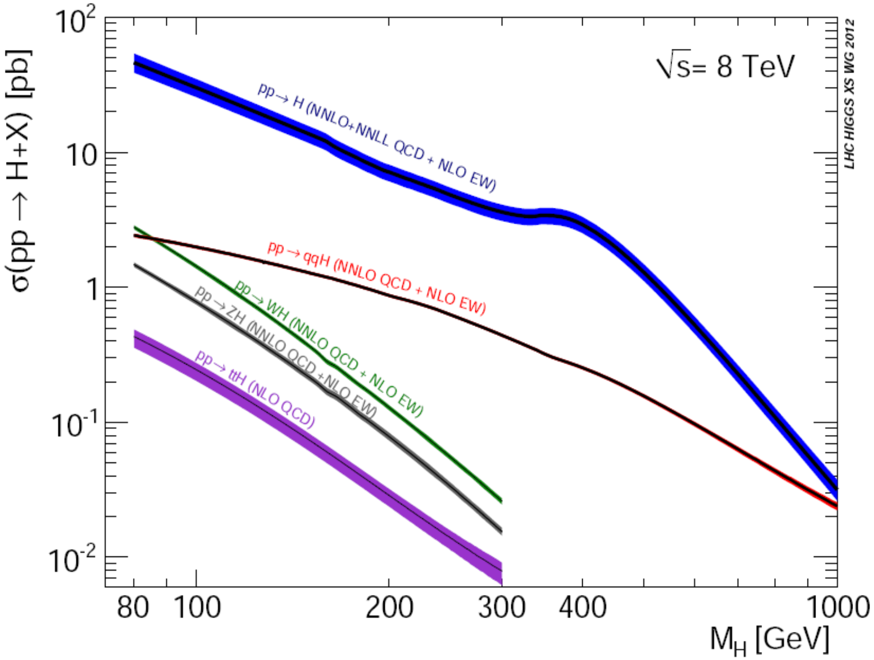
ATLAS Higgs Boson Search Update - the Ditau Decay

Xin Chen
Tsinghua University

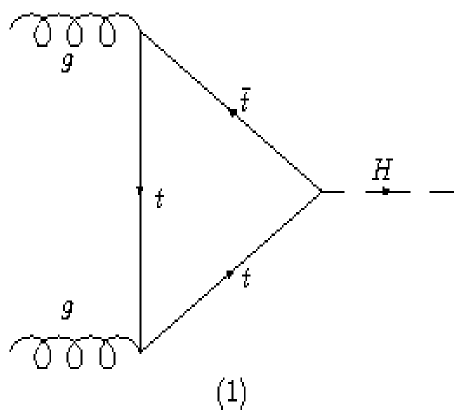
TeV 工作组学术研讨会
中山大学, May 15-18, 2014

Higgs production modes

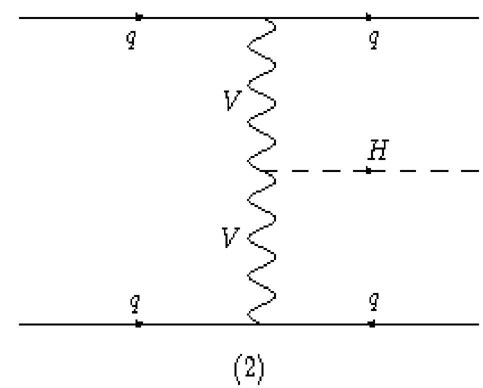
Higgs production cross sections:



Main Higgs production channels:

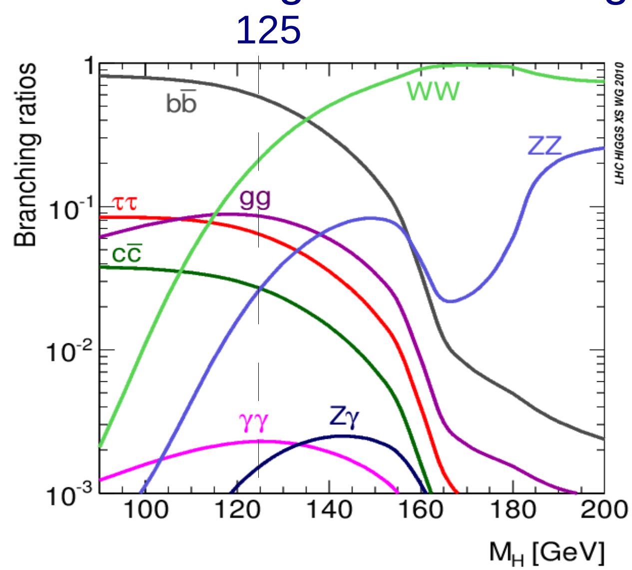


Gluon-gluon fusion



Vector-boson fusion

Branching ratios for a Higgs mass of 125 GeV:



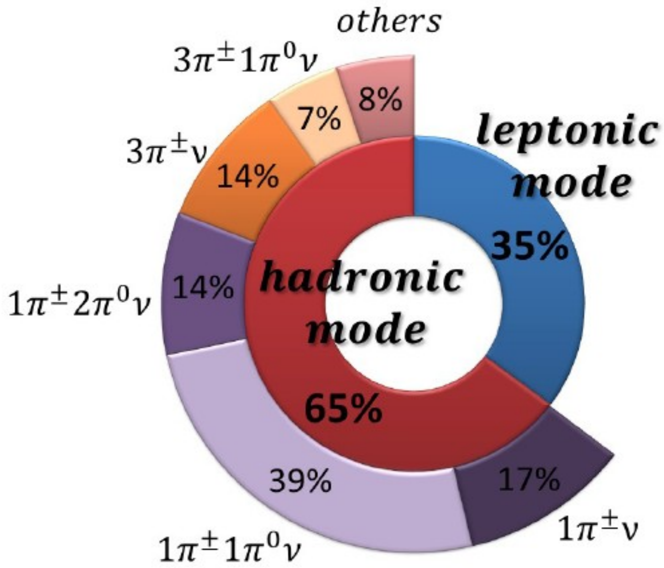
channel	BR
bb	57.7%
WW	21.5%
ττ	6.3%
ZZ	2.6%
γγ	0.23%

Fermions coupling:

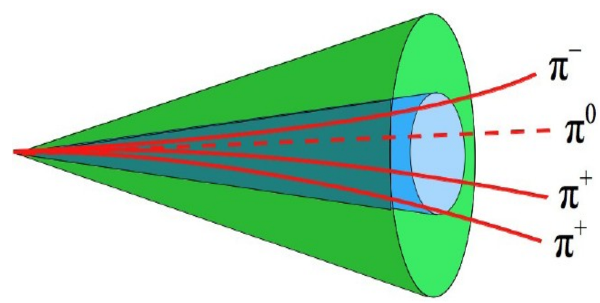
$$L_{Yukawa}=\frac{-\lambda_f v}{\sqrt{2}}\bar{f}\,f-\frac{\lambda_f}{\sqrt{2}}\bar{f}\,f\,H$$

Direct search for Hττ decay
will confirm it is a SM Higgs

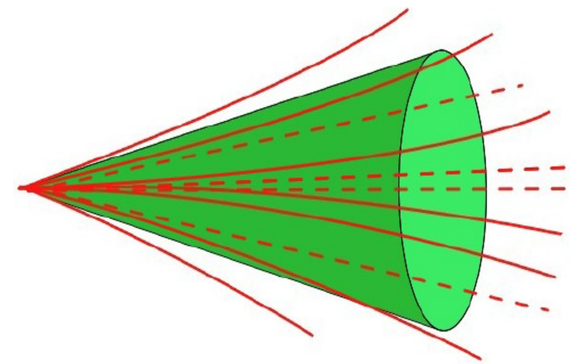
$H \to \tau\tau$



τ-hadron



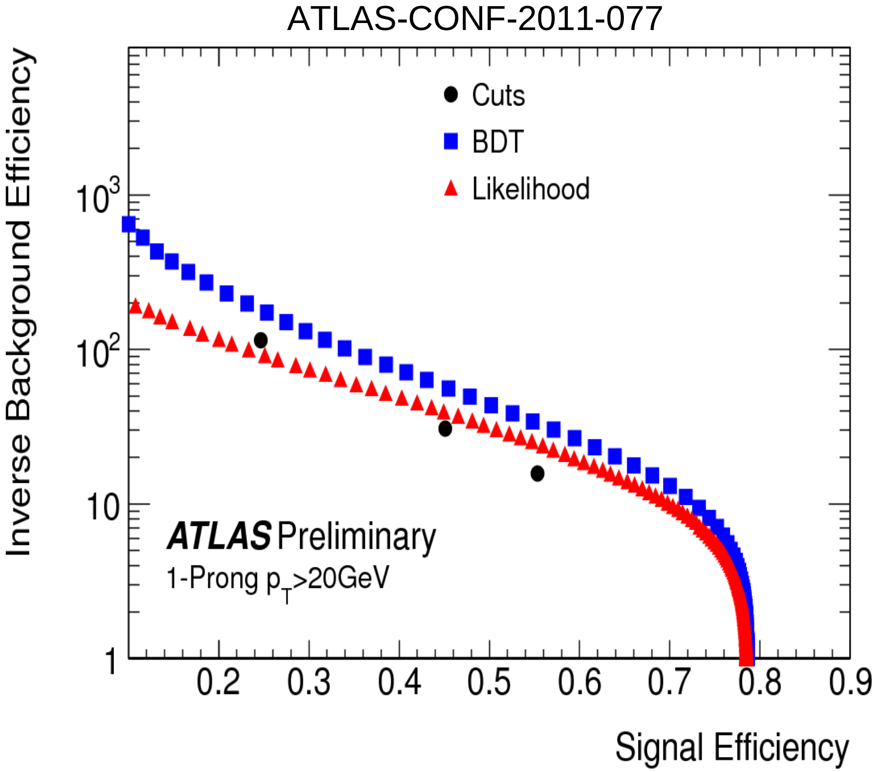
jet



ee, $e\mu$, $\mu\mu$, $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$

$\tau_{lep}\tau_{lep}$ $\quad$ $\tau_{lep}\tau_{had}$ $\quad$ $\tau_{had}\tau_{had}$

For a typical working point around 50% efficiency, about 2% fake rate is expected, for both ATLAS and CMS



