H→ τ τ at the LHC

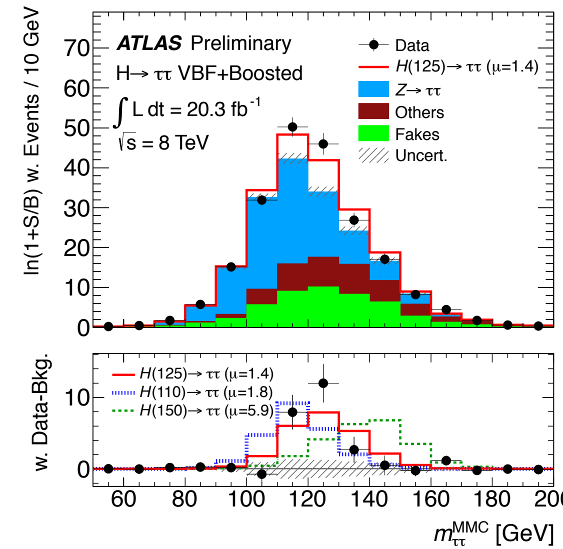
Higgs Maxwell Workshop
University of Edinburgh
26 February 2014

Katy Grimm

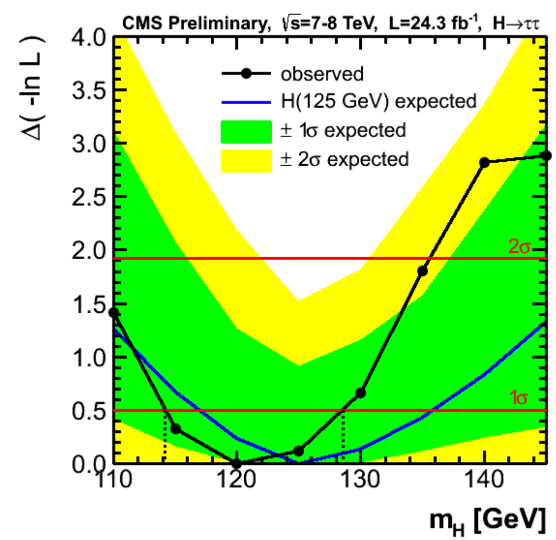
Lancaster University

Introduction: The Latest Evidence for SM BEH scalar boson

- **H → Fermions!**
- 26 November 2013 CERN Seminar:
  - ATLAS showed evidence for Htautau: 3.2 σ expected, 4.1 σ observed with 20.3 fb^{-1} of 8 TeV data
  - Conf Note ATLAS-CONF-2013-108

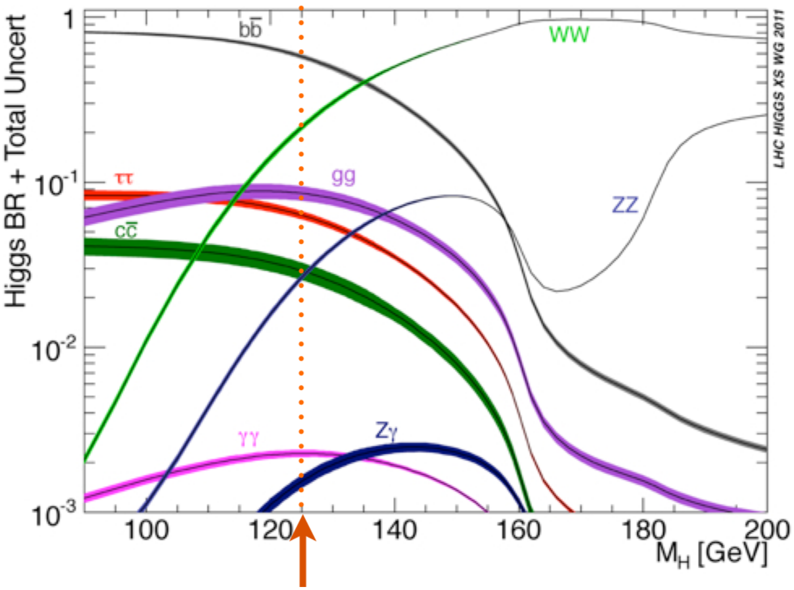


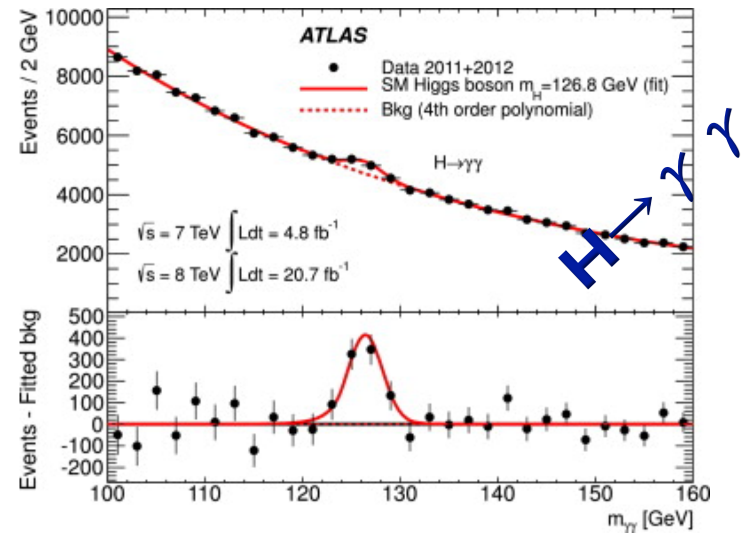
- 3 December 2013 CERN-LHC Seminar:
  - CMS showed evidence for Htautau 3.6 σ expected, 3.4 σ evidence. 19.7 fb-1 of 8 TeV + 4.9 of 7 TeV
  - Paper in arxiv on 20 Jan arXiv:1401.5041v1

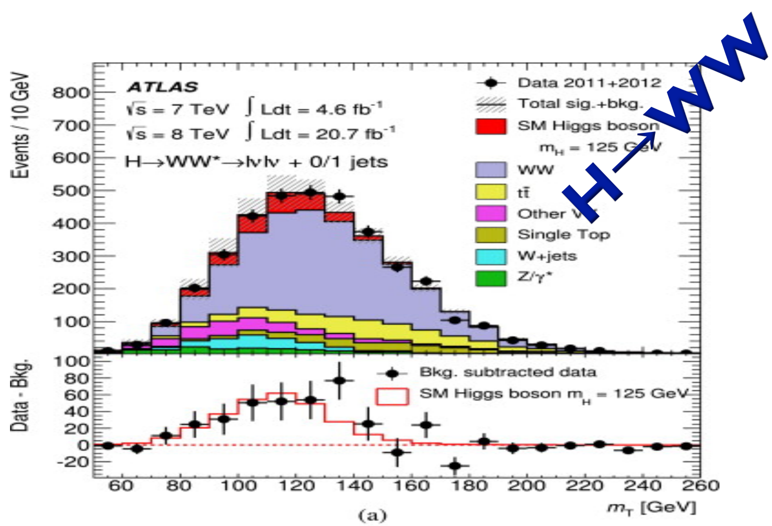


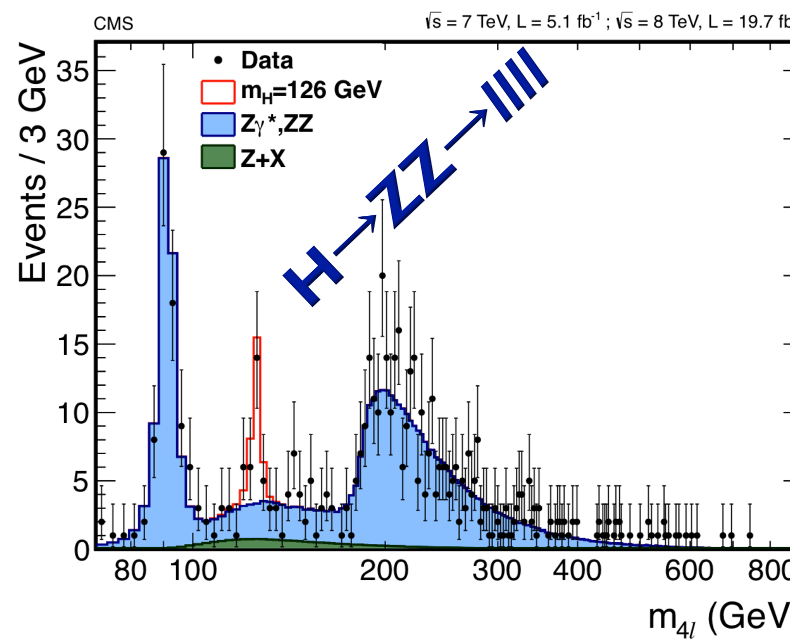
H Decay Modes, Searches at the LHC

H Branching Ratios



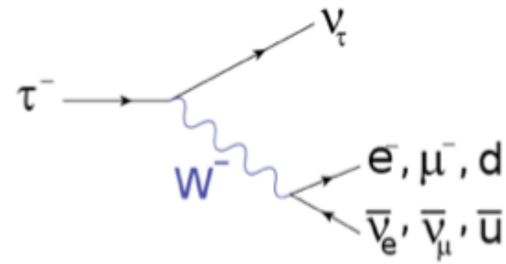




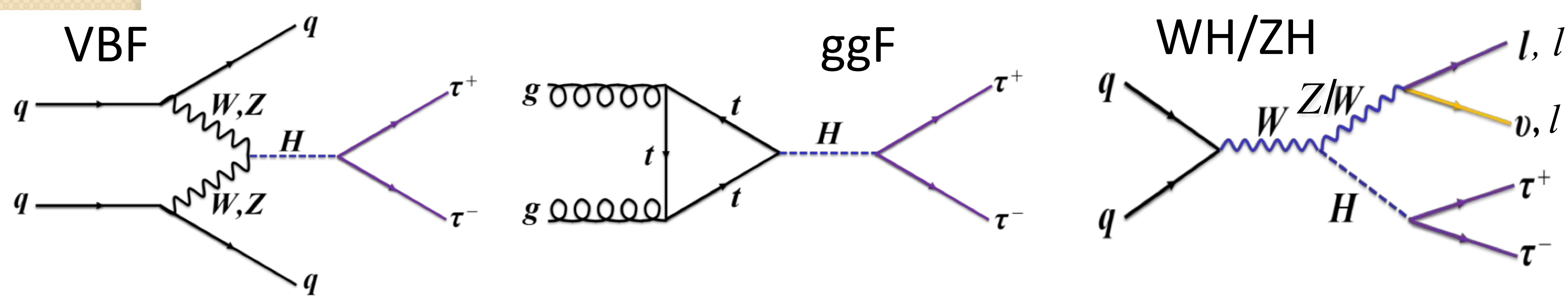


H→ $\tau\ \tau$ at ATLAS and CMS

- The three final states of H decay are:
  - H→ $\tau\ \tau$ →**lepton lepton + 4 ν, BR = 12.4%**
  - H→ $\tau\ \tau$ →**lepton hadron + 3 ν, BR = 45.6%**
  - H→ $\tau\ \tau$ →**hadron hadron + 2 ν, BR = 42%**



