

## CHAPTER 7

---

# Results

---

Results of the  $H \rightarrow \tau\tau$  analysis are discussed. This is described in additional detail in the recent ATLAS  $H \rightarrow \tau\tau$  publication [2]. The description of the fit procedure draws heavily from the recent ATLAS  $H \rightarrow WW^*$  publication [94].

The yields of signal and background processes are extracted from a statistical analysis of the data and the predictions described in Chapters 5 and 6. A likelihood function is maximized which provides the “best fit” of the prediction to the data in each bin of the BDT discriminators within their prescribed uncertainties.

The likelihood contains the data and prediction from six signal regions: three final states and two categories. Control regions are also used to constrain some background uncertainties. For example, a region of data rich in top events is used in the  $\tau_\ell\tau_{\text{had}}$  analysis to constrain the normalization on the portion of the top background predicted with simulation.

### 7.1 Fit procedure

The statistical analysis maximizes a likelihood function  $\mathcal{L}$  which depends on the signal strength  $\mu$ , the set of uncertainties (or nuisance parameters)  $\theta_i$ , and the observed number of data events  $N$ . The signal strength is defined as the measured  $H \rightarrow \tau\tau$  cross-section divided by the predicted  $H \rightarrow \tau\tau$  cross-section. A measured  $\mu = 1$  would imply exactly the same number of  $H \rightarrow \tau\tau$  events are observed as expected.

#### 7.1.1 Likelihood function

The likelihood function is shown in Fig. 7.1. It is the product of four terms.

$$\mathcal{L} = \underbrace{\prod_{i,b} f(N_{ib} \mid \mu \cdot S_{ib} \cdot \prod_r^{\text{Syst.}} \nu_{br}(\theta_r) + \sum_k \beta_k \cdot B_{kib} \cdot \prod_s^{\text{Syst.}} \nu_{bs}(\theta_s))}_{\text{Poisson for SR with signal strength } \mu; \text{ predictions } S, B} \cdot \underbrace{\prod_l f(N_l \mid \sum_k \beta_k \cdot B_{kl})}_{\text{Poisson for profiled CRs}} \cdot \underbrace{\prod_t^{\text{Syst. in } \{r,s\}} g(\vartheta_t \mid \theta_t)}_{\text{Gauss. for syst.}} \cdot \underbrace{\prod_k f(\xi_k \mid \zeta_k \cdot \theta_k)}_{\text{Poiss. for MC stats}}$$

Figure 7.1: The likelihood equation considered for maximization [94].

The first term of the likelihood is a Poisson function, in each bin  $b$  of each category  $i$ , describing the probability of observing  $N$  events given the predicted number of signal events  $S$  and background events  $B$ . The number of signal events is scaled by the signal strength  $\mu$  which is shared among bins and categories. The number of background events is scaled by a parameter  $\beta_k$  which can be category-dependent. The signal and background predictions are modified by response functions  $\nu(\theta)$  for each nuisance parameter  $\theta$ .

The response function for bin-wise systematic uncertainties is given by  $\nu(\theta) = 1 + \theta \cdot \frac{\Delta_\theta}{N}$ , where  $\Delta_\theta$  is the given uncertainty on  $\theta$  in units of  $N$ , i.e.,  $N \cdot \nu(\theta = \pm 1) = N \pm \Delta_\theta$  for  $N = S, B$ . The response function for statistical uncertainties on the predictions is  $\nu(\theta) = \theta$ .

The second term is a Poisson function, in each control region  $l$ , describing the probability of observing  $N$  events given the predicted number of background events  $B$ . The number of background events is scaled by a parameter  $\beta_k$  which is shared with the first term. The statistical power of the control region data is then used to constrain the background normalization in the signal region.

The third term is a Gaussian function, for each systematic uncertainty, describing the probability of a given nuisance parameter being pulled away from its nominal value. This is a penalty term for choosing a less likely value of the nuisance parameter than the nominal value. The choice of Gaussian function to describe the uncertainty is convention.

The fourth term is a Poisson function, in each bin of each category, describing the uncertainty due to finite sample size of the signal and background predictions. This is especially relevant for predictions which have natural statistical limitations, e.g., predicting the  $Z \rightarrow \tau\tau$  background with  $Z \rightarrow \mu\mu$  events from data.