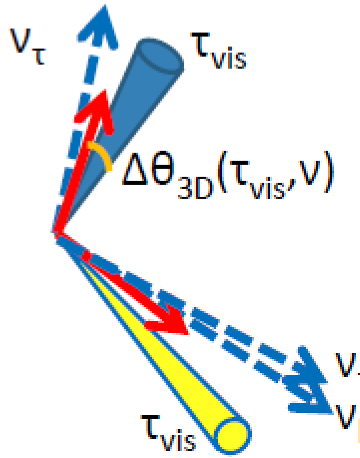
Higgs mass reconstruction

- Visible mass: invariant mass of the visible tau decay products

- Collinear mass: $$m_{\tau\tau} = \frac{m_{\ell\ell}}{\sqrt{x_1 \cdot x_2}}$$

  - Assume neutrinos and visible decay products from the tau are collinear, then ditau mass can be calculated as

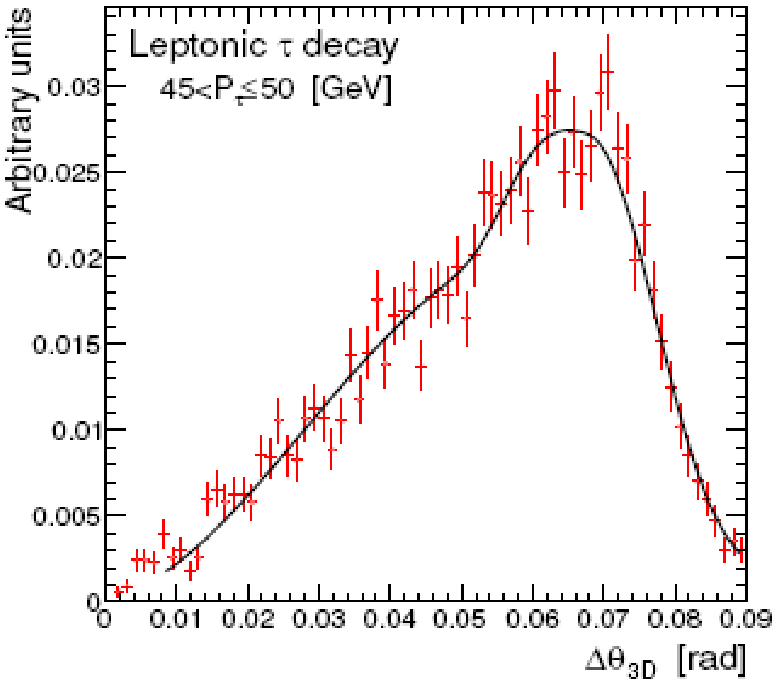
  - $x_{1,2}$ are the fractions of momenta carried away by the visible decay products from the tau



- Missing Mass Calculator (MMC):

  - Mass estimation by requiring the mutual orientations between neutrinos and other tau decay products are consistent with the mass and decay kinematics of a tau

  - Scan in the allowed phase space region (including MET variables) for the most likely solutions



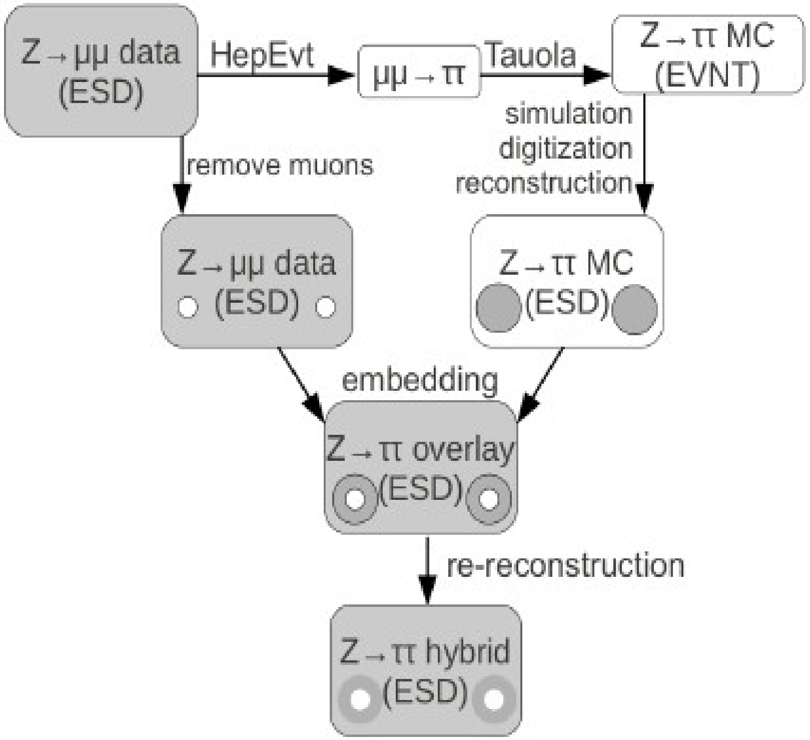
$H \rightarrow \tau\tau$ Preselection and Categorization

$\tau_{lep}\tau_{lep}$	$\tau_{lep}\tau_{had}$	$\tau_{had}\tau_{had}$
Exactly 2 leptons Opposite signs Single electron trigger Dilepton trigger $30 < m_{LL} < 75(100)$, MET>40(20) for SF(DF) $\Sigma p_{T,lep} > 35$, $\Delta\varphi_{LL} < 2.5$ $0.1 < x_{1,2} < 1$	Exactly 1 lepton + 1 tau Opposite signs Single lepton trigger $p_{T,lep} > 26$ $p_{T,tau} > 20$ m_T(lep,MET)<70 $m_{vis} > 40$	Exactly 2 taus Opposite signs Ditau trigger $p_{T,tau} > (35,25)$ At least 1 tau tight ID $0.8 < \Delta R_{\tau\tau} < 2.8$, $\Delta\eta_{\tau\tau} < 1.5$ MET>20 MET Centrality
b-jet (p_T>25) veto	b-jet (p_T>30) veto	N/A
2-jet VBF : $p_{T,J} > (40,30)$, $\Delta\eta_{JJ} > 2.2$	2-jet VBF : $p_{T,J} > (50,30)$, $\Delta\eta_{JJ} > 3.0$	2-jet VBF : $p_{T,J} > (50,30/35)$, $\Delta\eta_{JJ} > 2.0$
Boosted: $p_{T,H} > 100$ $p_{T,J1} > 40$	Boosted: $p_{T,H} > 100$	Boosted: $p_{T,H} > 100$

VBF and Boosted categories are optimized for Vector Boson Fusion and combined (VBF+gg-fusion) $H\tau\tau$ signals, respectively

$Z \rightarrow \tau\tau$ background

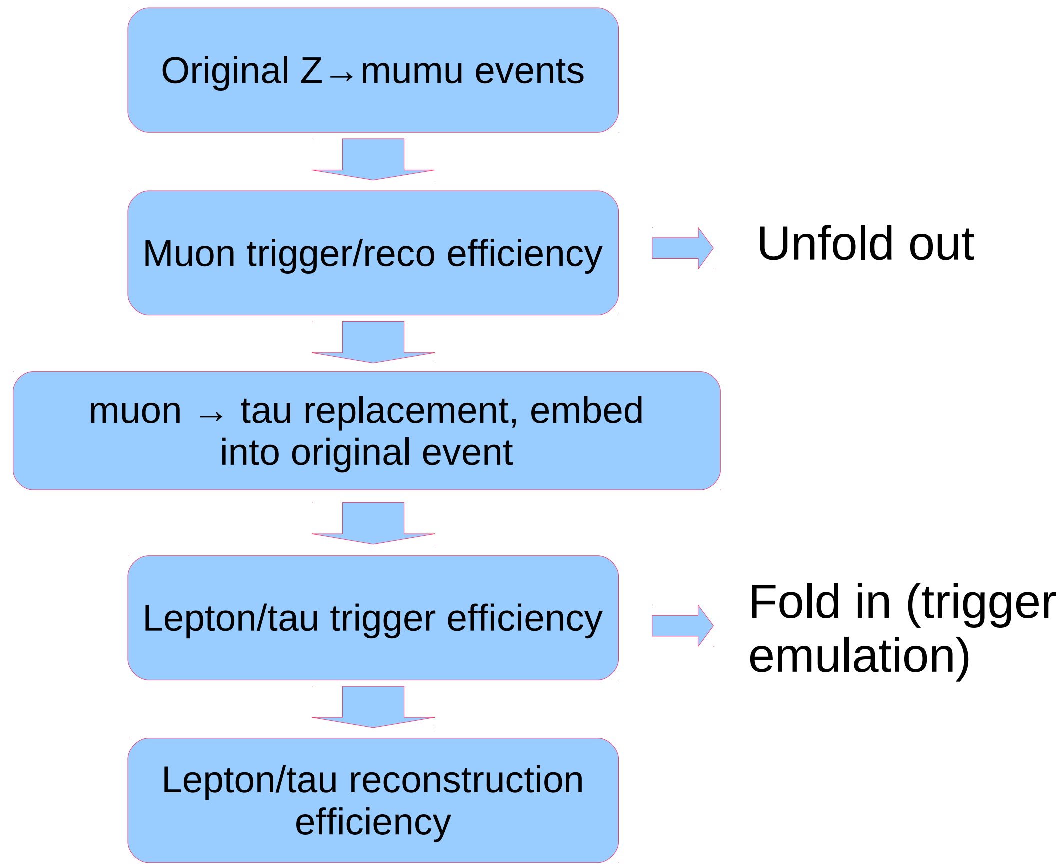
- $Z \rightarrow \tau\tau$ (dominant background) is estimated from data using embedding:



1) Replace muons from $Z \rightarrow \mu\mu$ data by taus and decay the taus

2) Embed the simulated tau decay products into the original event

Correction for trigger and acceptance:



Boosted Decision Trees

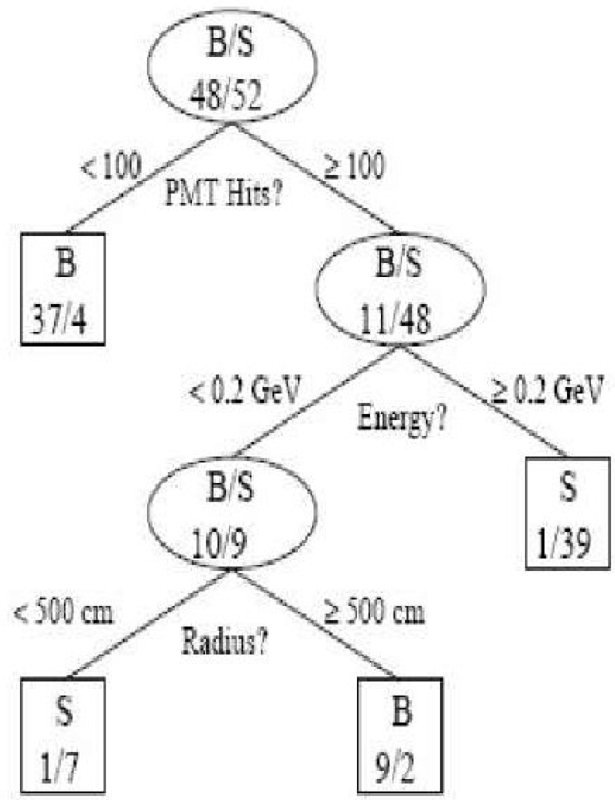
- Decision Tree (DT): repeatedly go through the discriminating variables, classify the events as signal (S) or background (B) with the best variable and cuts – a tree is formed based on which a decision (S or B) is made

- Boosting: misclassified events are given higher weights in the next DT building – multiple trees and averaged tree results.