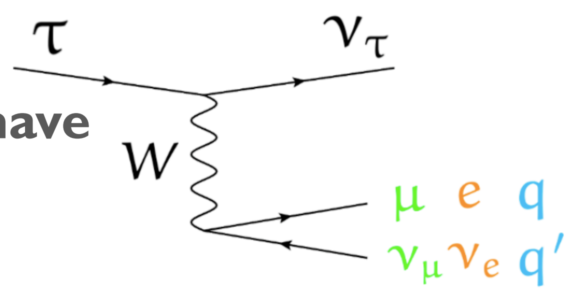
- In addition signatures based on H production are sought:



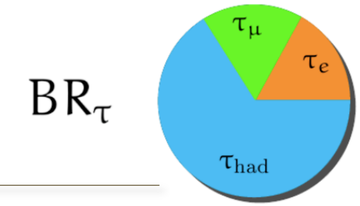
- CMS uses a series of cut-based categories to narrow in on the signatures, and ATLAS uses a Boosted Decision Tree

Tau Identification

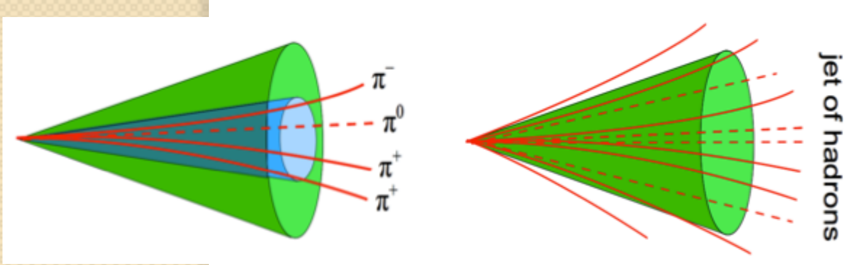
- m_τ = 1.78 GeV; Decay length = 87 μ m – taus almost always **decay within the beam pipe.**



- **Hadronic decays are highly collimated and have low track multiplicity– 1 or 3 charged pions**



- **TauID (and lepton suppression) uses Multivariate Techniques that take advantages of shower-shape and track information**



➢ **1 Prong** (BR = 49.5%):

Corresponds mostly to:

$$\tau^\pm \rightarrow \pi^\pm \nu_\tau$$

or to:

$$\tau^\pm \rightarrow \rho^\pm (\rightarrow \pi^0 \pi^\pm) \nu_\tau$$

➢ **3 Prong** (BR = 15.2%):

Corresponds mostly to:

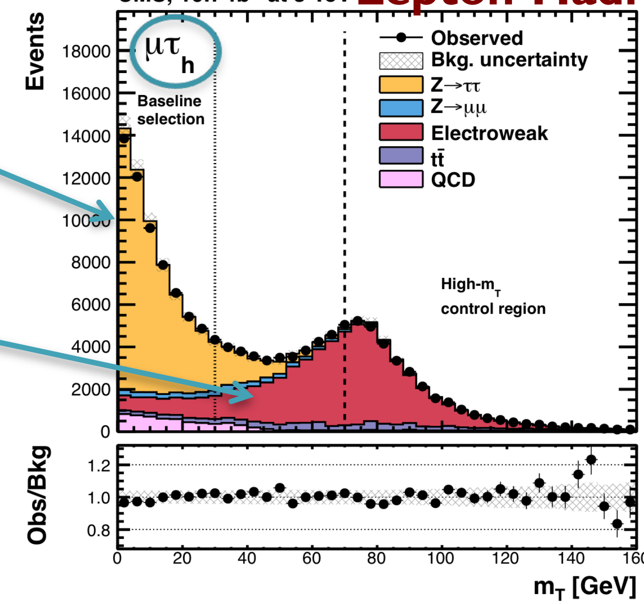
$$\tau^\pm \rightarrow a_1^\pm (\rightarrow \rho^0 \pi^\pm \rightarrow 3\pi^\pm) \nu_\tau$$

Backgrounds

Highly dependent on the tau channel:

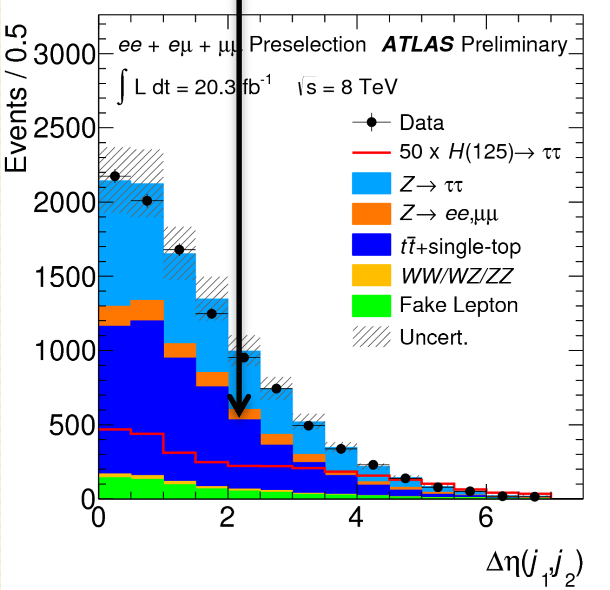
Ztautau is important for all channels

Lepton-Hadron

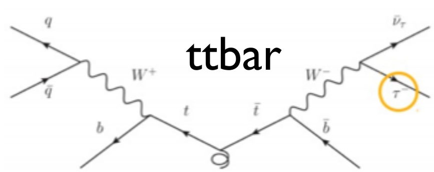


Electro-weak backgrounds including **W+jets, diboson, and top-quark** are significant for lepton+tau channel and lepton lepton channel

Lepton-Lepton

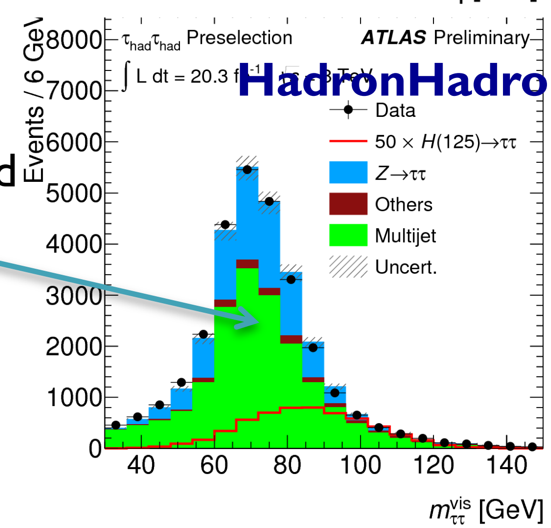


ttbar



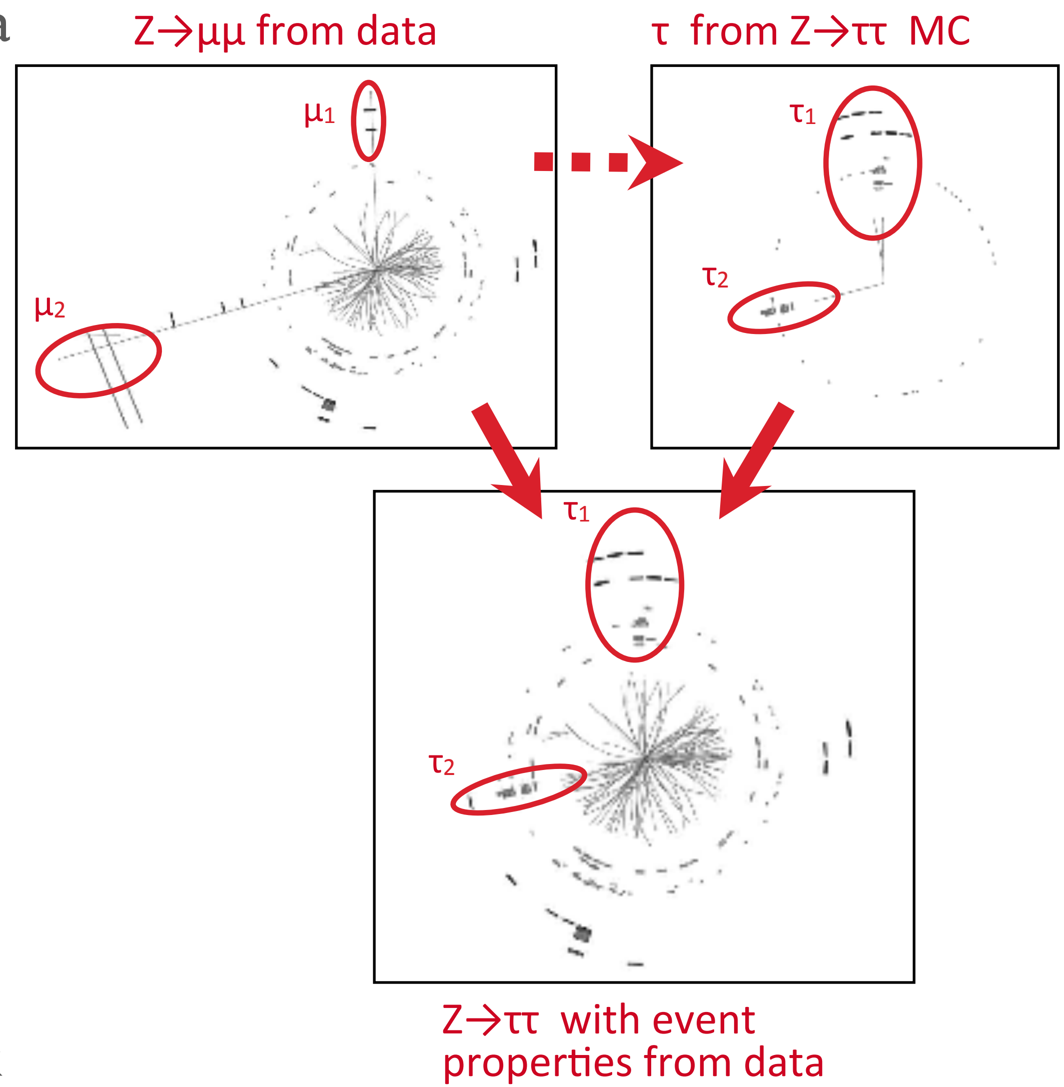
QCD multijet background is most significant for the $\tau_h\tau_h$ channel