To extract the observed  $\mu$  and other quantities, the likelihood is maximized with respect to  $\mu$  and the associated nuisance parameters  $\theta$ . This maximization yields the most probable  $\mu$ . The likelihood is evaluated at  $\vartheta = 0$  and  $\xi = \zeta$  which imposes that the most probable value of each systematic nuisance parameter is zero and statistical nuisance parameter is one.

### 7.1.2 Features

The first term of the likelihood is maximized when the prediction exactly matches the observed data events in a given bin. The prediction is allowed to change within the maximization as the nuisance parameters are pulled. This allowed change is restricted by the Gaussian constraints imposed on these uncertainties, which are maximized when the nuisance parameters are unchanged from their pre-fit prediction. For example, the jet energy scale is allowed to be moved by  $+1\sigma$ , but this incurs a penalty in the likelihood of  $\frac{g(1|0,1)}{g(0|0,1)} \approx 0.6$ .

The normalization of the embedded  $Z \rightarrow \tau\tau$  is allowed to float freely within each final state, i.e., there is no constraint beyond the signal region statistics. In the  $\tau_\ell\tau_{\text{had}}$  channel, a single normalization nuisance parameter is used for  $Z \rightarrow \tau\tau$  which is shared in each BDT bin of each category. This is sensible because the defining features of the categories – dijet kinematics and  $p_T^Z$  – are expected to be well predicted by the embedding method.

The likelihood maximization extracts many fitted parameters, including the signal strength  $\mu$ , the pulls of the nuisance parameters  $\theta_i$ , and the uncertainties on the nuisance parameters in units of the pre-fit uncertainties (i.e.,  $\Delta_{\theta, \text{pre-fit}} = 1$ ).

### 7.1.3 Test statistic

The test statistic is defined as:

$$q(\mu) = -2 \ln \left( \frac{\mathcal{L}(\mu, \theta_i)}{\mathcal{L}_{\text{max.}}} \right) \bigg|_{\theta_i = \hat{\theta}_{i, \mu}} \quad (7.1)$$

and is evaluated as a function of  $\mu$ . The denominator  $\mathcal{L}_{\text{max.}}$  is the unconditional maximum of the likelihood as a function of  $\mu$  and the nuisance parameters  $\theta_i$  and is therefore just a number. The numerator  $\mathcal{L}(\mu, \theta_i)$  is the conditional maximum for a given  $\mu$  as a function of the nuisance parameters, thereby asking which signal strength is most likely.

The fitted uncertainty on a nuisance parameter is defined as the distance  $\theta_{\text{right}} - \theta_{\text{left}}$  such that  $q(\mu) = 1$  when scanning a given nuisance parameter to the right and left of the fitted  $\theta$ . If  $\mathcal{L}$  follows a Gaussian distribution, the integral from  $\theta_{\text{left}}$  to  $\theta_{\text{right}}$  corresponds to 68% ( $1\sigma$ ) of the distribution.

### 7.1.4 Impact of uncertainties on $\mu$

The post-fit *impact* of a given nuisance parameter on the signal strength is defined as:

$$\Delta_{\hat{\mu}, i} = \hat{\mu} \left( \hat{\theta}_i \pm \Delta_{\hat{\theta}_i} \right) - \hat{\mu} \quad (7.2)$$

where hats indicate post-fit values.  $\hat{\mu}(\hat{\theta}_i \pm \Delta_{\hat{\theta}_i})$  indicates the fitted  $\mu$  of the full fit but with  $\theta_i$  fixed to its post-fit value  $\hat{\theta}$  varied by the post-fit uncertainty  $\pm\Delta_{\hat{\theta},i}$ , where all other nuisance parameters are floating. If the fitted  $\mu$  is robust against changes to a nuisance parameter, its impact will be small.

## 7.2 Fit results

The data and fitted signal and background predictions are shown in Fig. 7.2. The majority of bins in the BDTs are dominated by background, and good modeling is observed. In the signal-like regime of BDT output near 1, the predicted contribution from  $H \rightarrow \tau\tau$  signal is visible in the VBF category of each final state, and the signal hypothesis is favored over the background-only hypothesis.

