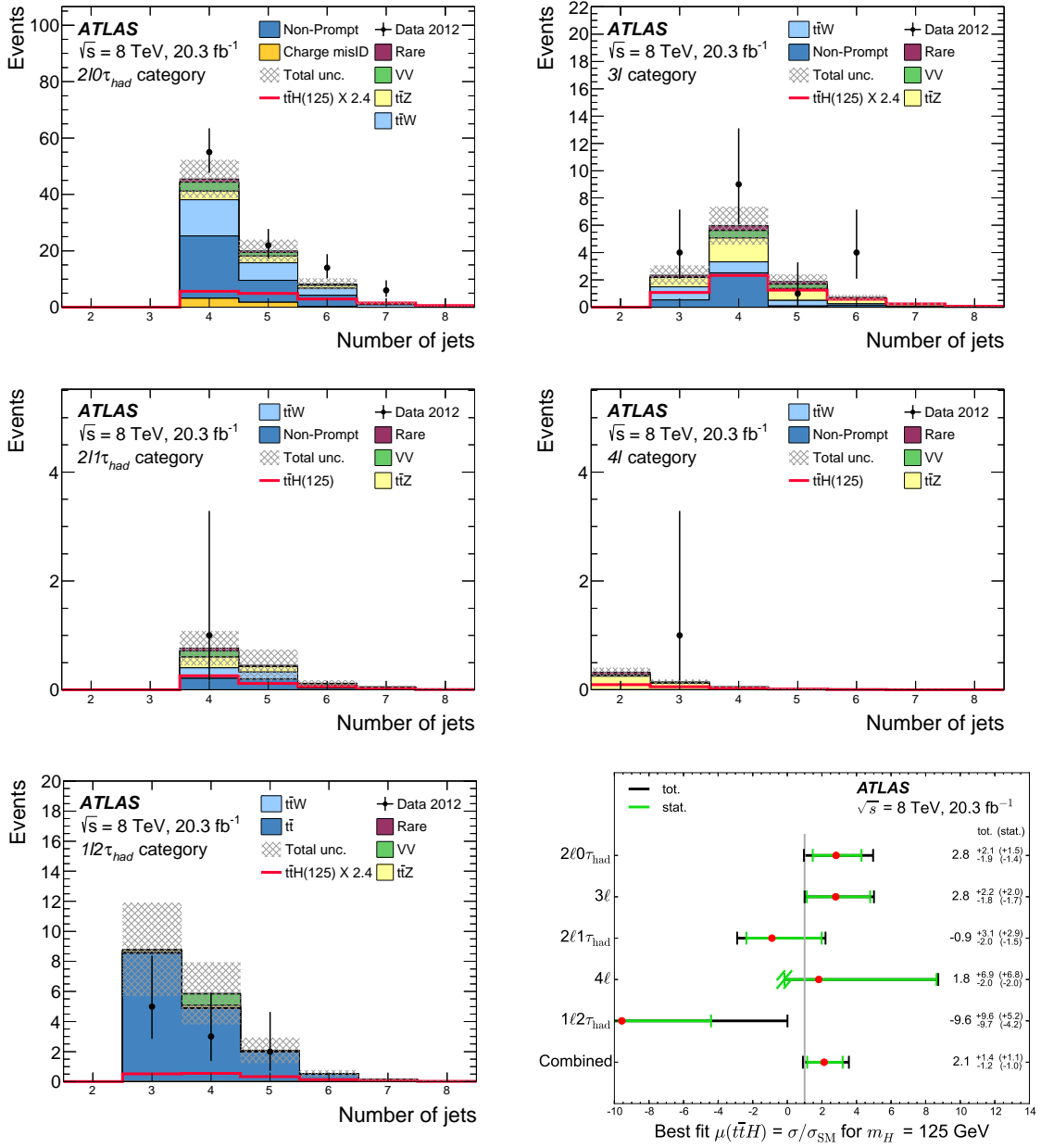
Figure 3: The spectrum of the number of jets expected and observed in each signal region as well as the best-fit values of the signal strength parameter. For display purposes the six $2\ell0\tau_{had}$ categories ($ee/e\mu/\mu\mu$ and $= 4/\geq 5$ jets) are combined into one plot, as are the two 4ℓ categories (Z-enriched and Z-depleted). The hatched bands show the total uncertainty on the background prediction in each bin. The non-prompt and charge mis-id background spectra are taken from simulation of $t\bar{t}$, single top, $Z \rightarrow \ell^+\ell^-$+jets, and other small backgrounds, with normalization as described in the text (in particular the $= 4/\geq 5$ jet regions of the $2\ell0\tau_{had}$ plot have the ratio given by the data-driven prediction). The overlaid red line shows the $t\bar{t}H$ signal from the Standard Model. For visibility, the $t\bar{t}H$ signal is multiplied by a factor of 2.4 in the $2\ell0\tau_{had}$, 3ℓ, and $1\ell2\tau_{had}$ plots. The lower right figure shows the best-fit values of the signal strength parameter $\mu = \sigma_{t\bar{t}H,obs}/\sigma_{t\bar{t}H,SM}$. For the 4ℓ Z-depleted category, $\mu < -0.17$ results in a negative expected total yield and so the lower uncertainty is truncated at this point [9].