HadronHadron



$Z \rightarrow \tau\ \tau$ Simulation: Embedding

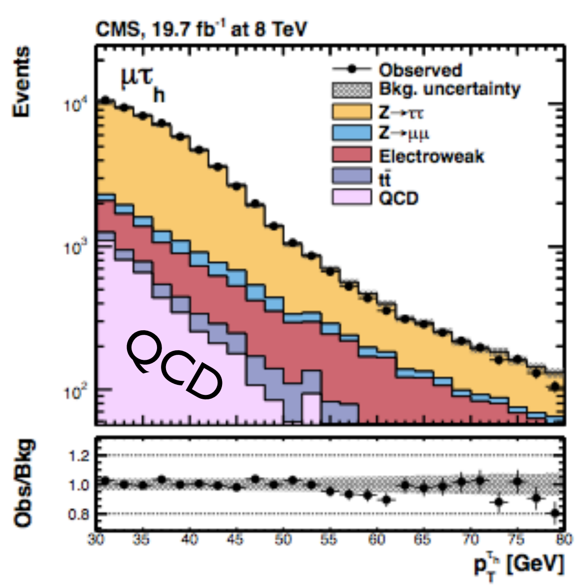
- Both ALTAS and CMS use Embedding: all properties of a $Z \rightarrow \tau\ \tau$ event except the taus are modeled by $Z \rightarrow \mu\ \mu$ data
- Remove μ from data
  - Simulate τ including spin
  - Add τ in place of the μ
- Major advantages:
  - Directly model with data: Z-boson kinematics, jets, MET resolution, pile-up, and VBF/EWK production
  - MC would not have a signal-free Ztautau region to check modeling.



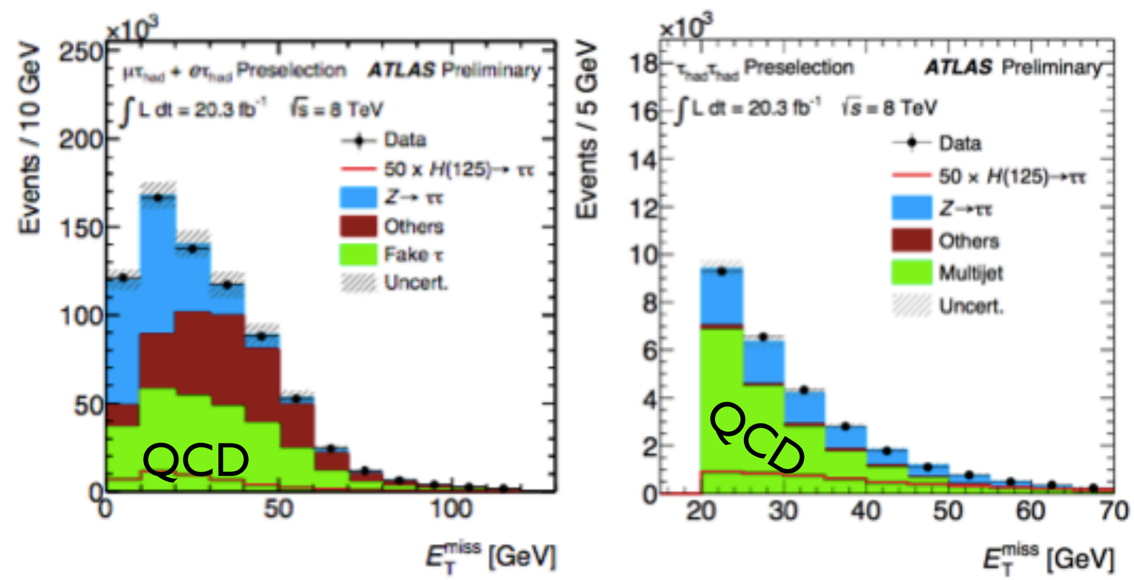
Modeling Multijet Background

- QCD Multijet background is estimated directly from data both at CMS and ATLAS
- The shape is modeled using a sample of one of the following
  - taus/leptons that fail the isolation requirement,
  - taus that fail ID requirements, or
  - taus that have the same charge

CMS Tau pT

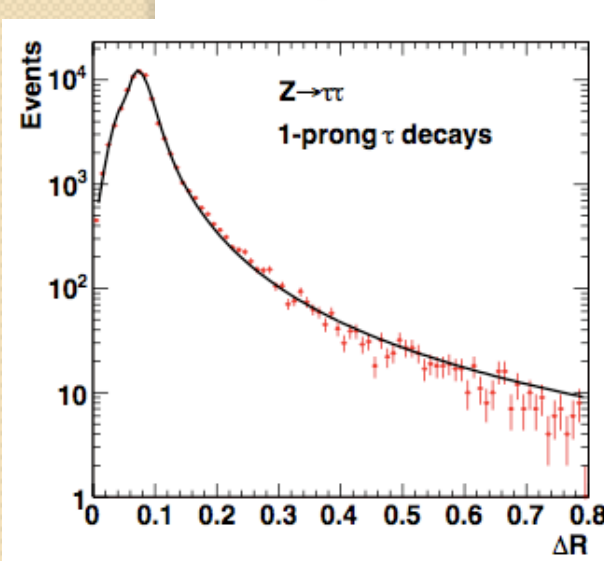


ATLAS MET (lephad, hadhad)

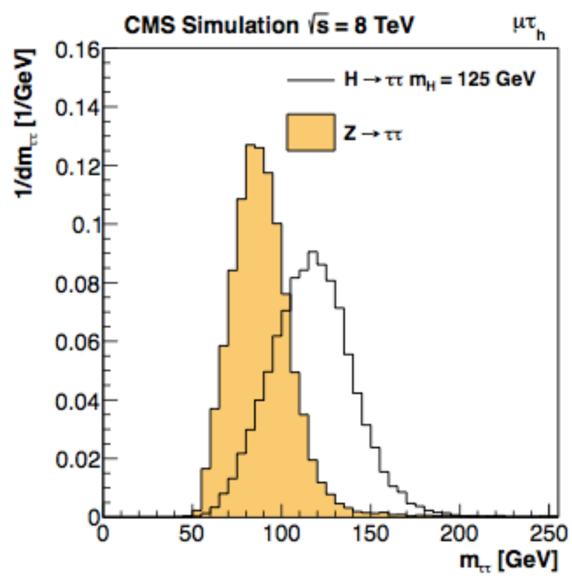


Reconstructing di-tau Mass

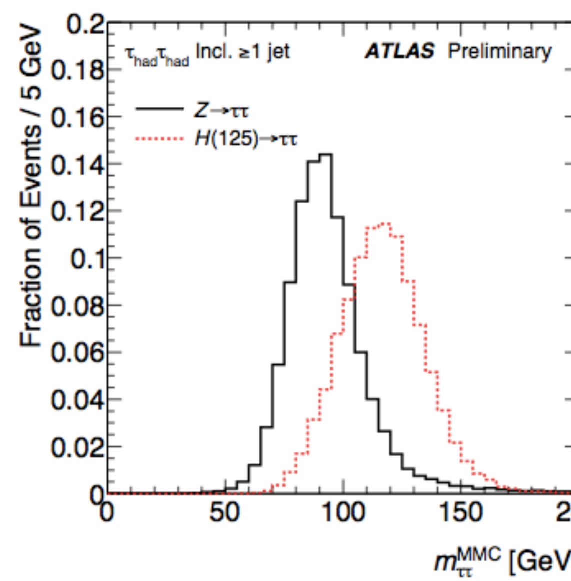
- Good $m_{\tau\tau}$ resolution provides separation between H and $Z\rightarrow\tau\tau$
- **Final state neutrinos are a problem. Ditau Mass is reconstructed using kinematic probabilities of neutrinos w.r.t. taus e.g.** dR($\tau\ \nu$)



$m_{\tau\tau}$ CMS



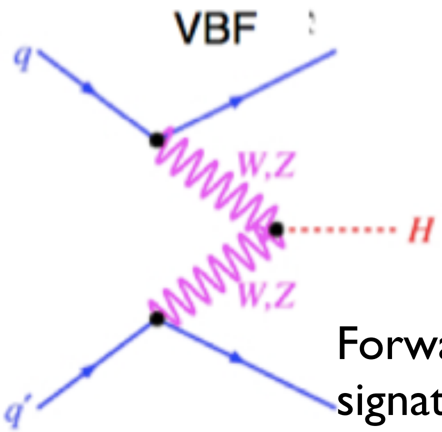
$m_{\tau\tau}$ ATLAS



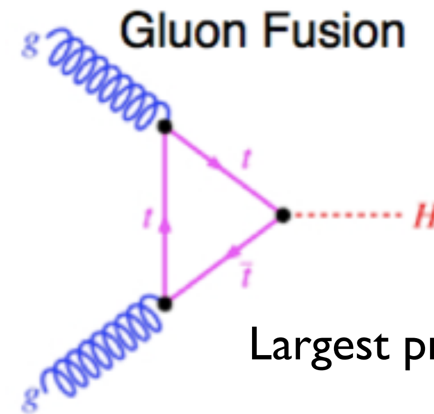
Mass Resolution varies between ~10%-20% depending on channel