The yields in the VBF  $H \rightarrow \tau_\ell\tau_{\text{had}}$  category and in the most signal-like regime are shown in Table 7.1 after the global fit. In the highest BDT bin, the signal hypothesis is favored over the background-only hypothesis.

At the value of the Higgs boson mass obtained from the combination of the ATLAS  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^*$  measurements, the signal strength obtained is  $1.4 \pm 0.3$  (syst.)  $\pm 0.3$  (stat.)  $\pm 0.1$  (theo.). The signal strength breakdown by category and channel is shown in Fig. 7.3.

The probability  $p_0$  of obtaining a result at least as signal-like as observed in the data if no signal

Table 7.1: Data and the predicted yields of signal and background in the VBF  $\tau_\ell\tau_{\text{had}}$  category after the global fit.

Process/Category	VBF $\tau_\ell\tau_{\text{had}}$		
	All bins	Second to last bin	Last bin
BDT output bin			
$j \rightarrow \tau_{\text{had}}$	$1680 \pm 50$	$8.2 \pm 0.9$	$5.2 \pm 0.7$
$Z \rightarrow \tau\tau$	$877 \pm 29$	$7.6 \pm 0.9$	$4.2 \pm 0.7$
Top	$82 \pm 15$	$0.3 \pm 0.4$	$0.5 \pm 0.4$
$Z \rightarrow \ell\ell$ ( $\ell \rightarrow \tau_{\text{had}}$ )	$54 \pm 26$	$1.0 \pm 0.7$	$0.30 \pm 0.28$
Diboson	$63 \pm 11$	$1.0 \pm 0.4$	$0.48 \pm 0.20$
ggF $H \rightarrow \tau\tau$	$16 \pm 6$	$1.0 \pm 0.4$	$1.2 \pm 0.6$
VBF $H \rightarrow \tau\tau$	$31 \pm 8$	$4.5 \pm 1.1$	$9.1 \pm 2.2$
$WH \rightarrow \tau\tau$	$< 1$	$< 0.1$	$< 0.1$
$ZH \rightarrow \tau\tau$	$< 1$	$< 0.1$	$< 0.1$
Total background	$2760 \pm 40$	$18.1 \pm 2.3$	$10.7 \pm 2.7$
Total signal	$48 \pm 12$	$5.5 \pm 1.3$	$10.3 \pm 2.5$
Data	2830	22	21

## 7. RESULTS

---

were present is calculated using the test statistic  $q(0)$ . For  $m_H = 125.36$  GeV, the observed  $p_0$  is  $2.7 \times 10^{-6}$ , which corresponds to a deviation from the background-only hypothesis of  $4.5\sigma$ . This can be compared to an expected significance of  $3.4\sigma$ .

To emphasize the most signal-like regime, two distributions are built from the six categories used in the global fit and shown in Fig. 7.4. First, the signal region bins of the BDT discriminator are re-ordered by  $\log_{10}(S/B)$ , where  $S/B$  is the signal-to-background ratio calculated assuming  $\mu = 1.4$  in each bin. The expected signal yield for both  $\mu = 1$  and  $\mu = 1.4$  is shown, as well as the fitted background yield for the background-only hypothesis. The signal hypothesis is clearly favored. Second, the  $m_{\tau\tau}^{\text{MMC}}$  distribution is shown summed across all categories and weighted event-by-event by  $\ln(1 + S/B)$ , which enhances the events compatible with the signal hypothesis. The excess of events in this mass distribution is consistent with the expectation for a Standard Model Higgs boson with  $m_H = 125$  GeV.