Table 1: Expected and observed yields in each channel. Uncertainties shown are the sum in quadrature of systematic uncertainties and Monte Carlo simulation statistical uncertainties. "Non-prompt" includes the misidentified τ_{had} background to the $1\ell 2\tau_{\text{had}}$ category. Rare processes (tZ, $t\bar{t}WW$, triboson production, $t\bar{t}t\bar{t}$, tH) are not shown as a separate column but are included in the total expected background estimate [9].

Category	q mis-id	Non-prompt	$t\bar{t}W$	$t\bar{t}Z$	Diboson	Expected bkg.	$t\bar{t}H$ ($\mu=1$)	Observed
$ee + \geq 5j$	1.1 ± 0.5	2.3 ± 1.2	1.4 ± 0.4	0.98 ± 0.26	0.47 ± 0.29	6.5 ± 1.8	0.73 ± 0.14	10
$e\mu + \geq 5j$	0.85 ± 0.35	6.7 ± 2.4	4.8 ± 1.2	2.1 ± 0.5	0.38 ± 0.30	15 ± 3	2.13 ± 0.41	22
$\mu\mu + \geq 5j$	$-$	2.9 ± 1.4	3.8 ± 0.9	0.95 ± 0.25	0.69 ± 0.39	8.6 ± 2.2	1.41 ± 0.28	11
$ee + 4j$	1.8 ± 0.7	3.4 ± 1.7	2.0 ± 0.4	0.75 ± 0.20	0.74 ± 0.42	9.1 ± 2.1	0.44 ± 0.06	9
$e\mu + 4j$	1.4 ± 0.6	12 ± 4	6.2 ± 1.0	1.5 ± 0.3	1.9 ± 1.0	24 ± 5	1.16 ± 0.14	26
$\mu\mu + 4j$	$-$	6.3 ± 2.6	4.7 ± 0.9	0.80 ± 0.22	0.53 ± 0.30	12.7 ± 2.9	0.74 ± 0.10	20
3ℓ	$-$	3.2 ± 0.7	2.3 ± 0.7	3.9 ± 0.8	0.86 ± 0.55	11.4 ± 2.3	2.34 ± 0.35	18
$2\ell 1\tau_{\text{had}}$	$-$	$0.4^{+0.6}_{-0.4}$	0.38 ± 0.12	0.37 ± 0.08	0.12 ± 0.11	1.4 ± 0.6	0.47 ± 0.08	1
$1\ell 2\tau_{\text{had}}$	$-$	15 ± 5	0.17 ± 0.06	0.37 ± 0.09	0.41 ± 0.42	16 ± 5	0.68 ± 0.13	10
4ℓ Z-enr.	$-$	$\lesssim 10^{-3}$	$\lesssim 3 \times 10^{-3}$	0.43 ± 0.12	0.05 ± 0.02	0.55 ± 0.15	0.17 ± 0.02	1
4ℓ Z-dep.	$-$	$\lesssim 10^{-4}$	$\lesssim 10^{-3}$	0.002 ± 0.002	$\lesssim 2 \times 10^{-5}$	0.007 ± 0.005	0.025 ± 0.003	0

the likelihood. The $\mu = 1$ hypothesis assumes Standard Model Higgs boson production and decay with $m_H = 125\,$GeV. For all other values of μ only the $t\bar{t}H$ production cross section is scaled while the Higgs boson branching fractions are fixed to their SM values.