ATLAS Signal Region Categories

- In ATLAS, sub-categories within each channel are created based on signal kinematics



Forward jet signature



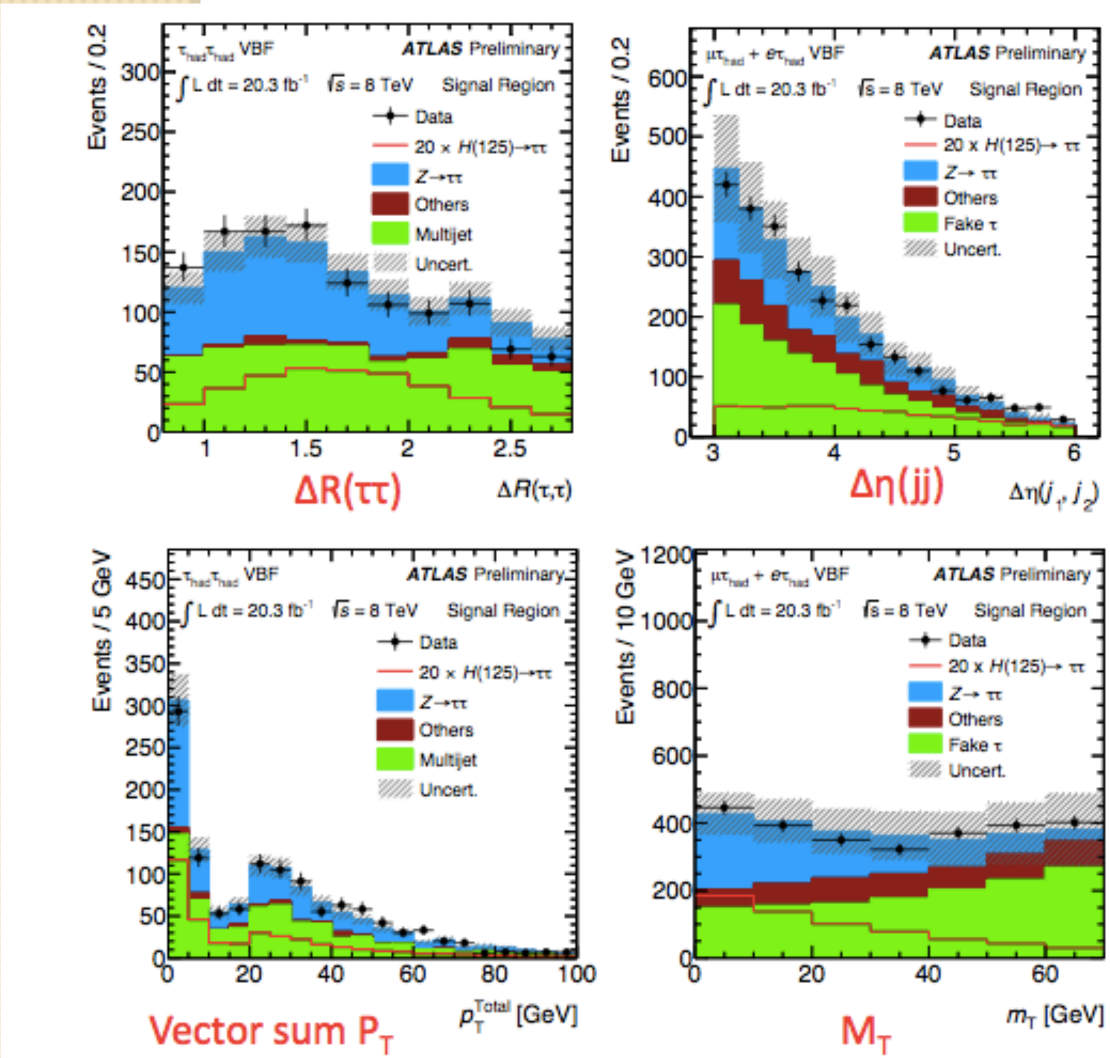
Largest production mode

ATLAS

Category	Selection	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$	$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$
VBF	$p_{\mathrm{T}}(j_1) >$ (GeV)	40	50	50
	$p_{\mathrm{T}}(j_2) >$ (GeV)	30	30	30/35
	$\Delta\eta(j_1, j_2) >$	2.2	3.0	2.0
	$b-$jet veto for jet $p_{\mathrm{T}} >$ (GeV)	25	30	-
	$p_{\mathrm{T}}^{H} >$ (GeV)	-	-	40
Boosted	$p_{\mathrm{T}}(j_1) >$ (GeV)	40	-	-
	$p_{\mathrm{T}}^{H} >$ (GeV)	100	100	100
	$b-$jet veto for jet $p_{\mathrm{T}} >$ (GeV)	25	30	-

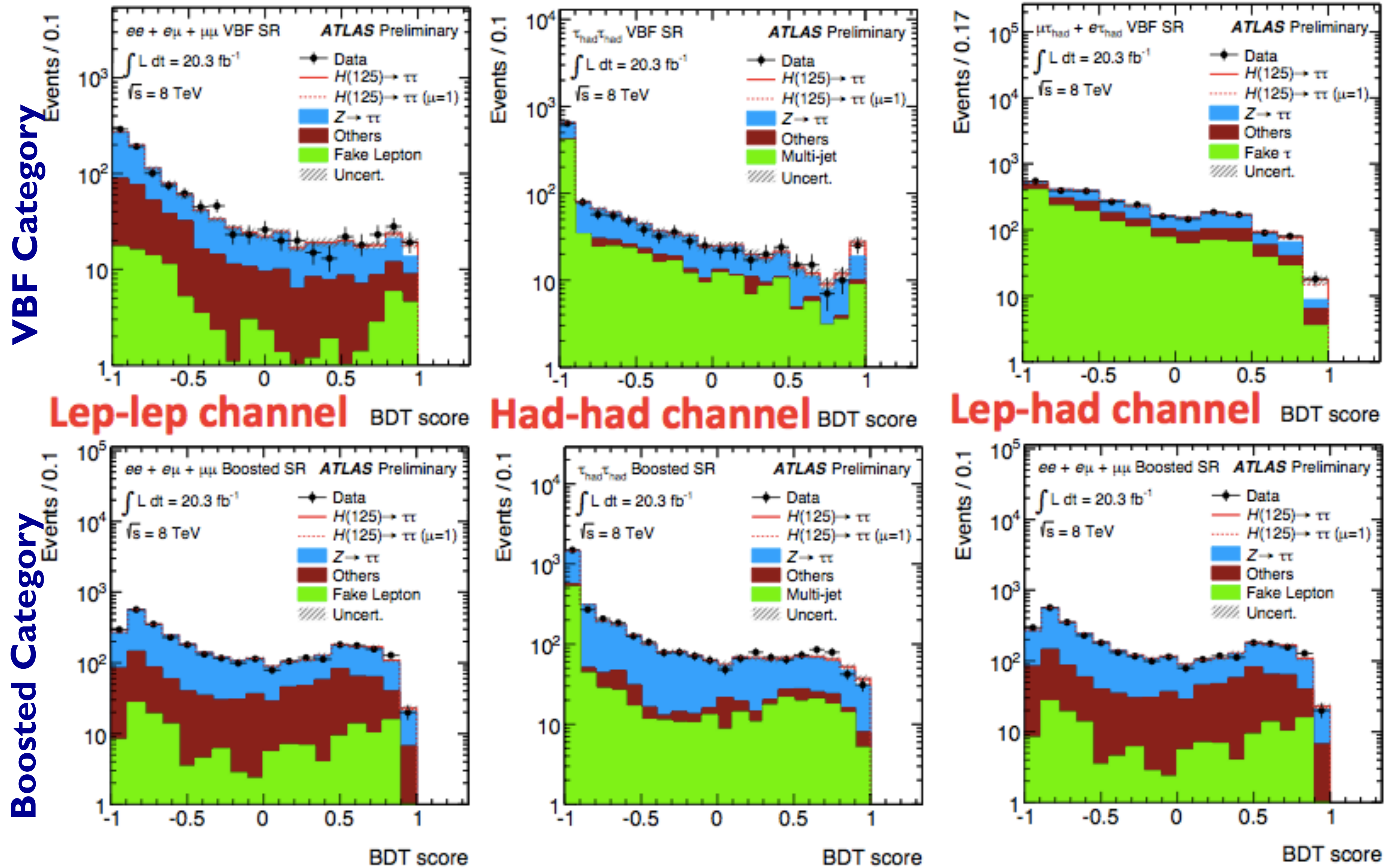
ATLAS BoostedDecisionTrees

- Each channel and sub-category has it's own BDT training.