- Must define 2 independent samples: one for training, one for testing

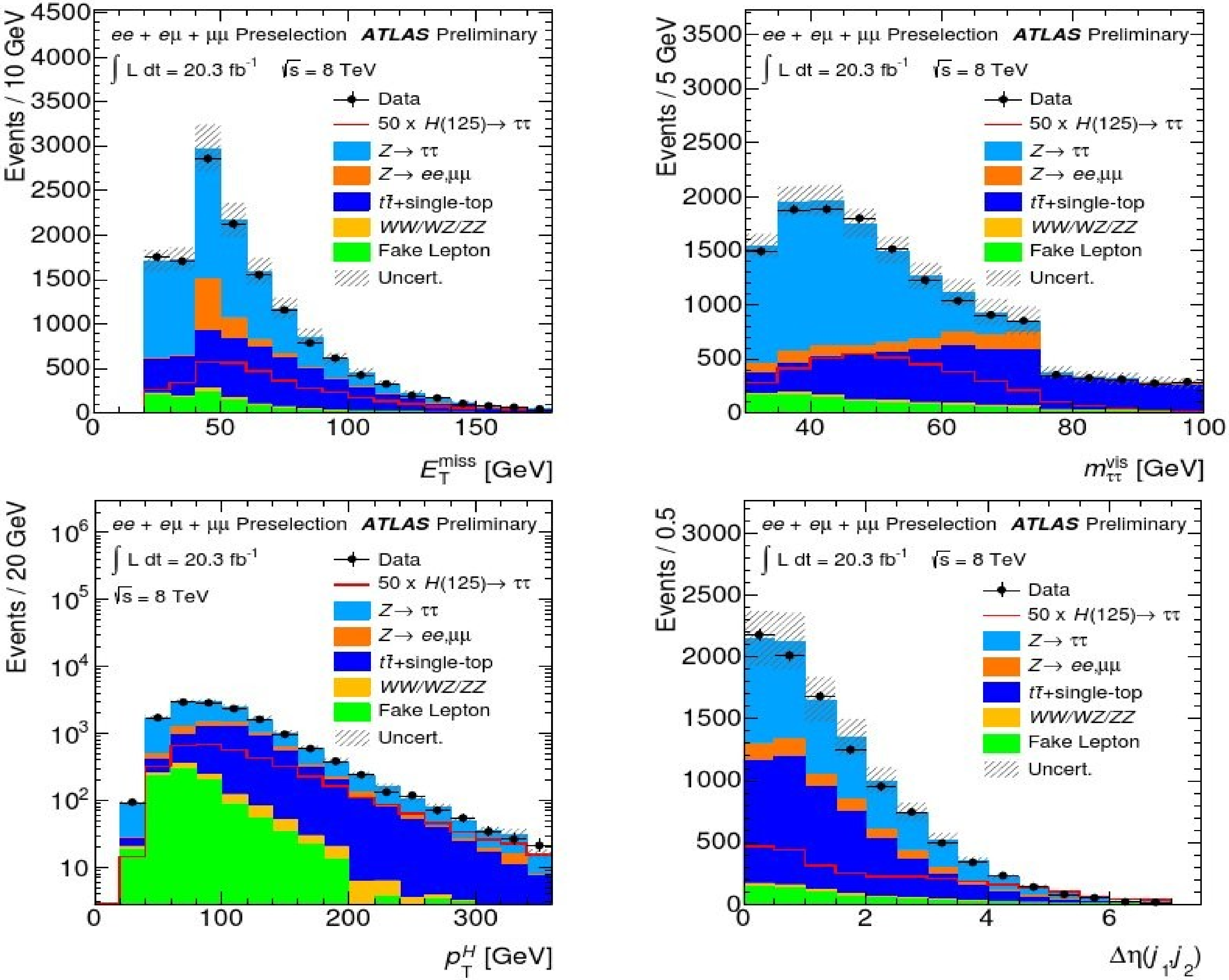
- BDT turns a group of weak classifiers (input variables) into a final strong one (BDT score) useful for data analysis with complex phase space distributions

  - Robust against statistical fluctuations in the training sample
  - Insenstitive to inclusion of very weak discriminating variables, or variables with large correlations with others

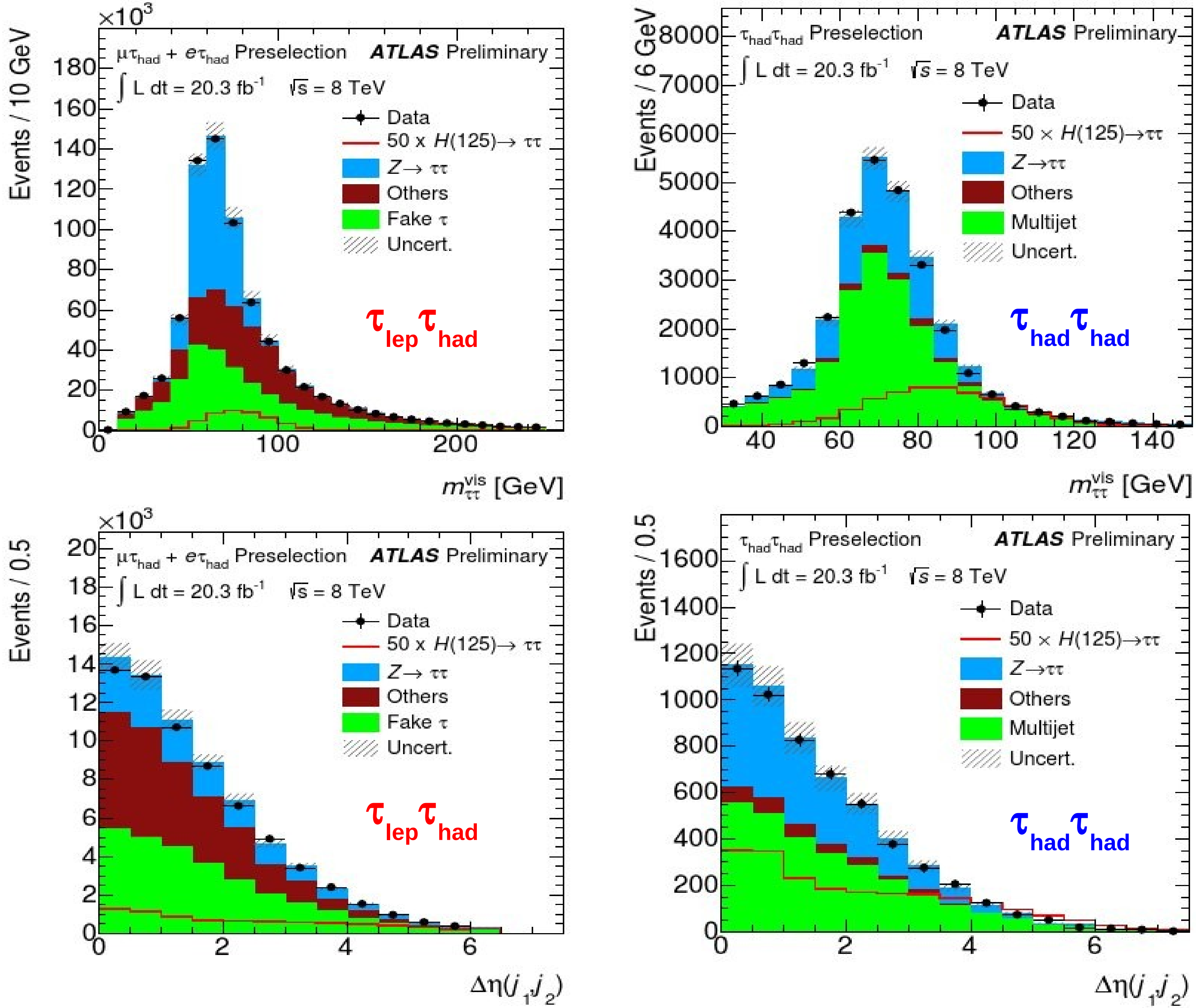


BDT is now a common MVA technique in ATLAS and CMS

Example BDT input variables for $\tau_{lep}\tau_{lep}$

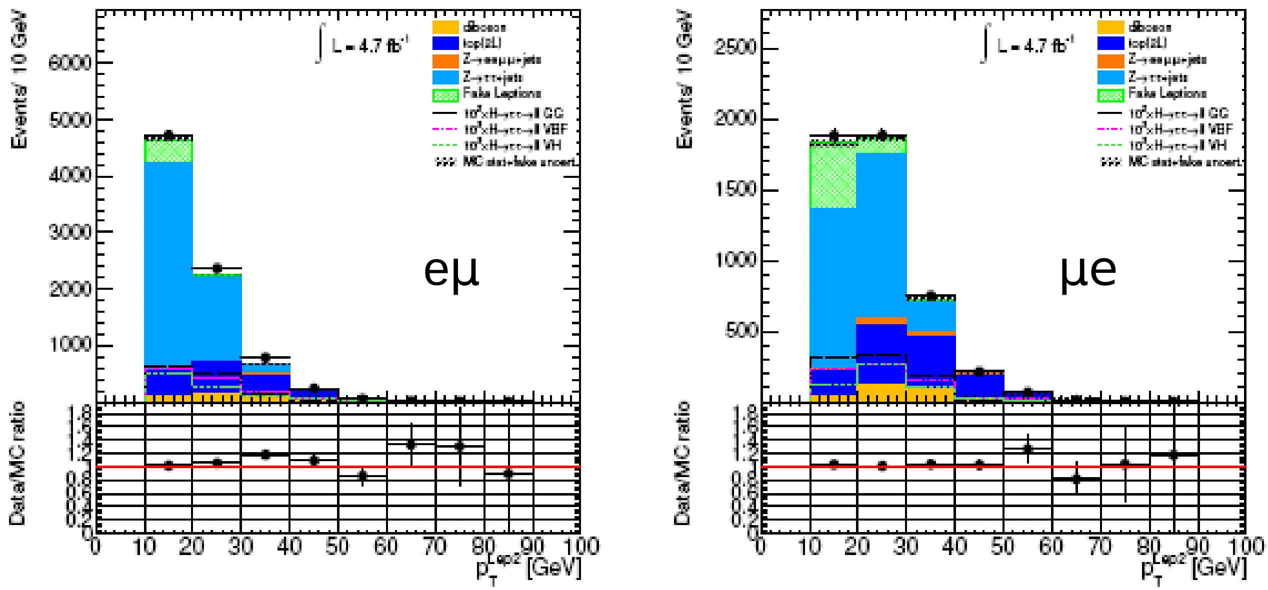


Example BDT input variables for $\tau_{lep}\tau_{had}/\tau_{had}\tau_{had}$



Background estimation in $\tau_{lep}\tau_{lep}$

- **Fake** (at least one fake lepton, W+jets/QCD multi-jet/semileptonic top): fake template fit after earlier cuts