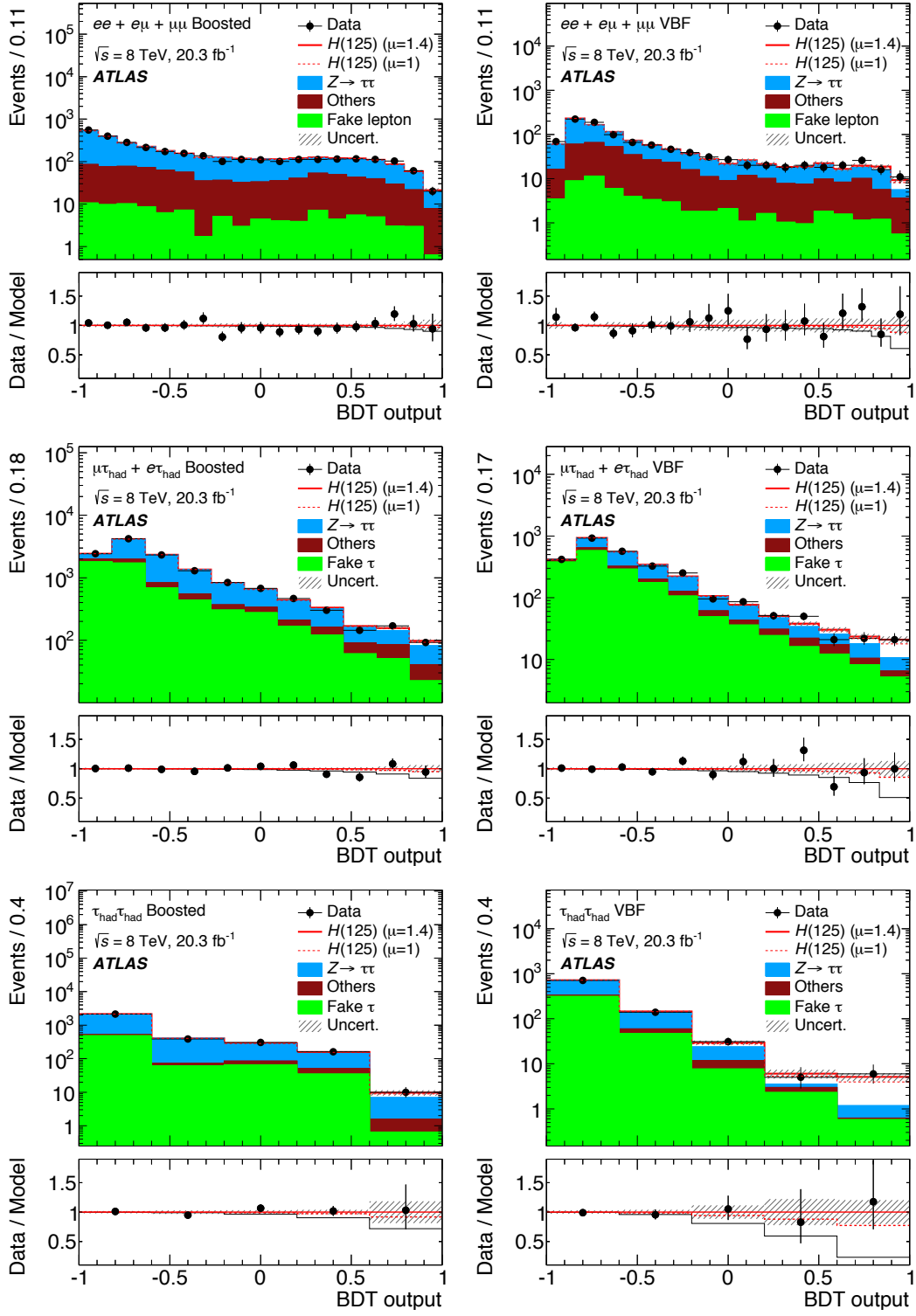