- Train the BDT to recognize signal and background based on these training variables:
  - DiTau Mass & Separation
  - DiJet Mass & Separation
  - Direction of MET wrt Taus
  - How momentum points in the event: scalar and vector sum pT
- The BDT is the final discriminating variable

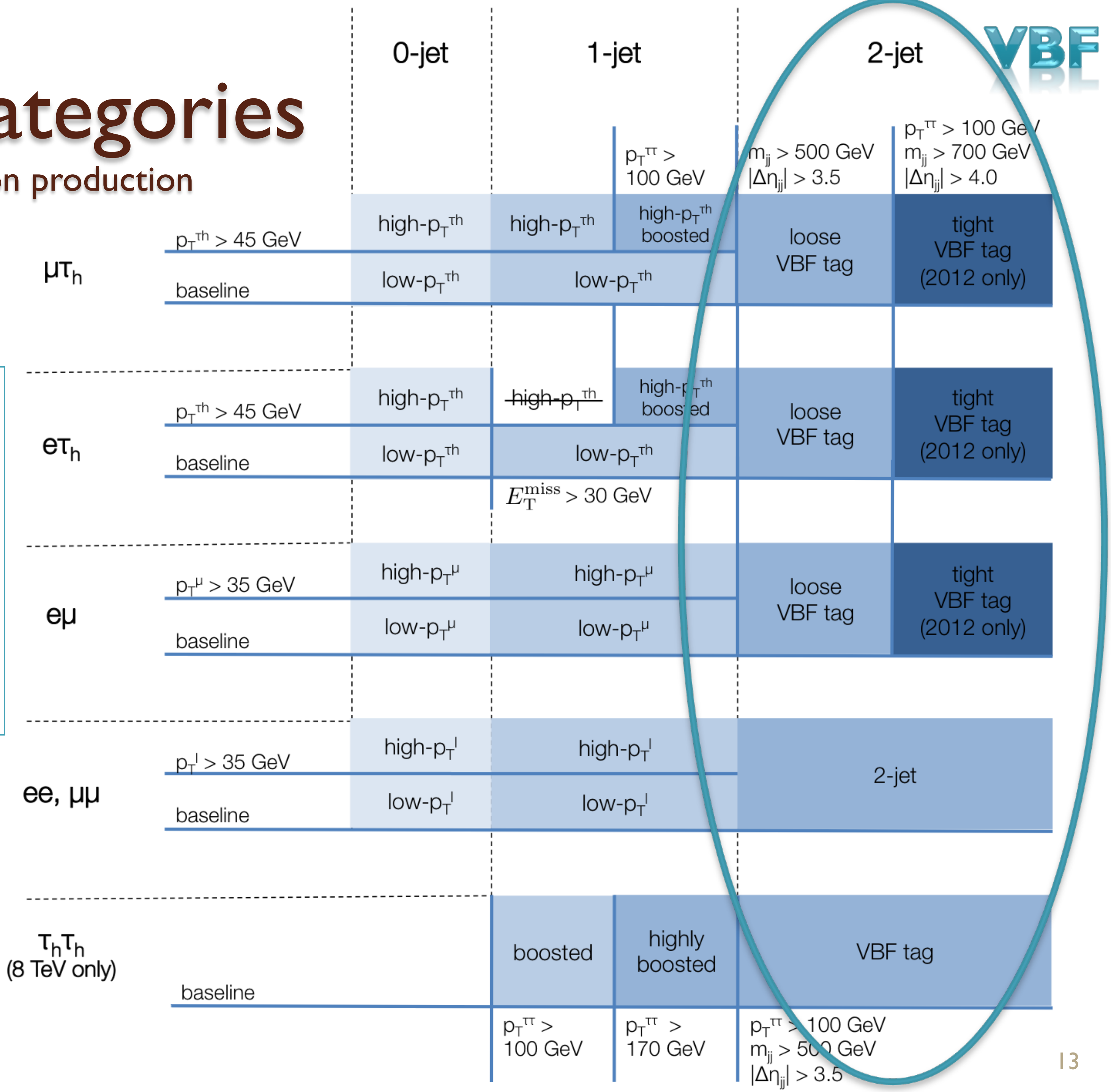
ATLAS BDT Scores



CMS Categories

VBF & GluonFusion production

Variables used in cutting:
Jet Multiplicity
DiJet Mass & Separation
DiTau pT
H pT, Tau or lep pT
Central Jet Veto



CMS Categories

Associated Production:VH

Variables used in cutting:
Lepton& Tau pT
Lepton & Tau Eta
Lepton & Tau isolation

Isolation:

$$I^L \equiv \sum_{\text{charged}} p_T + \max\left(0, \sum_{\text{neutral}} p_T + \sum_{\gamma} p_T - \frac{1}{2} \sum_{\text{charged, PU}} p_T\right)$$

$$R^L \equiv I^L / p_T^L,$$

Trigger(pT threshold)

WH

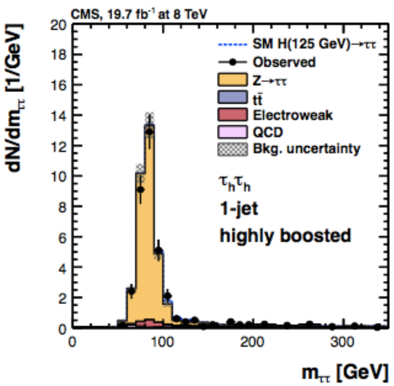
Channel	HLT requirement	Lepton selection		
$\mu + \mu\tau_h$	$\mu(17)$ & $\mu(8)$	$p_T^{\mu_1} > 20$ $p_T^{\mu_2} > 10$ $p_T^{\tau_h} > 20$	$\lvert\eta^\mu\rvert < 2.4$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^\mu < 0.1$–0.2 $I^{\tau_h} < 2$
$e + \mu\tau_h$/ $\mu + e\tau_h$	$e(17)$ & $\mu(8)$ $e(8)$ & $\mu(17)$	$p_T^{\ell_1} > 20$ $p_T^{\ell_2} > 10$ $p_T^{\tau_h} > 20$	$\lvert\eta^e\rvert < 2.5$ $\lvert\eta^\mu\rvert < 2.4$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^\ell < 0.1$–0.2 $I^{\tau_h} < 2$
$\mu + \tau_h\tau_h$	$\mu(24)$	$p_T^{\mu} > 24$ $p_T^{\tau_{h,1}} > 25$ $p_T^{\tau_{h,2}} > 20$	$\lvert\eta^\mu\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^\mu < 0.1$ $I^{\tau_h} < 2$–3
$e + \tau_h\tau_h$	$e(20)$ & $\tau_h(20)$ $e(22)$ & $\tau_h(20)$	$p_T^{e} > 24$ $p_T^{\tau_{h,1}} > 25$ $p_T^{\tau_{h,2}} > 20$	$\lvert\eta^e\rvert < 2.1$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^e < 0.1$–0.15 $I^{\tau_h} < 2$

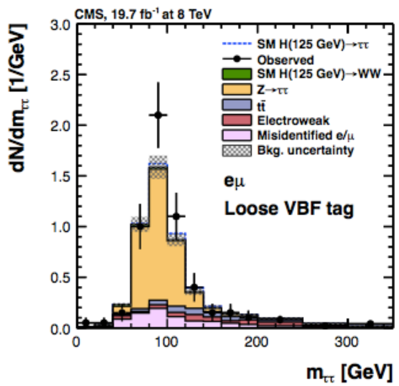
ZH

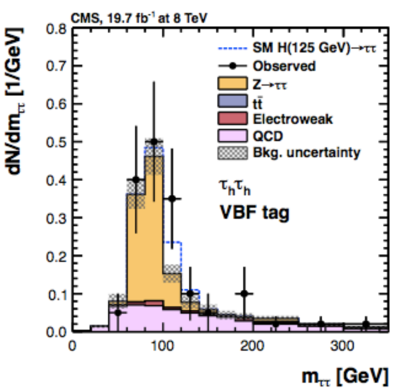
Resonance	HLT requirement	Lepton selection		
$Z \to \mu\mu$	$\mu(17)$ & $\mu(8)$	$p_T^{\mu_1} > 20$ $p_T^{\mu_2} > 10$	$\lvert\eta^\mu\rvert < 2.4$	$R^\mu < 0.3$
$Z \to ee$	$e(17)$ & $e(8)$	$p_T^{e_1} > 20$ $p_T^{e_2} > 10$	$\lvert\eta^e\rvert < 2.5$	$R^e < 0.3$
$H \to \mu\tau_h$		$p_T^{\mu} > 10$ $p_T^{\tau_h} > 15$	$\lvert\eta^\mu\rvert < 2.4$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^\mu < 0.3$ $I^{\tau_h} < 2$
$H \to e\tau_h$		$p_T^{e} > 10$ $p_T^{\tau_h} > 15$	$\lvert\eta^e\rvert < 2.5$ $\lvert\eta^{\tau_h}\rvert < 2.3$	$R^e < 0.2$ $I^{\tau_h} < 2$
$H \to \tau_h\tau_h$		$p_T^{\tau_h} > 15$	$\lvert\eta^{\tau_h}\rvert < 2.3$	$I^{\tau_h} < 1$
$H \to e\mu$		$p_T^{\ell} > 10$	$\lvert\eta^e\rvert < 2.5$ $\lvert\eta^\mu\rvert < 2.4$	$R^\ell < 0.3$

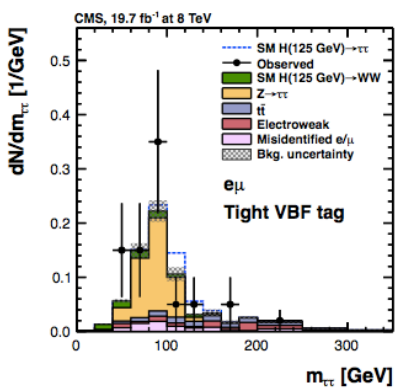
First ID $Z \to ll$, $60 < m_Z < 120$ GeV, then a $H \to ll$ from remaining leptons

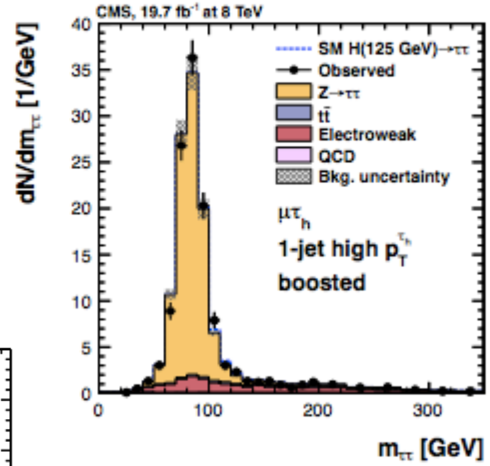
CMS Categories (10 out of 84)