Green: fake lepton background

- **Zll**: data/MC scale factors are derived in a control region in which $80<m_{ee,\mu\mu}<100$ GeV is applied (all other cuts are unchanged)
- **Top** (ttbar and single top): data/MC scale factors are derived in a control region with at least a b-jet (all other cuts are unchanged)
- **Diboson**: estimated by MC and validated in an independent region

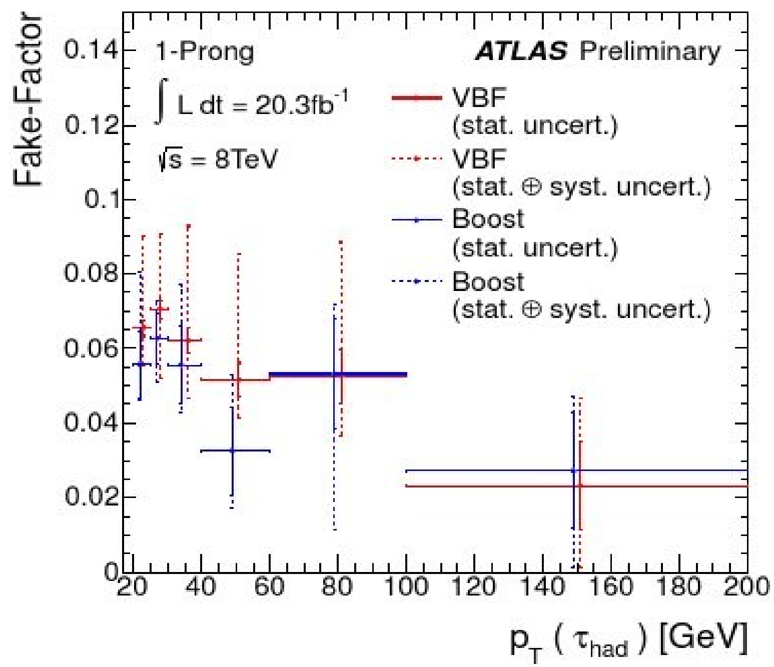
Background estimation in $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$

- $\tau_{lep}\tau_{had}$ uses a Fake Factor (FF) method to estimate the Fake (W+jets/QCD multi-jet):

$$N_{\tau}^{SR} = N_{anti-\tau}^{SR} \times FF$$

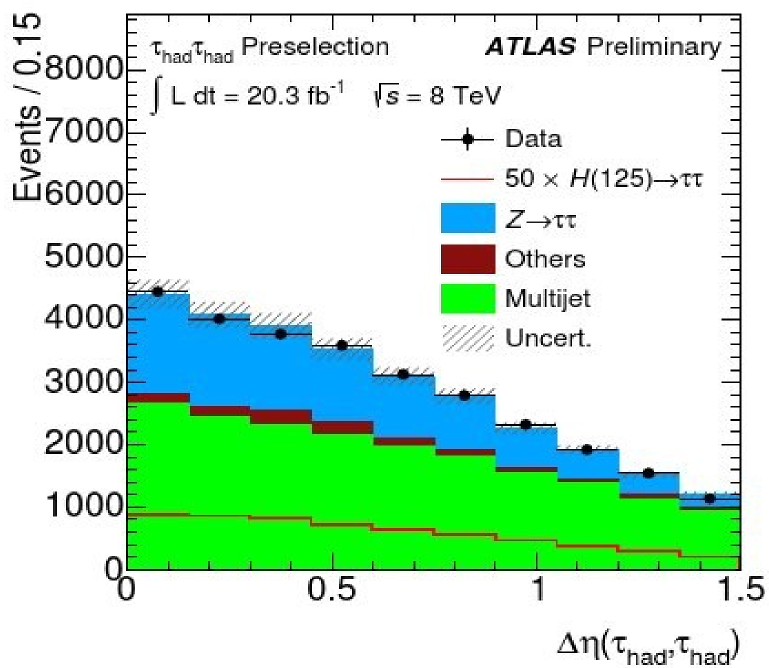
$$FF = \frac{N_{\tau}^{CR}}{N_{anti-\tau}^{CR}}$$

  - FF is estimated separately for W+jets (high m_T region) and QCD (loose not tight lepton) events

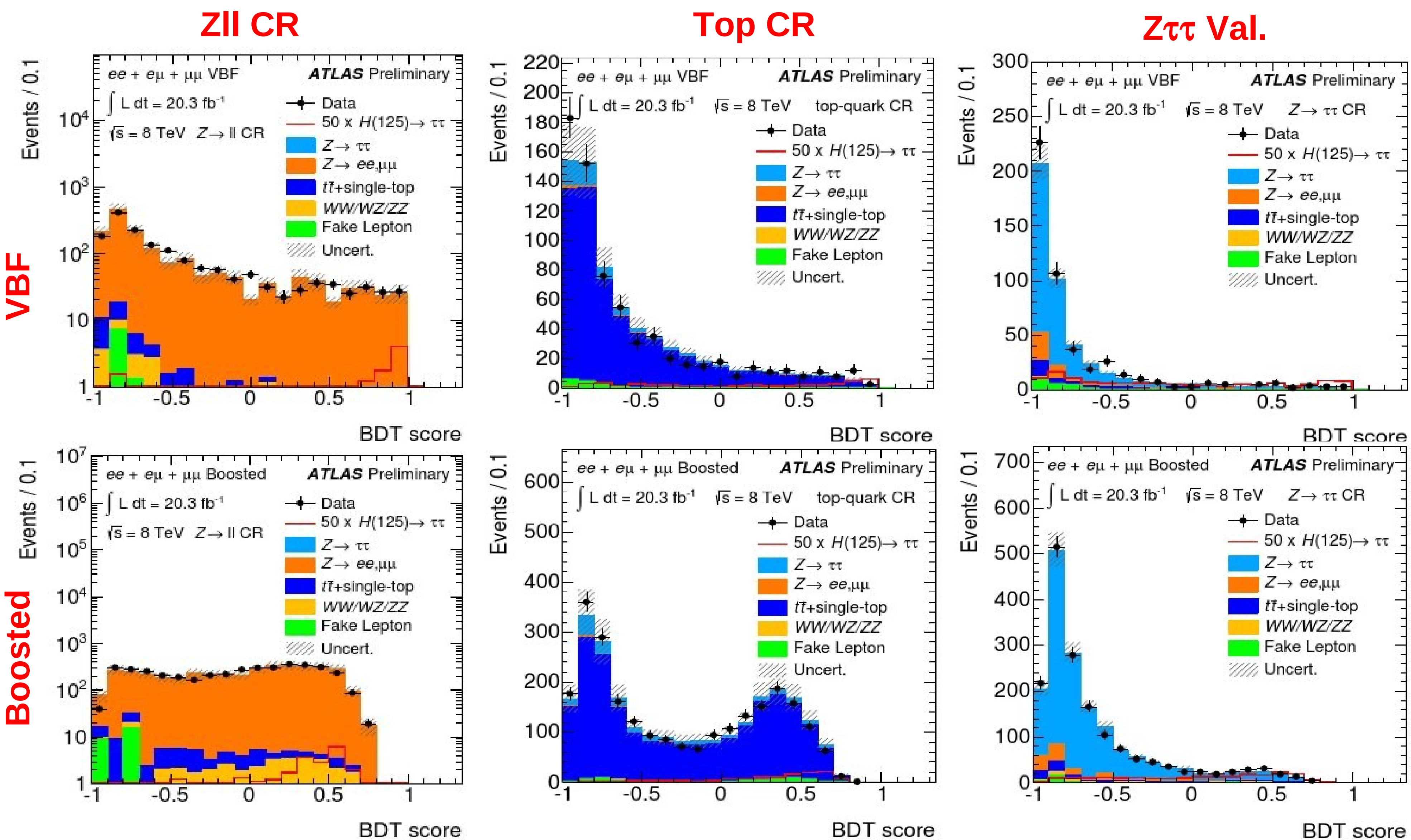


- $\tau_{had}\tau_{had}$ uses notOS events (charge1*charge2 ≥ 0) to get a QCD template

- $\Delta\eta_{\tau\tau}$ distribution in the 0-jet category events is used (via simultaneous fits) to dertermine the QCD and Zττ normalizations



BDT distributions in control/validation regions ($\tau_{lep}\tau_{lep}$)