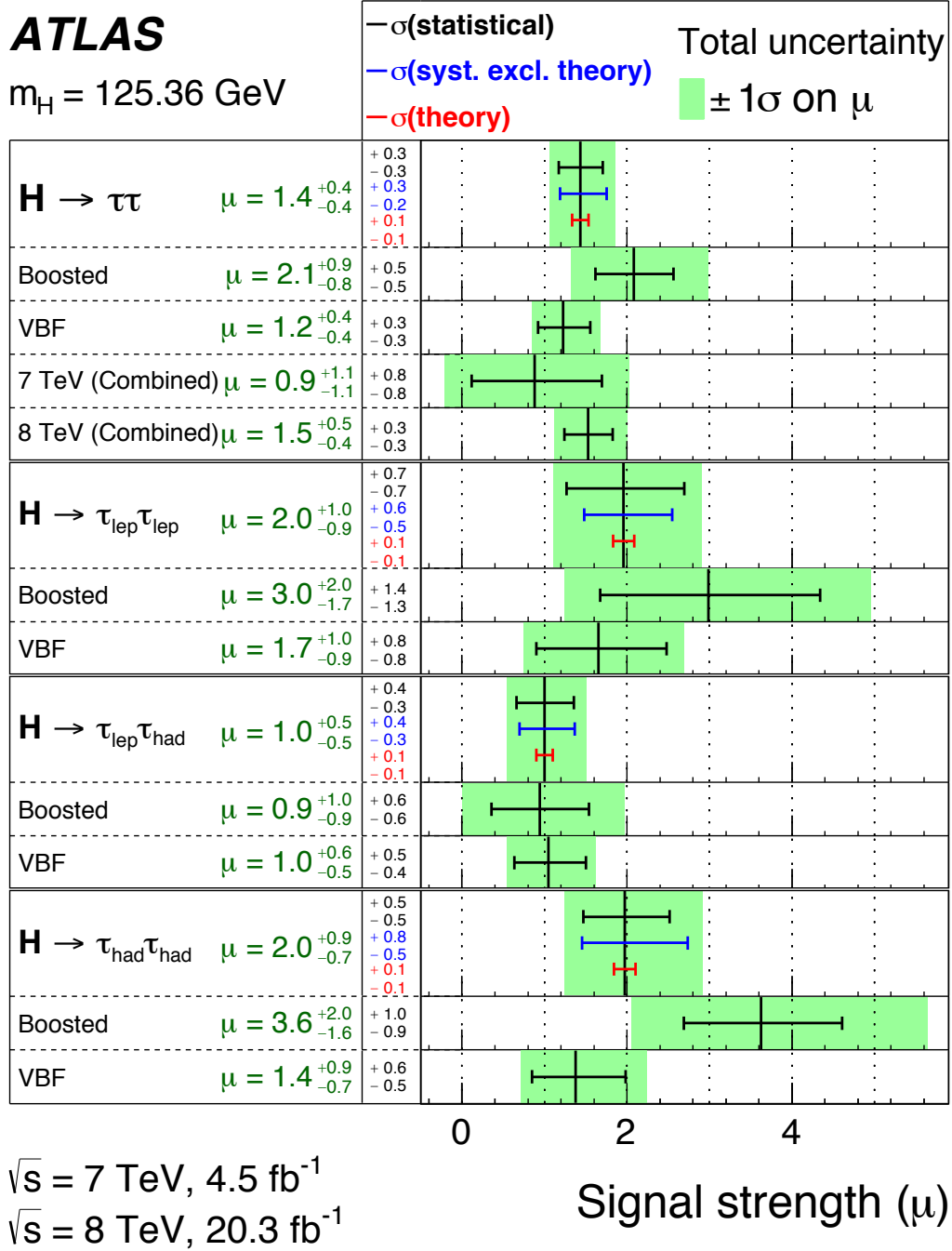