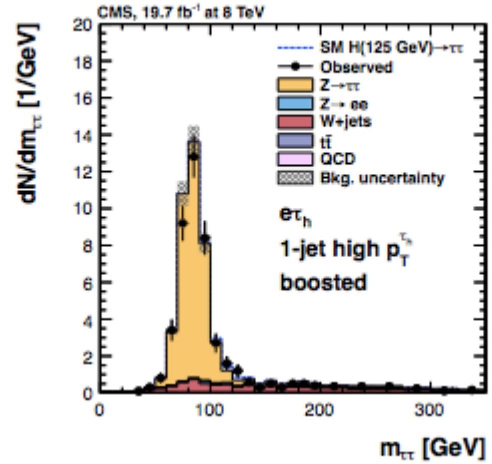


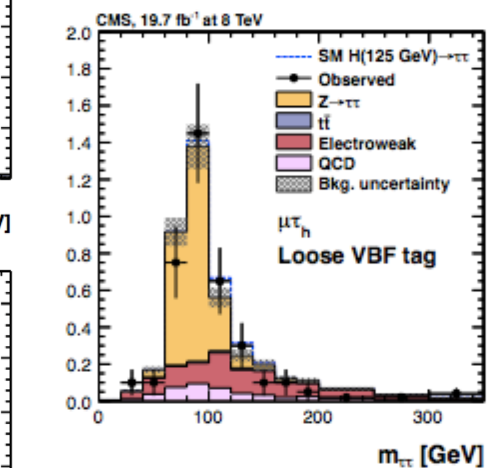


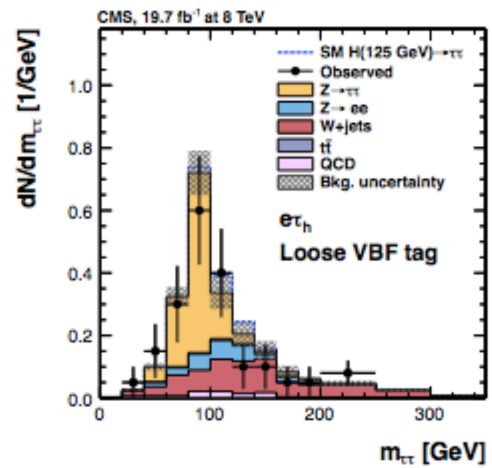




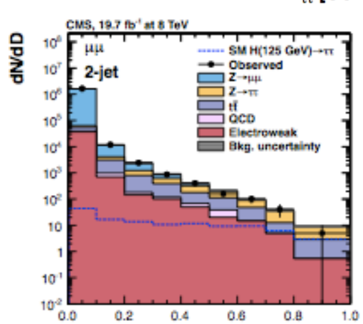


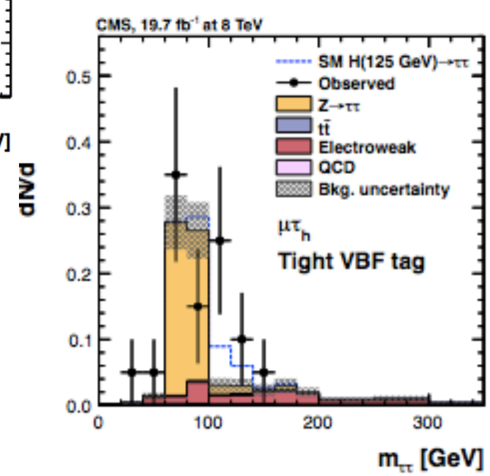


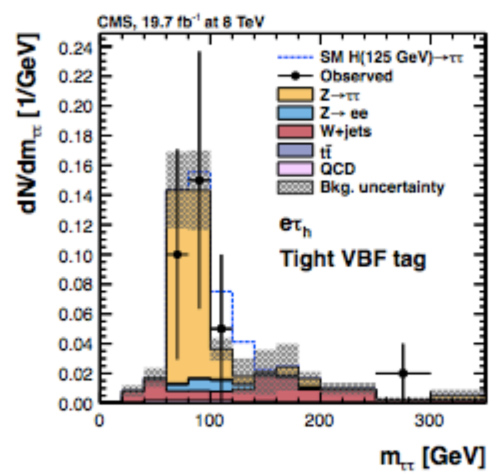




The Mass or Visible Mass is used in all but the two weakest categories, where a Decision Tree is used.







Systematic Uncertainties:

- 1) Uncertainties related to theory (most important for Signal MC)
- 2) Uncertainties from experimental sources

ATLAS:

Source of Uncertainty	Uncertainty on μ
Signal region statistics (data)	0.30
$Z \to \ell\ell$ normalization ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ boosted)	0.13
ggF $d\sigma/dp_T^H$	0.12
JES η calibration	0.12
Top normalization ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ VBF)	0.12
Top normalization ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ boosted)	0.12
$Z \to \ell\ell$ normalization ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ VBF)	0.12
QCD scale	0.07
di-τ_{had} trigger efficiency	0.07
Fake backgrounds ($\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$)	0.07
τ_{had} identification efficiency	0.06
$Z \to \tau^+\tau^-$ normalization ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$)	0.06
τ_{had} energy scale	0.06

$$Signal\ strength\ \mu = \frac{\sigma_{measured}}{\sigma_{SM}}$$

- Leading theory uncertainty is due to effect of top, bottom, charm quark masses in gluon-gluon loop, affecting $P_T(H)$ spectrum in gluon fusion produced H (~30%)

CMS overview (not the full list)

Tau ID & Trigger	6 to 10% on rate
$\mathrm{Tau}_{\mathrm{had}}$ Energy Scale	3% on the Energy; 1-29% on yields
Z → ll normalization	20-80% on rate
Top & Diboson norm.	4-15%
qq' → H: PDF & Scale variations;	4% & 3% on rate
gg→H; generator differences & missing HO terms	30% & (10-41)% on rate

The Combined Fit: Determining the Signal Strength

- To determine the signal strength in the data, a pdf of each background and MC signal is made.
  - In ATLAS the discriminating variable is the BDT distribution
  - In CMS the discriminating variable is the $m_{\tau\tau}$, $m_{vis\,\tau\tau}$, or the BDT for the ee and $\mu\mu$ *channels*
- A global maximum likelihood fit is done simultaneously in all channels:
  - The backgrounds and signals are allowed to move within their systematic uncertainties
  - The normalizations for many backgrounds are floated
  - Best fit for the signal strength μ is extracted.

$$\textit{Signal strength}\ \mu = \frac{\sigma_{measured}}{\sigma_{SM}}$$

Results: By the numbers

ATLAS Fit results:
Signal observed with significance
3.2 σ expected, 4.1 observed
for signal hypothesis of m_H =125 GeV

Measured signal strength obs/SM:
μ =1.4 (+0.5 -0.4)

ATLAS: Events in highest BDT-score bins

		Lep-lep	Lep-had	Had-had
VBF	Signal	5.7±1.7	8.7±2.5	8.8±2.2
	Bkgd	13.5±2.4	8.7±2.4	11.8±2.6
	Data	**19**	**18**	**19**
Boosted	Signal	2.6±0.8	8.0±2.5	3.6±1.1
	Bkgd	20.2±1.8	32±4	11.2±1.9
	Data	20	34	15

High sensitivity CMS categories:

Event category	SM Higgs (m_H = 125 GeV) ggH	VBF	VH	Σ signal	Background	Data	$\frac{S}{S+B}$	σ_{eff} (GeV)
$\mu\tau_h$								
VBF tag 7 TeV	0.2	1.3	–	1.6 ± 0.1	22 ± 2	23	0.14	19.6
Loose VBF tag 8 TeV	1.1	3.4	–	4.5 ± 0.4	81 ± 7	76	0.17	17.0
Tight VBF tag 8 TeV	0.3	2.0	–	2.4 ± 0.2	15 ± 2	20	0.49	18.1
$e\tau_h$								
VBF tag 7 TeV	0.2	0.7	–	0.9 ± 0.1	14 ± 2	13	0.24	15.9
Loose VBF tag 8 TeV	0.6	1.8	–	2.4 ± 0.2	45 ± 4	40	0.14	16.7
Tight VBF tag 8 TeV	0.3	1.3	–	1.6 ± 0.1	9 ± 2	7	0.51	16.2
$\tau_h\tau_h$								
1-jet boosted 8 TeV	7.2	2.1	1.0	10.3 ± 1.7	1133 ± 49	1120	0.054	15.2
1-jet highly-boosted 8 TeV	5.6	1.6	1.2	8.4 ± 1.2	380 ± 23	366	0.14	13.1
VBF tag 8 TeV	0.5	2.4	–	3.0 ± 0.3	29 ± 4	34	0.32	14.3

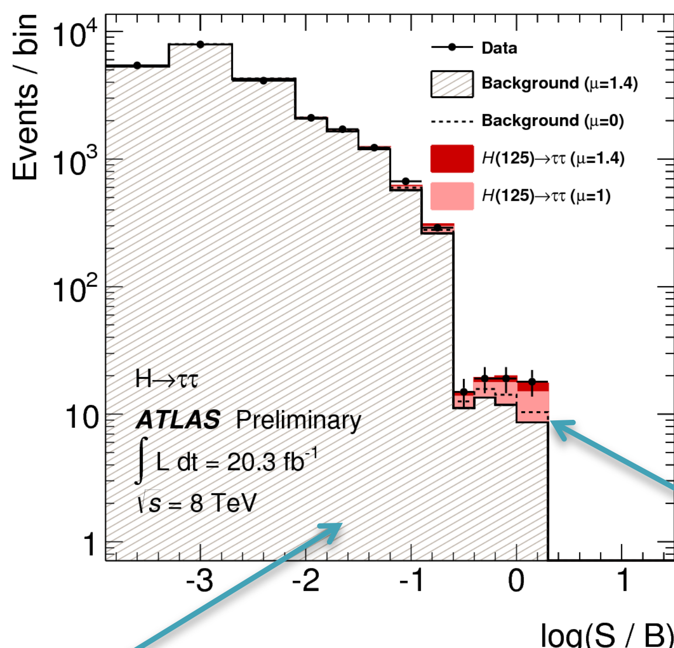
CMS Fit results:
Signal observed with significance **3.7 σ expected, 3.2 observed**
for signal hypothesis of m_H =125 GeV