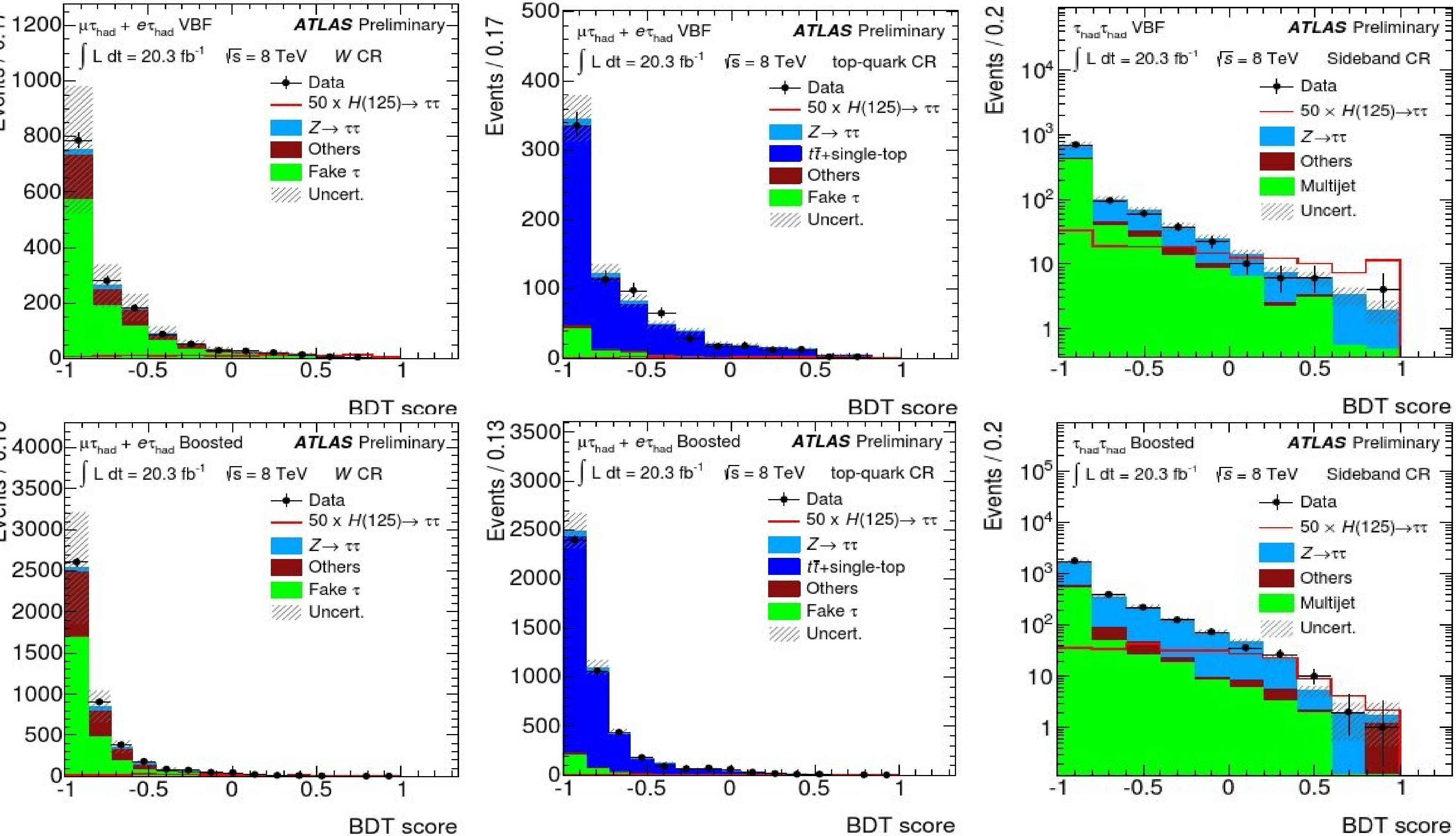
BDT has to be checked in control (independent of SR) and validation (part of SR) regions that are depleted of signal

BDT distributions in control/validation regions ($\tau_{lep}\tau_{had}/\tau_{had}\tau_{had}$)