Figure 7.2: Distributions of the 8 TeV BDT discriminants in all six analysis categories after the global fit [2].

The dominant uncertainties on the measurement of the signal strength parameters include statistical uncertainties on the data from the signal regions, uncertainties on the jet and tau energy scales, uncertainties on the normalization of the  $Z \rightarrow \tau\tau$  and top backgrounds, and theoretical uncertainties. The contributions of each of these significant sources to the uncertainty of the measured signal strength are summarized in Fig. 7.5.

### 7.3 High score events in data

Event displays of some of the most signal-like events in data in the VBF  $H \rightarrow \tau_\ell \tau_{\text{had}}$  and  $H \rightarrow \tau_{\text{had}} \tau_{\text{had}}$  analyses are shown in Figs. 7.6 and 7.7.

Figure 7.3: The fitted signal strength  $\mu$  split by category, final state, and data-taking period [2].



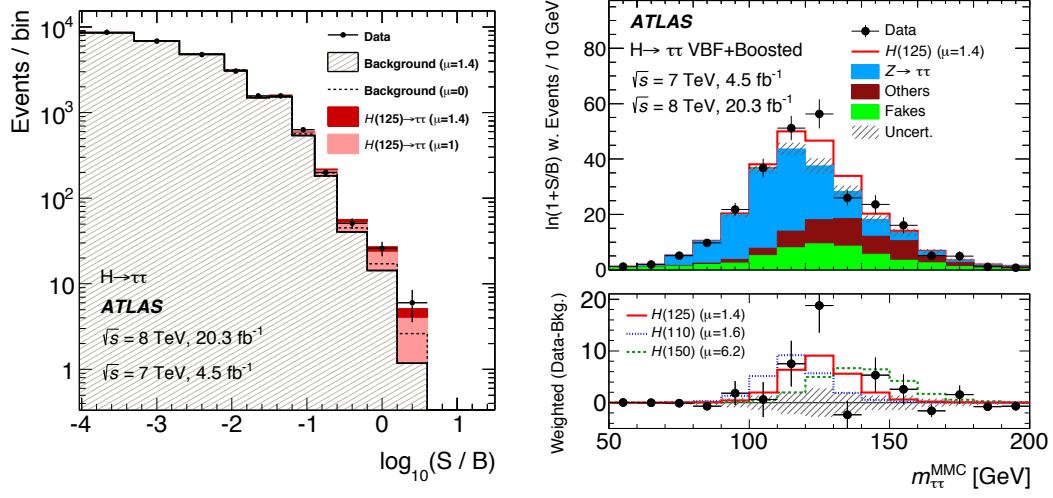


Figure 7.4: Plots of data and prediction which emphasize the most sensitive regions [2]. The individual BDT bins from all six categories are ordered by  $S/B$  and plotted on a shared axis (left) and entries in the  $m_{\tau\tau}^{\text{MMC}}$  distribution are weighted by  $\log(1 + S/B)$  (right).

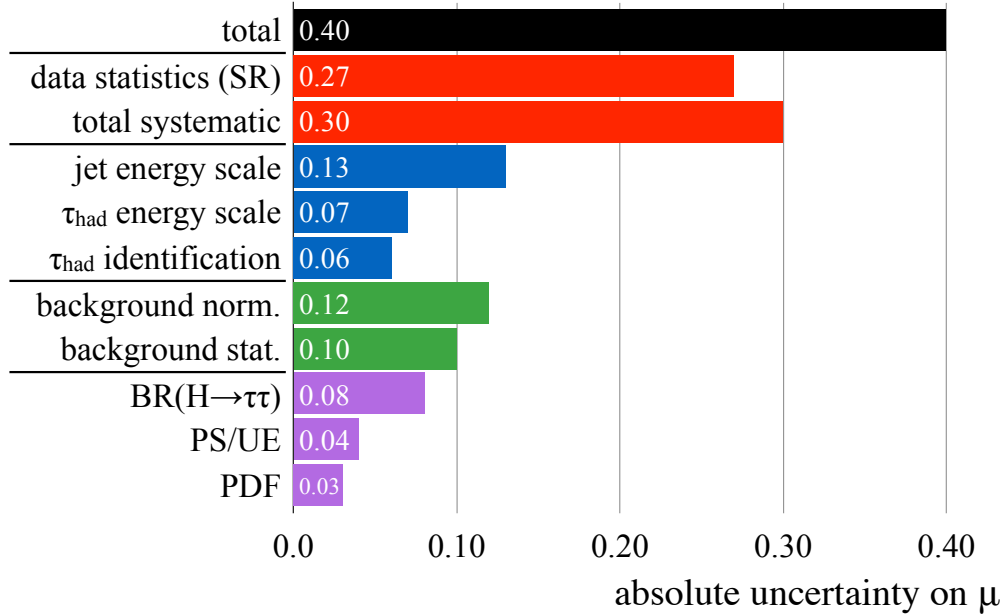


Figure 7.5: Comparison of the impact of the statistical and systematic uncertainties on the absolute uncertainty on  $\mu$  [2].