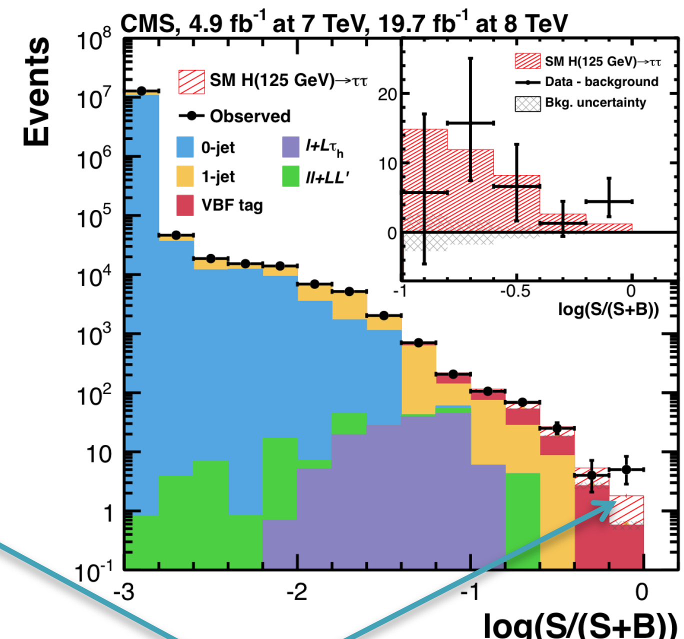
Measured signal strength obs/SM: μ =0.78 (+-0.27)

Representing the Signal

- Every event is associated to a Histogram Bin in the BDT or Mass Distribution. Plot the Signal/Background of the bin for every event:

- In effect, put high-sensitivity bins together and compare to data

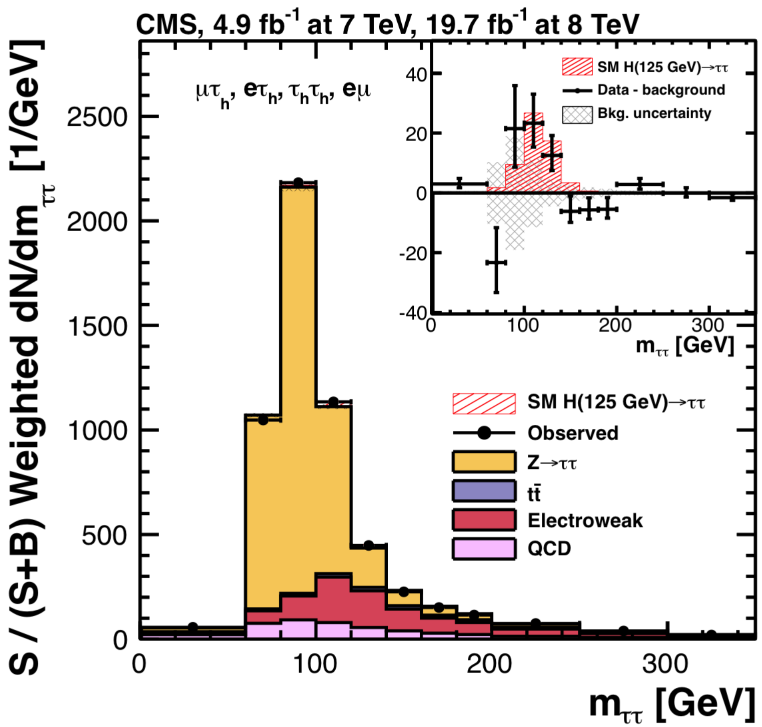


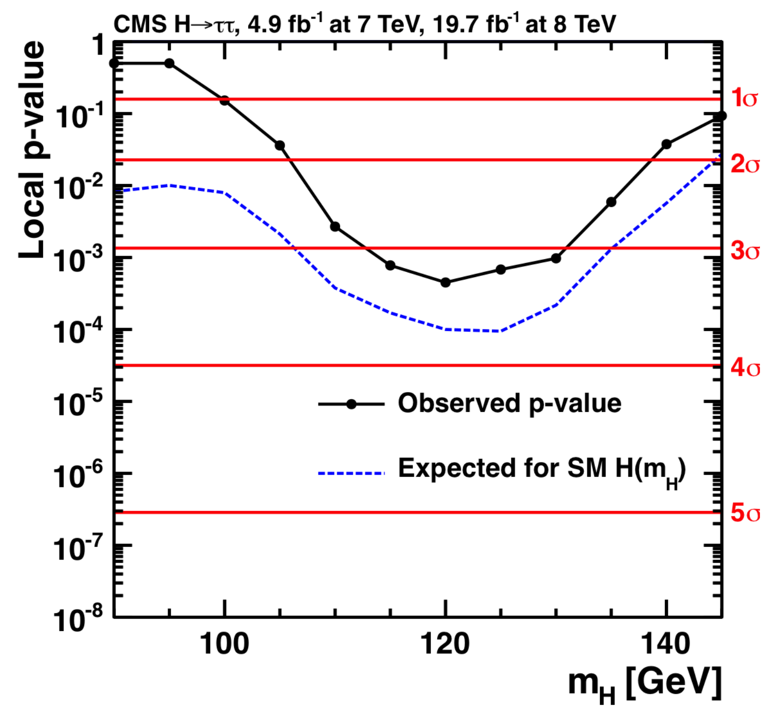


- Pink is predicted SM signal

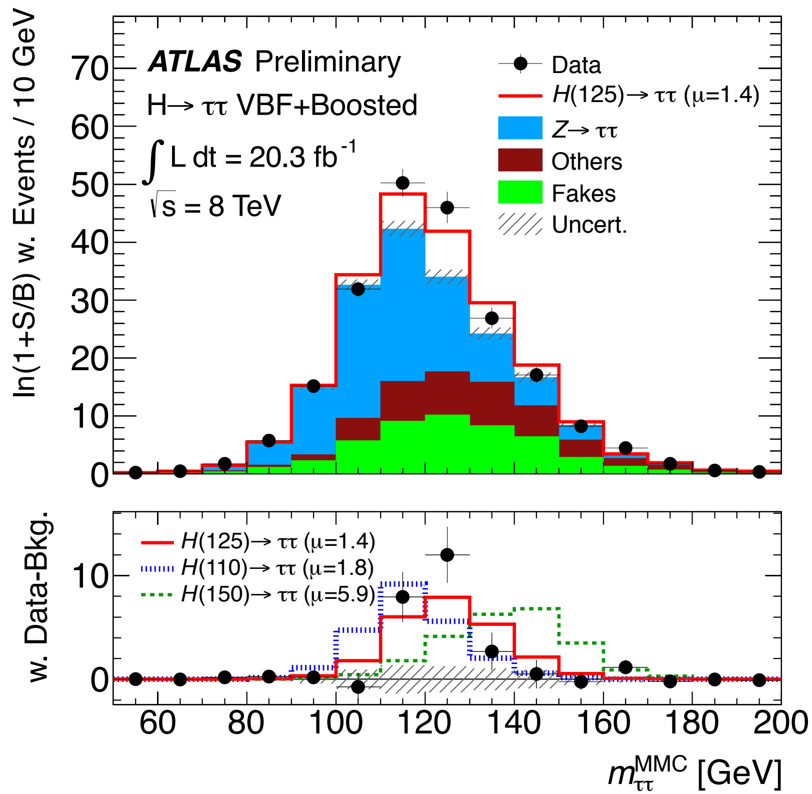
- Grey is predicted background-only.

CMS Compatibility with M_H=125 GeV





ATLAS Compatibility With M_H=125 GeV



Signals at M_H=110, 125 and 150 GeV are shown at best fit μ ; post-fit background normalizations

- This analysis was not designed to measure the H mass. But we can look at how well the excess matches various mass hypotheses
- Each event is weighted by ln(1+S/B) for its corresponding bin in BDT-score
- Excess of data events is consistent with presence of Higgs at 125 GeV