W+jets CR | **Top CR** | **Zττ sideband**

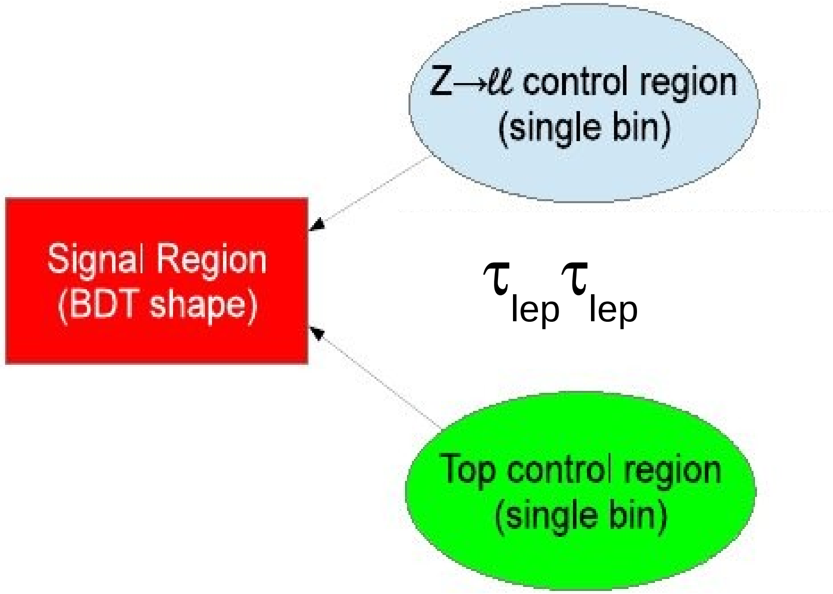
VBF

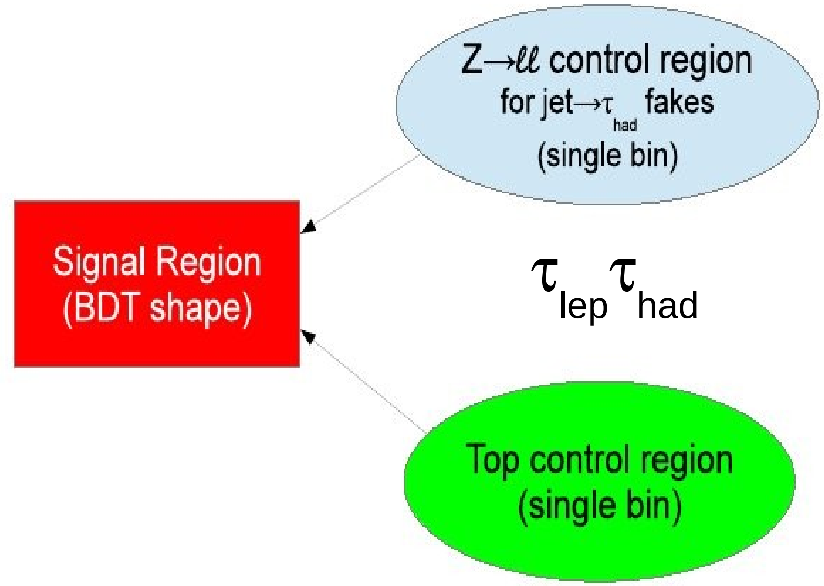
Boosted

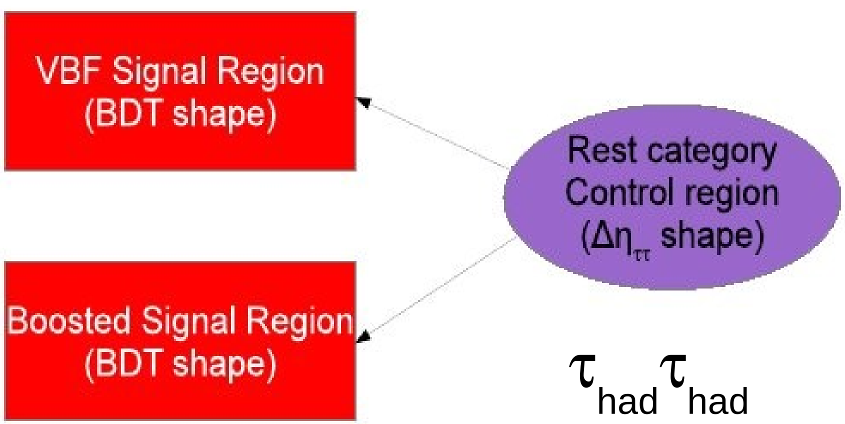


$\tau_{lep}\tau_{had}$ | $\tau_{lep}\tau_{had}$ | $\tau_{had}\tau_{had}$

Everything in a simultaneous fit





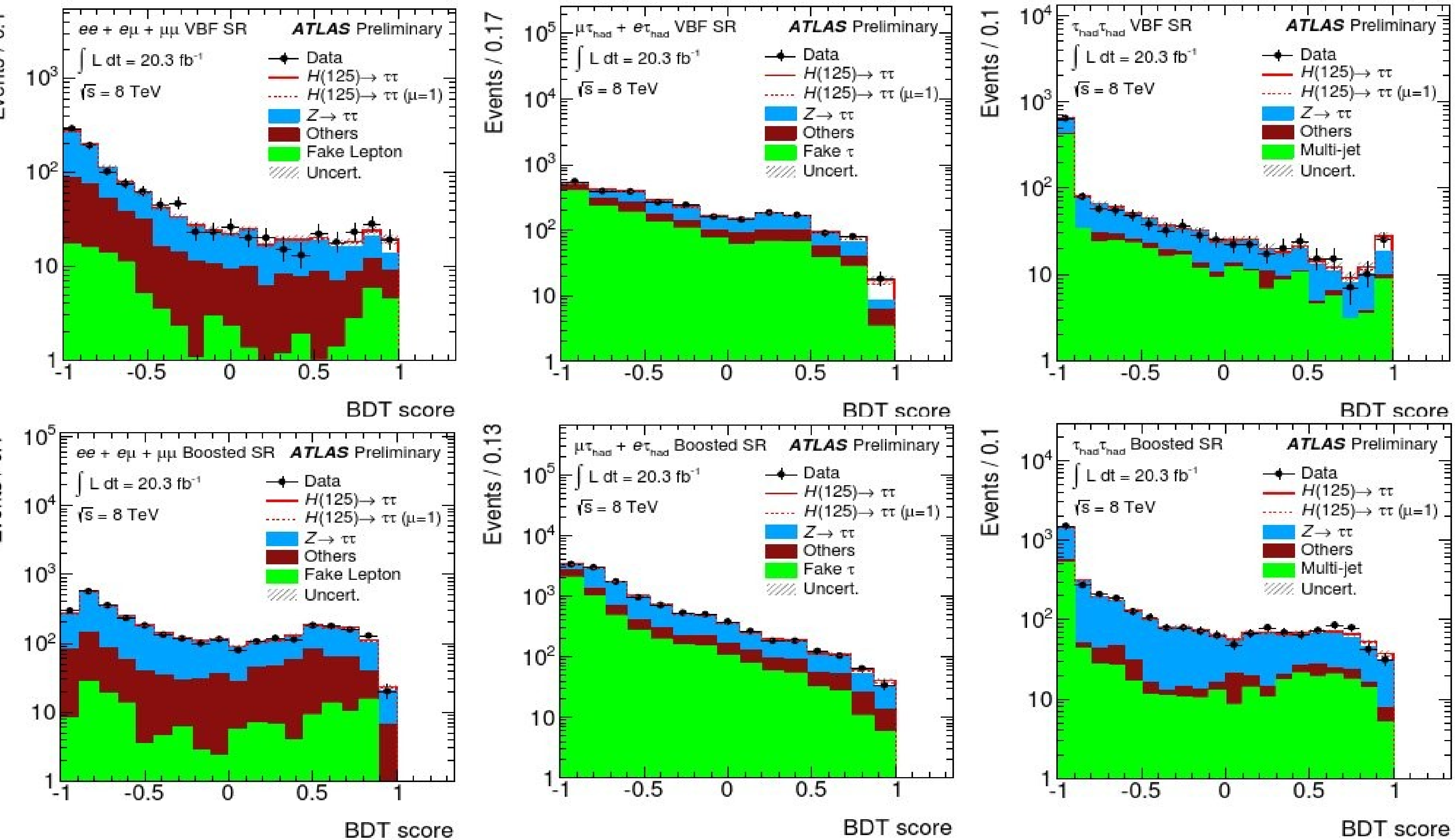


- Each channel has 2 signal regions (VBF and Boosted)
  - $\tau_{lep}\tau_{lep}$ has 2x2=4 control regions for Zll and top norms
  - $\tau_{lep}\tau_{had}$ has 2x2=4 control regions for Zll(jet->τ) and top norms
  - $\tau_{had}\tau_{had}$ has 1 control region for QCD norm

Main Systematics (in unit signal strength μ)

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

BDT distributions in the signal regions after the fit

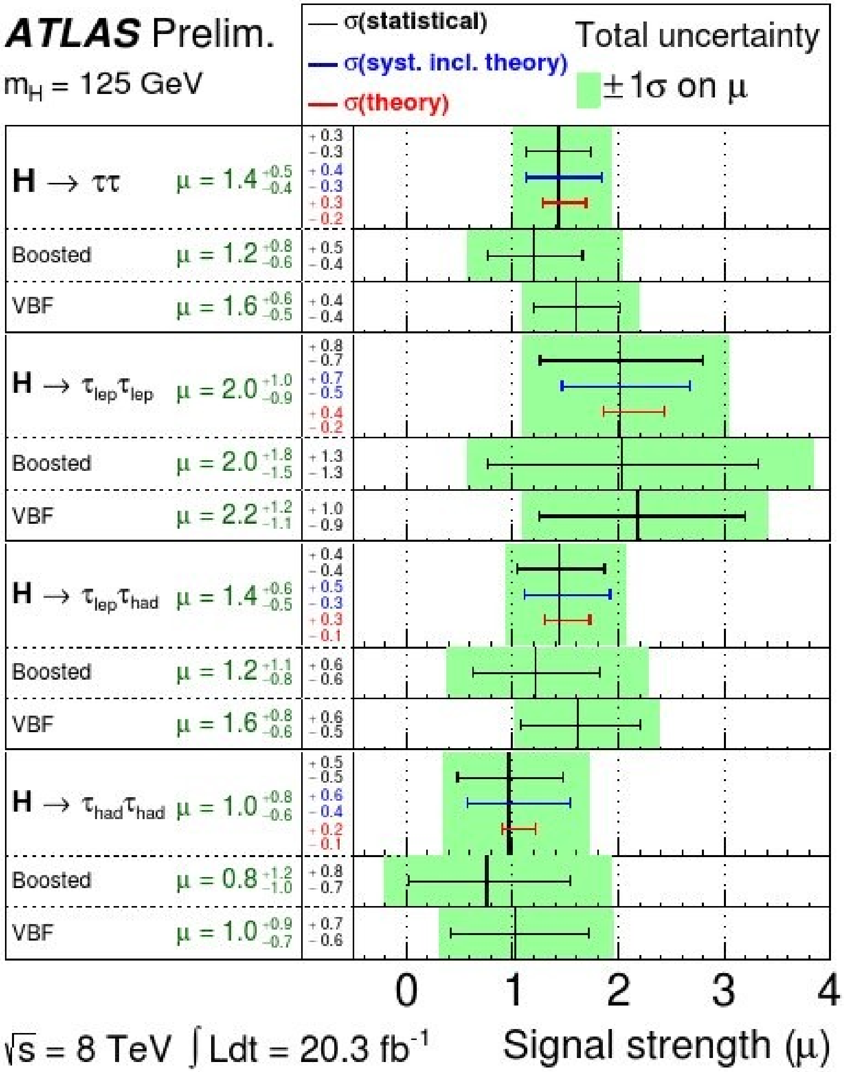


VBF

Boosted

$\tau_{lep}\tau_{lep}$ $\tau_{lep}\tau_{had}$ $\tau_{had}\tau_{had}$

Fitted yields and signal strength



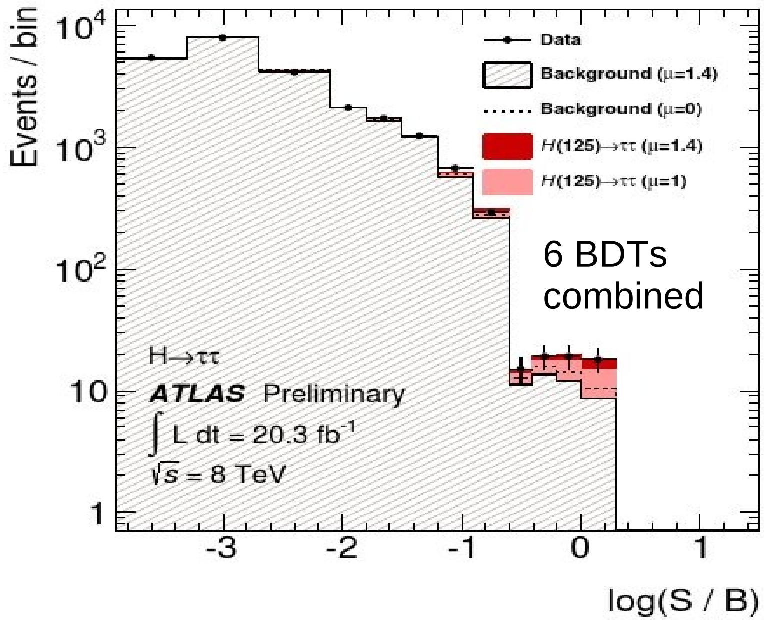
Yields in the most sensitive BDT bin:

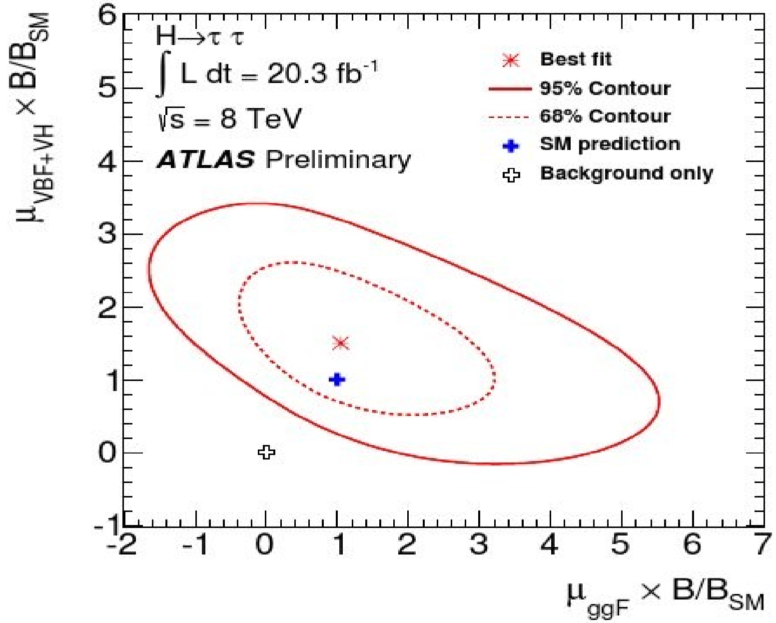
		LepLep	LepHad	HadHad
VBF	Signal	5.7±1.7	8.7±2.5	8.8±2.2
	Bkg	13.5±2.4	8.7±2.4	11.8±2.6
	Data	19	18	19
Boosted	Signal	2.6±0.8	8.0±2.5	3.6±1.1
	Bkg	20.2±1.8	32±4	11.2±1.9
	Data	20	34	15

The combined fitted signal Strength: $\mu = 1.4^{+0.5}_{-0.4}$

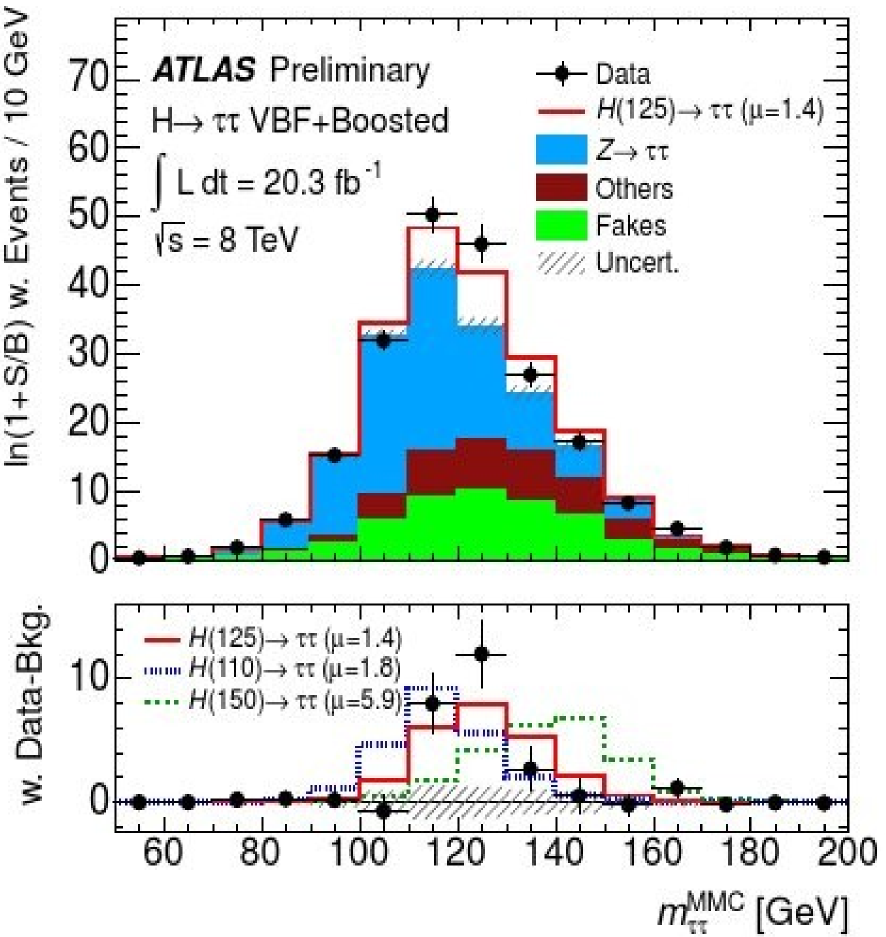
This corresponds to 4.1σ for 125 GeV (3.2σ expected)