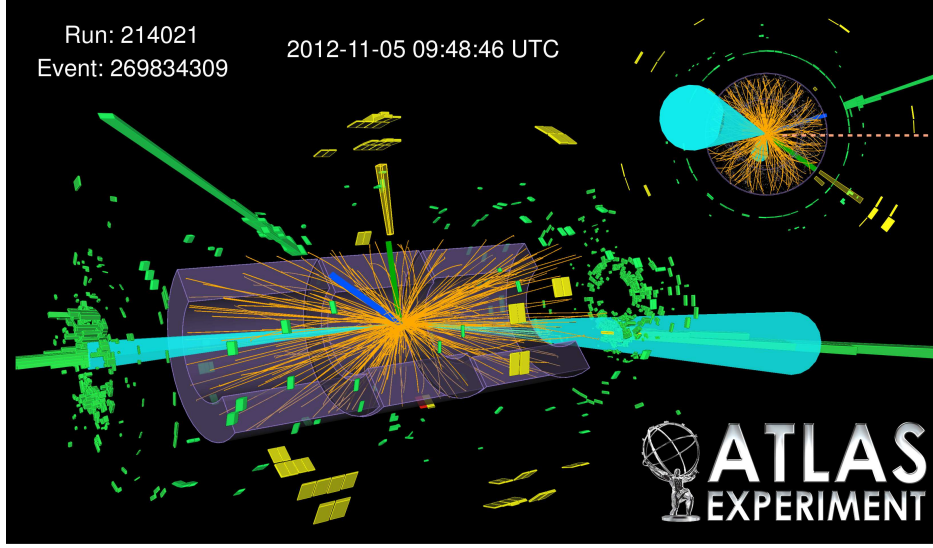


Figure 7.6: Display of one of the most signal-like events in the  $H \rightarrow \tau_\ell \tau_{\text{had}}$  VBF category in data [2]. The blue track matched to the green cluster indicates an electron, the green track matched to the yellow cluster indicates a  $\tau_{\text{had}}$ , the pink dotted line indicates the  $E_T^{\text{miss}}$  in the transverse plane, and the turquoise cones indicates the VBF jets. The reconstructed  $m_{\tau\tau}^{\text{MMC}} = 127$  GeV and  $m_{\text{jj}} = 1.53$  TeV.

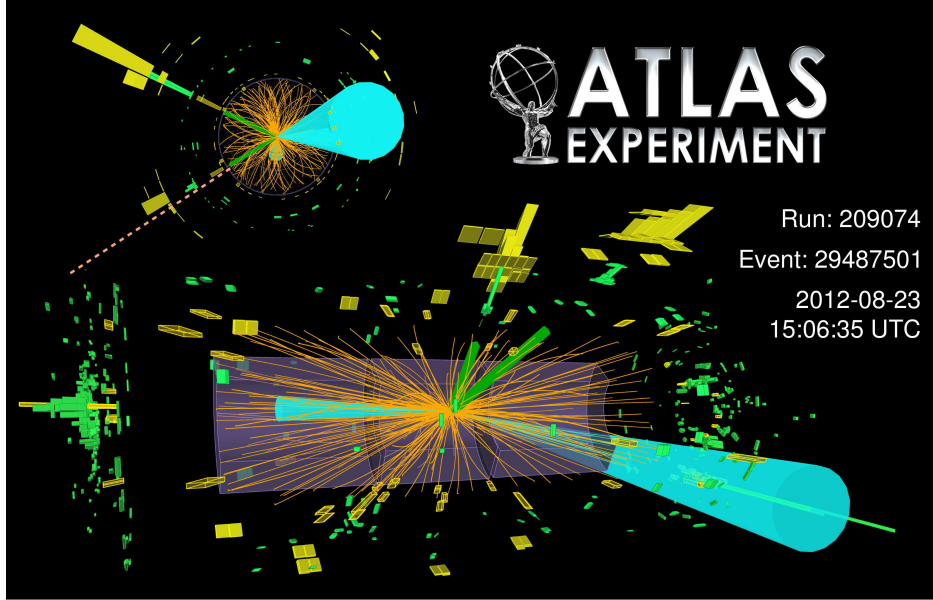


Figure 7.7: Display of one of the most signal-like events in the  $H \rightarrow \tau_{\text{had}} \tau_{\text{had}}$  VBF category in data [2]. The green tracks matched to the yellow clusters indicate the  $\tau_{\text{had}}$ , the pink dotted line indicates the  $E_T^{\text{miss}}$  in the transverse plane, and the turquoise cones indicates the VBF jets. The reconstructed  $m_{\tau\tau}^{\text{MMC}} = 123$  GeV and  $m_{\text{jj}} = 1.02$  TeV.