The results of the fit are shown in the lower right plot of Fig. 3. The results are compatible with the Standard Model expectation and with previous searches for $t\bar{t}H$ production in multilepton final states [10]. Combined over all categories, the value of μ is found to be $2.1^{+1.4}_{-1.2}$. In the presence of a signal of SM strength, the combined fit is expected to return $\mu = 1.0^{+1.2}_{-1.1}$. The $\mu = 0$ hypothesis has an observed (expected) p-value of 0.037 (0.18), corresponding to 1.8σ (0.9σ). The $\mu = 1$ hypothesis (the SM) has an observed p-value of 0.18, corresponding to 0.9σ. The observed (expected) 95% confidence level (CL) upper limit, combining all channels, is $\mu < 4.7$ (2.4). More information about limit construction can be found in Ref. [9].

These results are sensitive to $t\bar{t}W$ and $t\bar{t}Z$ production and use theoretical SM predictions for their cross sections as experimental measurements are still statistically limited. The best-fit μ value as a function of these cross sections is [9]:

$$\mu(t\bar{t}H) = 2.1 - 1.4 \left(\frac{\sigma(t\bar{t}W)}{232\,\text{fb}} - 1 \right) - 1.3 \left(\frac{\sigma(t\bar{t}Z)}{206\,\text{fb}} - 1 \right).$$

An event display of the observed data event in the signal region of the 2l+1τ_{had} category is illustrated in Fig. 4.

5. Conclusions

A search for $t\bar{t}H$ production in multilepton final states has been performed using 20.3 fb^{-1} of proton–proton collision data at $\sqrt{s} = 8$ TeV recorded by the ATLAS experiment at the LHC. The best-fit value of the ratio μ of the observed production rate to that predicted by the Standard Model is $2.1^{+1.4}_{-1.2}$. This result is consistent with the Standard Model expectation. A 95% confidence level limit of $\mu < 4.7$ is set. The expected limit in the absence of $t\bar{t}H$ signal is $\mu < 2.4$. The observed (expected) p-value of the no-signal hypothesis corresponds to 1.8σ (0.9σ).
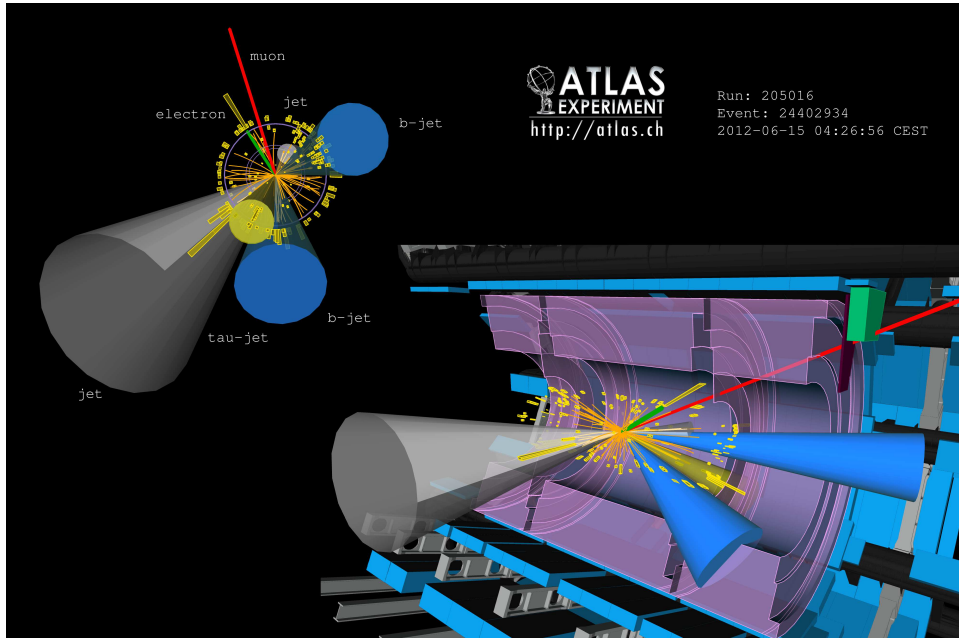
Figure 4: Event display of the observed data event in the signal region of the 2l+1τ_{had} category. Visible are the reconstructed electron, muon, the 2 light jets, the 2 b-tagged jets and the τ_{had} candidate [9].

References

[1] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[2] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[3] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[4] G. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[6] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 S08003 (2008)

[7] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].

[8] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 S08004 (2008)

[9] 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. B **749** (2015) 519, arXiv:1506.05988 [hep-ex].

[10] CMS Collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP **1409** (2014) 087, (erratum *ibid* **1410** (2014), 106), arXiv:1408.1682 [hep-ex].