Signal interpretation



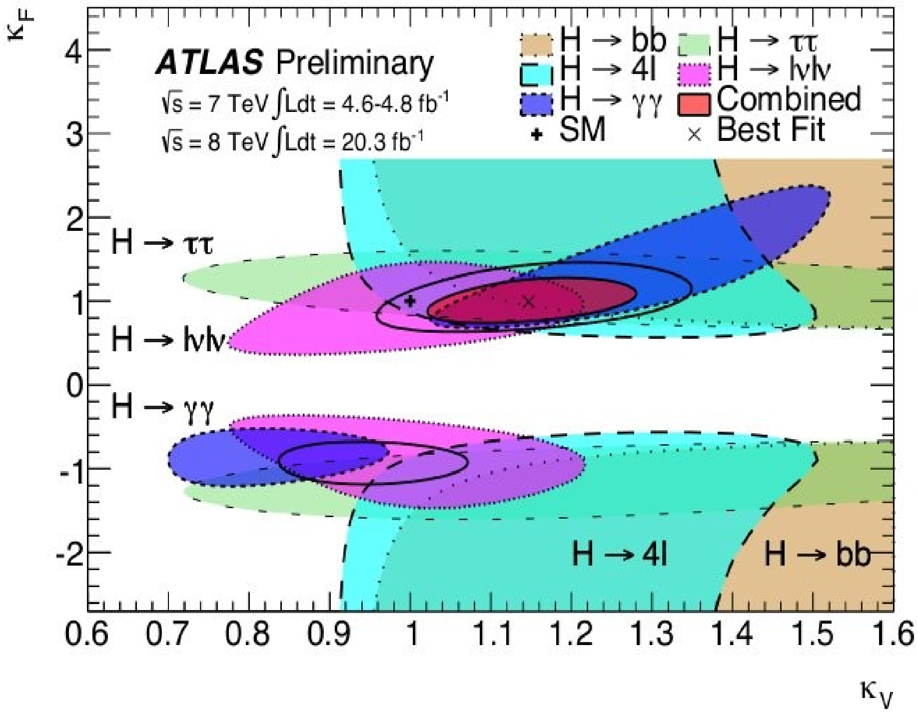


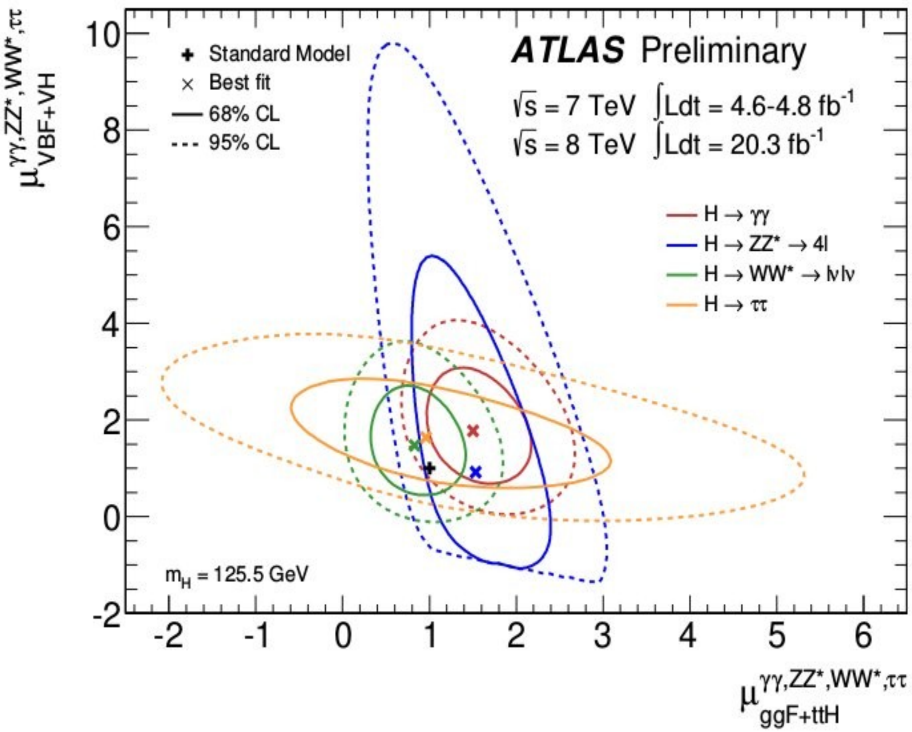
MMC weighted by ln(1+S/B)
in each BDT bin:



Signal is more compatible
with 125 GeV

H→ττ in the combination



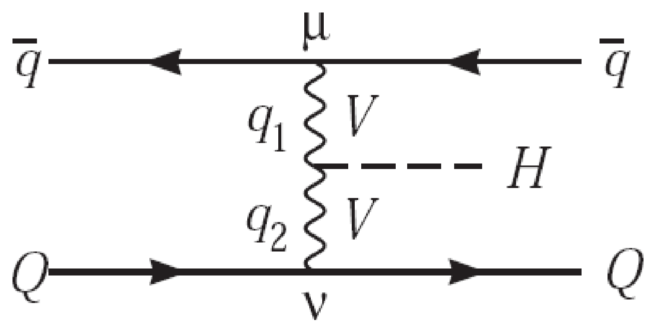


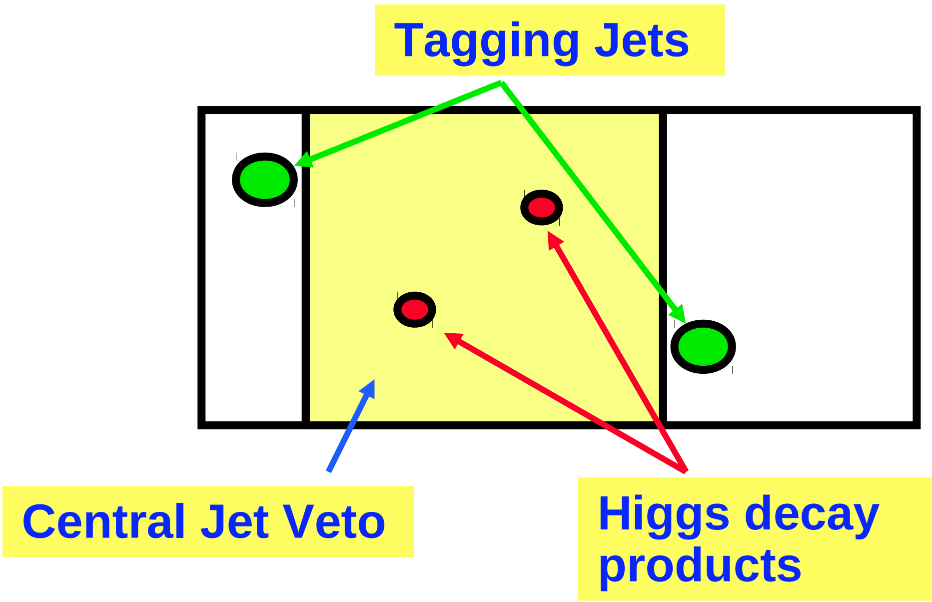
- κ_F (κ_V): scaling factor on the fermionic (bosonic) Higgs coupling
- μ_{VBF+VH} ($\mu_{ggF+tttH}$): scaling factor on the VBF+VH (ggF+tttH) producion
- H→ττ channel constrains mainly the fermionic coupling and the VBF production mode

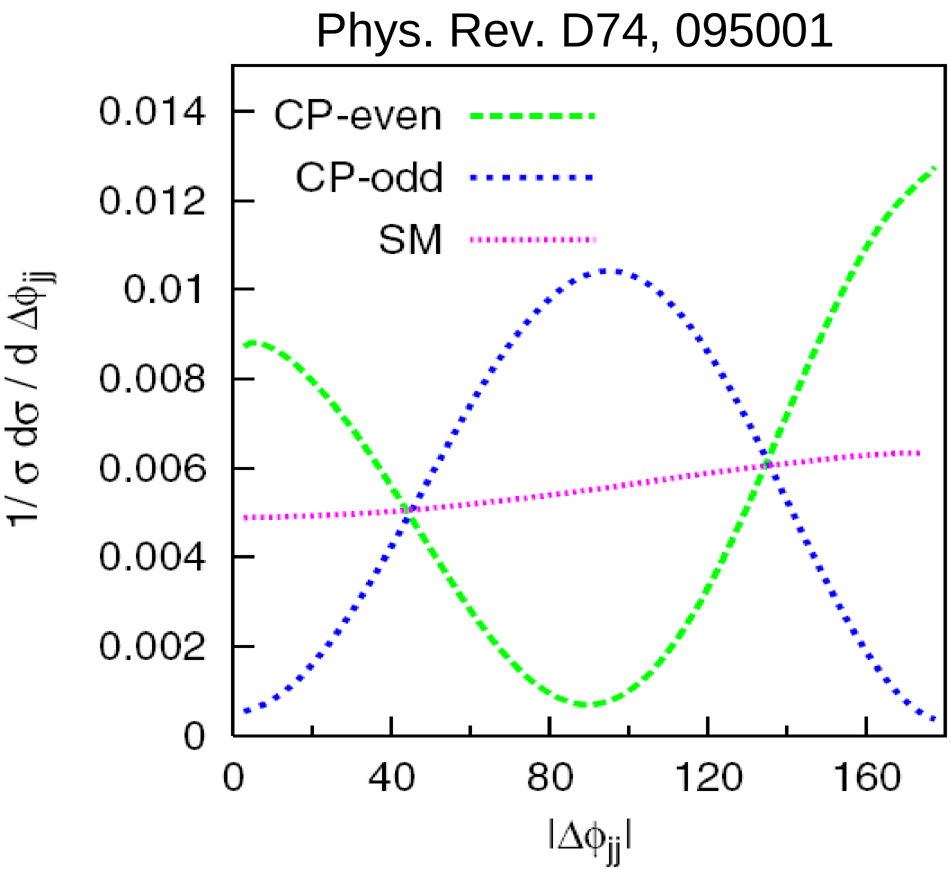
H→ττ CP on the HVV production side

The VBF channel is very powerful in revealing the HVV vertex structure:

$$T^{\mu\nu}(q_1,q_2) = \underbrace{a_1(q_1,q_2)\; g^{\mu\nu}}_{\text{SM}} + \underbrace{a_2(q_1,q_2)\;[q_1\cdot q_2 g^{\mu\nu} - q_2^\mu q_1^\nu]}_{\text{CP-even}} + \underbrace{a_3(q_1,q_2)\;\varepsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma}}_{\text{CP-odd}}$$