Future Htautau Measurements

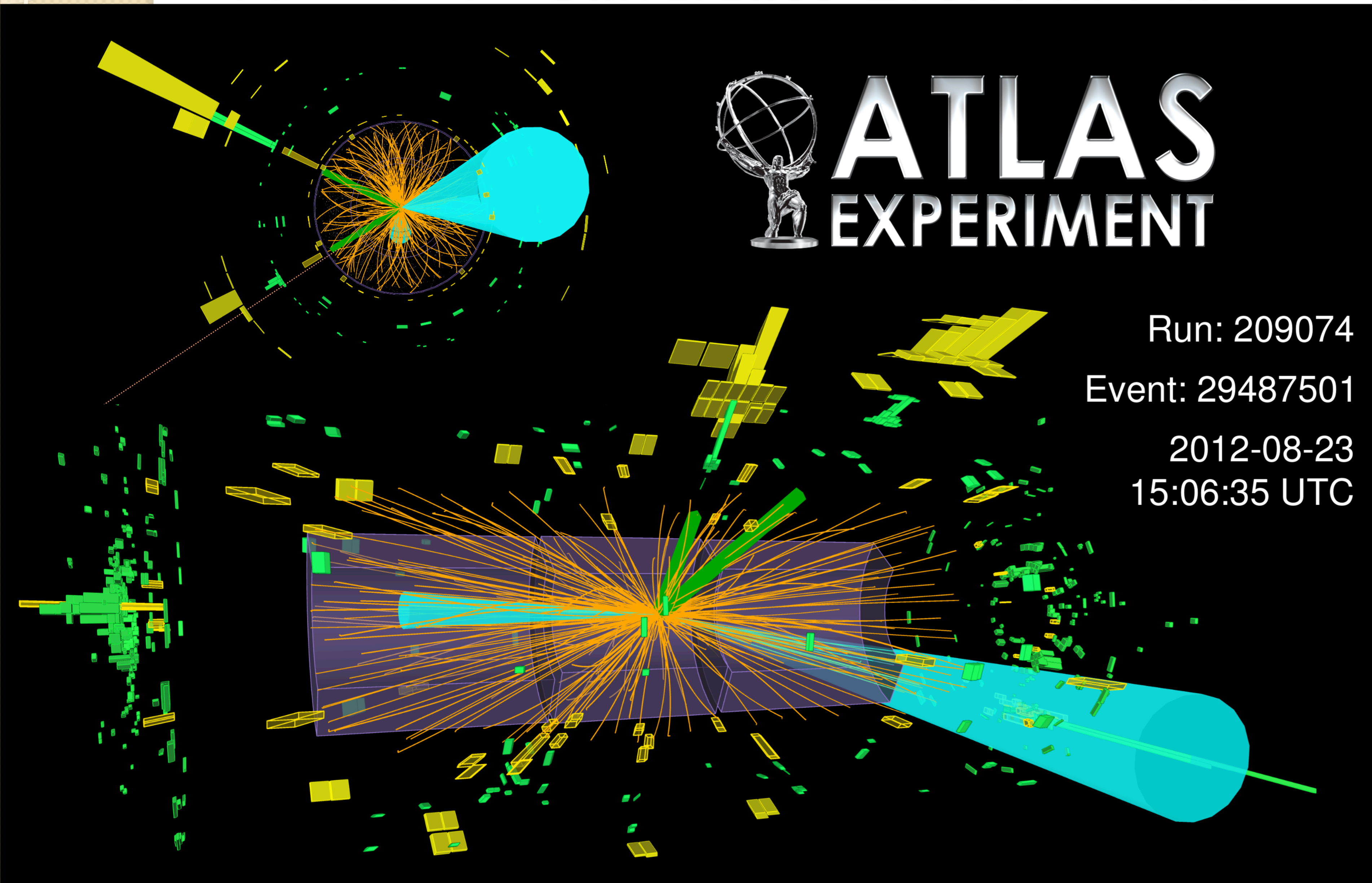
- H Fermion coupling
- Possibility to measure H Charge Parity properties through polar angle distributions with specific τ -spin correlations? (S. Berge, W. Bernreuther, B. Niepelt, H. Spiesberger arXiv:1108.0670v2)
- ~3000 fb^{-1}? Higgs Self-coupling with bbtautau (M. J. Dolan, C. Englert ,M. Spannowsky, JHEP10(2012)112)

Summary

- Both ATLAS and CMS see > 3 σ evidence for H→ $\tau\ \tau$, the first direct evidence for H decaying to fermions at the LHC
- Measured signal strengths are consistent with the Standard Model:
  - ATLAS: μ =1.4 (+0.5 -0.4)
  - CMS: μ =0.78 (+-0.27)

Backup

H→ τ τ Candidate Event In Had-



ATLAS Monte Carlo Samples

MC Sample	Generator
GGF	POWHEG w/(NLO) QCD (CT10 PDF)
VBF	POWHEG w/(NLO) QCD (CT10 PDF)
VH	LO QCD PYTHIA (CTEQ6L1)
W/Z+jets	ALPGEN. MLM matching scheme between the hard process and the parton shower (calculated with LO matrix elements for up to five jets)
ttbar	MC@NLO (CT10)
single top (t/s channel, *Wt)*	ACERMC (CTEQ6L1)
WZ, ZZ, & WW for leplep	HERWIG (CTEQ6L1)
WW (lephad and hadhad)	ALPGEN interfaced to HERWIG (CTEQ6L1)

Soft-gluon resummation up to NNLLog order. The finite quark-mass effects are taken into account in POWHEG. The parton shower, hadronization and underlying event simulations are provided by PYTHIA.

CMS MC

- GGF & VHF Higgs Production: Powheg
- VH & ttH: Pythia
- Z+jets, W+jets, tt+jets, diboson: MadGraph
- single top: Powheg

- Powheg and Madgraph generators are interfaces with pythia for parton shower and fragmentation
- The PYTHIA parameters affecting the description of the underlying event are set to the Z2 tune for the 7 TeV samples and to the Z2$*$ tune for the 8 TeV samples. All generators are interfaced with TAUOLA for the simulation of the τ*-lepton* decays. The Higgs boson *pT* spectrum from POWHEG is reweighted to the spec- trum obtained from a next-to-next-to-leading-order (NNLO) calculation using HRES.

Di-Tau Mass Calculation

- 4 equations, 6-8 unknowns

$$\begin{aligned}
\not{E}_{\mathrm{T}_x} &= p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \cos\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \cos\phi_{\mathrm{mis}_2} \\
\not{E}_{\mathrm{T}_y} &= p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \sin\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \sin\phi_{\mathrm{mis}_2} \\
M_{\tau_1}^2 &= m_{\mathrm{mis}_1}^2 + m_{\mathrm{vis}_1}^2 + 2\sqrt{p_{\mathrm{vis}_1}^2 + m_{\mathrm{vis}_1}^2}\sqrt{p_{\mathrm{mis}_1}^2 + m_{\mathrm{mis}_1}^2} \\
&\quad -2p_{\mathrm{vis}_1} p_{\mathrm{mis}_1} \cos\Delta\theta_{vm_1} \\
M_{\tau_2}^2 &= m_{\mathrm{mis}_2}^2 + m_{\mathrm{vis}_2}^2 + 2\sqrt{p_{\mathrm{vis}_2}^2 + m_{\mathrm{vis}_2}^2}\sqrt{p_{\mathrm{mis}_2}^2 + m_{\mathrm{mis}_2}^2} \\
&\quad -2p_{\mathrm{vis}_2} p_{\mathrm{mis}_2} \cos\Delta\theta_{vm_2}
\end{aligned} \tag{3}$$

Calculate all possible dR:

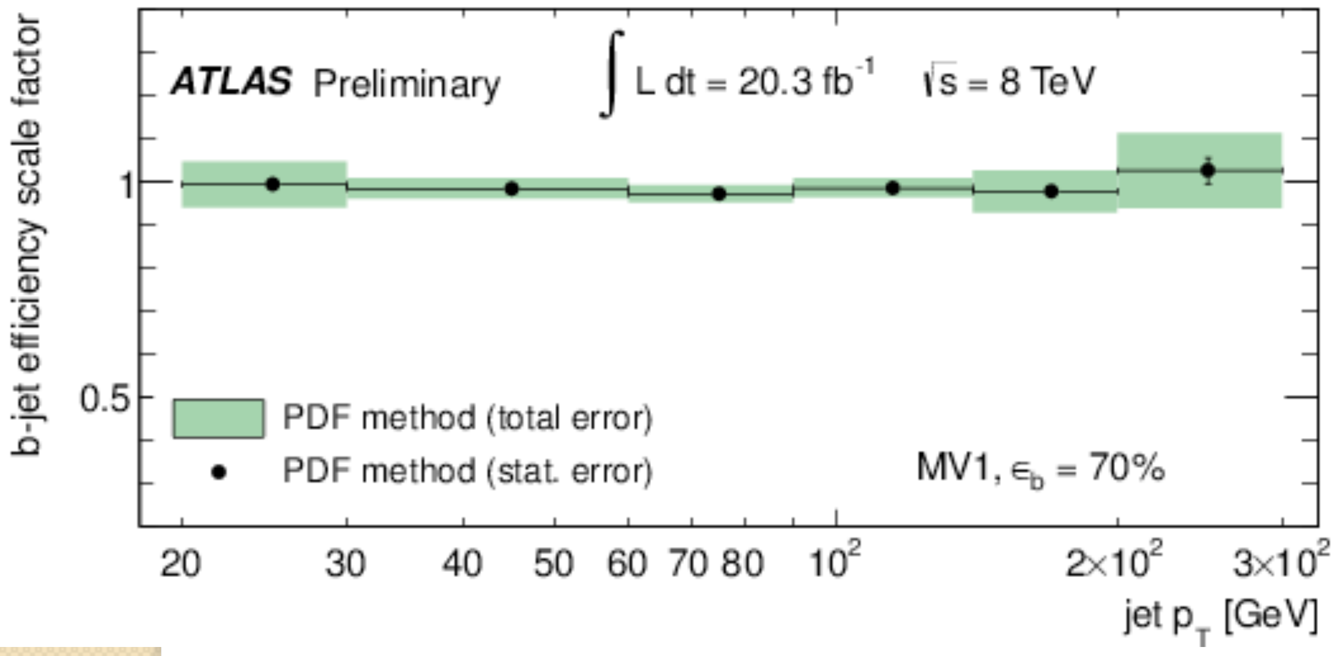
$$\Delta R = \sqrt{(\eta_{\mathrm{vis}} - \eta_{\mathrm{mis}})^2 + (\phi_{\mathrm{vis}} - \phi_{\mathrm{mis}})^2}$$

Triggers

Trigger	p_T threshold(s) [GeV]	$\tau_{lep}\tau_{lep}$	$\tau_{lep}\tau_{had}$	$\tau_{had}\tau_{had}$
Electron	24	•	•	
Muon	24		•	
Di-electron	12 ; 12	•		
Di-muon	18 ; 8	•		
Electron and Muon	12 ; 8	•		
Electron and τ_{had}	18 ; 20		•	
Muon and τ_{had}	15 ; 20		•	
Di-τ_{had}	29 ; 20			•

Table 1: Triggers used for each channel. When more than one trigger is used, a logical OR of the triggers is taken and the trigger efficiency is calculated accordingly.

b-Tagging @ ATLAS



Data-to-MC scale factor vs P_T(jet)

- H→ $\tau\ \tau$ uses a NN *b*-tagging algorithm with a working point corresponding to 70% *b*-tagging efficiency
  - Provides rejection factor of ~5 and ~150 against *c*-jets and light flavor jets
  - *b*-jet data-to-MC scale factors (SF) derived using *ttbar* di-lepton events
  - *b*-jet SF uncertainty in the range 2-10%
    - ~2% in the most important P_T(jet) range for H→bb search
  - *c*-jet SF uncertainty in the range 8-15%