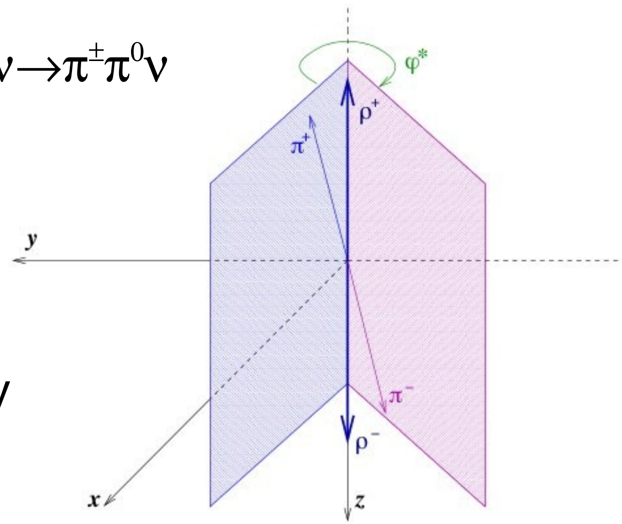




H→ττ CP on the decay side

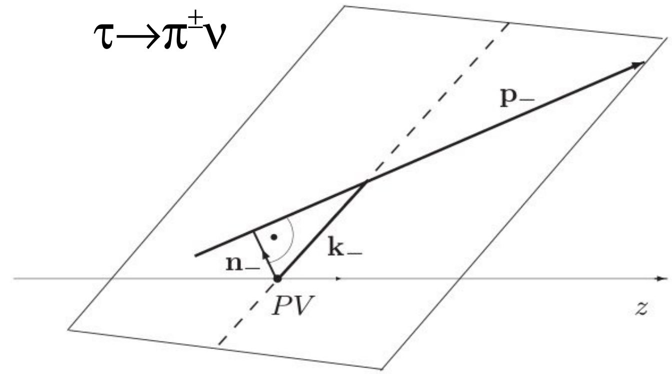
$$\mathcal{L}_Y = -N\left(\cos\phi\,\bar{\tau}\tau + \sin\phi\,\bar{\tau}i\gamma_5\tau\right)h$$

$\tau \to \rho\nu \to \pi^{\pm}\pi^{0}\nu$

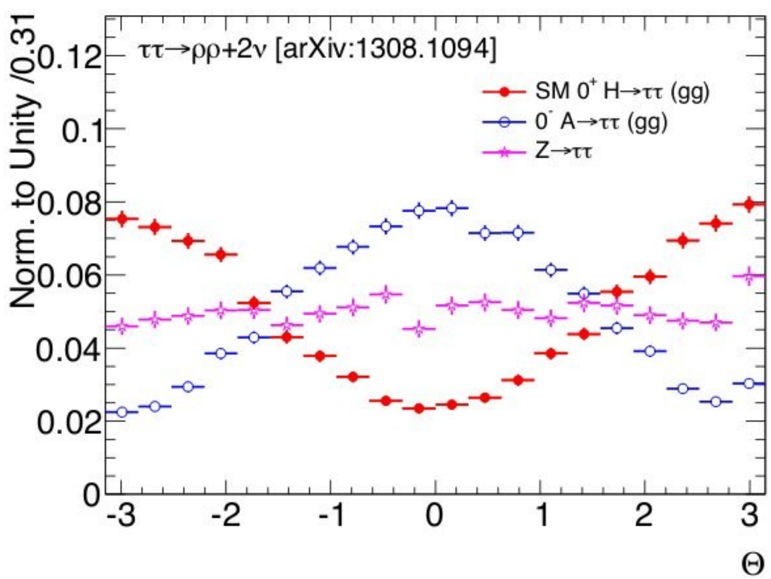


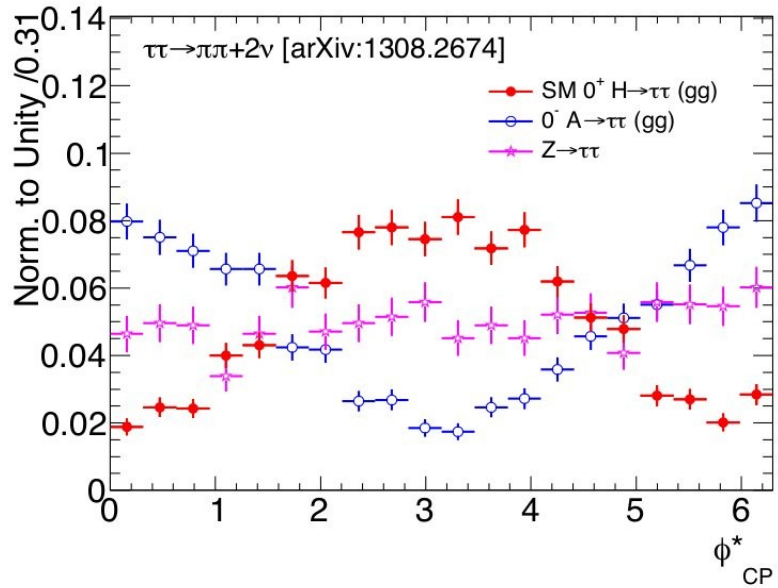
τ decay planes

$\tau \to \pi^{\pm}\nu$



Angle between impact paremeter vectors





- Both methods become difficult for boosted tau
- Correlation between production and decay (matrix element) ?

Summary and outlook

- The Higgs boson in its fermionic decay is close to discovery in the ditau mode (observed 4.1σ/expected 3.2σ). However, if we combine with the same results from CMS, it is more than 5σ already

- Boosted Decision Trees was used in the $H\rightarrow\tau\tau$ channel to enhance the search sensitivity

- The mass was found to be compatible with 125 GeV. The result better constrains Higgs fermionic and VBF couplings

- The channel will focus on the CP analysis with more luminosity expected in the coming years

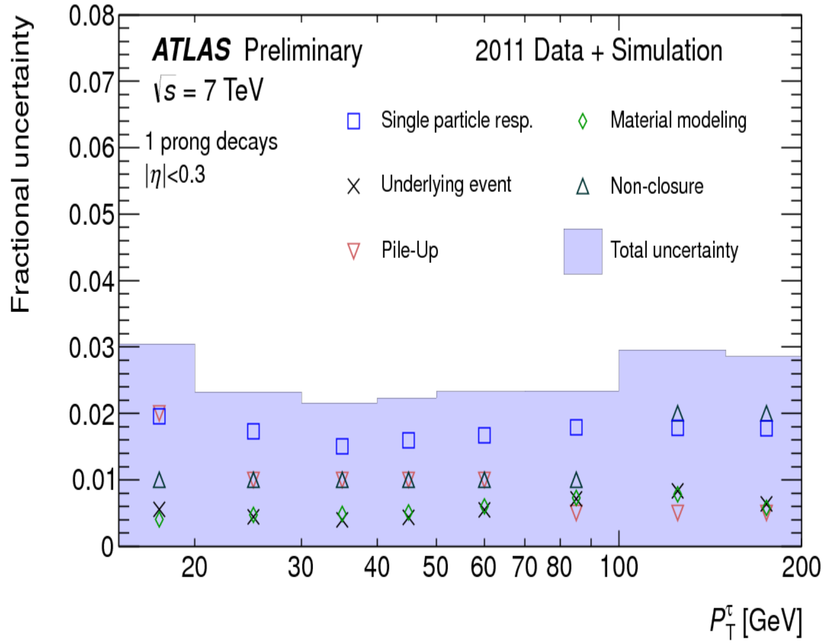
Backup

Tau reconstruction

- Seeded from AntiKt4 Local Hadron Calibration (LC) jets in |η|<2.5 (1/3-prongs). Ininial tau 4-vector defined by LC clusters within ΔR=0.2 of the jet barycenter

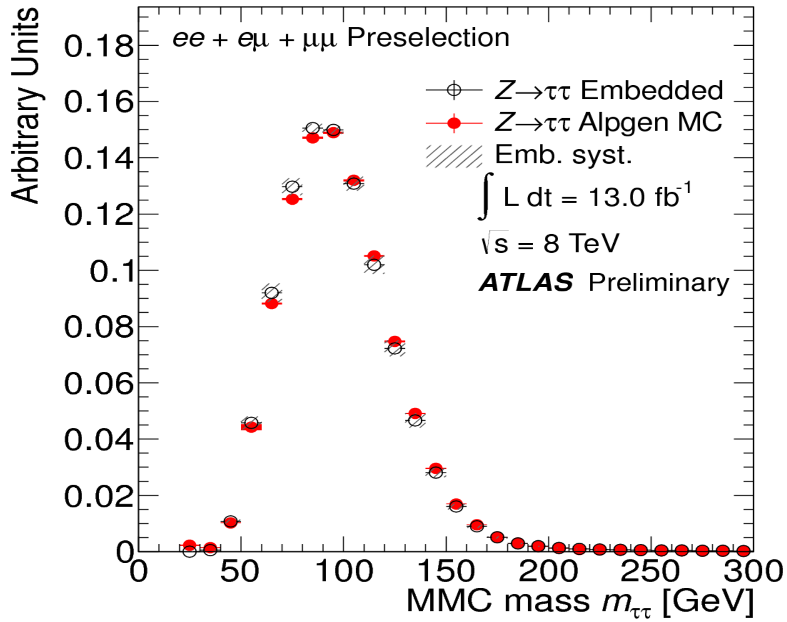
- The final tau energy scale is obtained by apply to the raw LC energy 1) a pileup offset subtraction, and 2) a response correction as a function of raw energy, η and number of primary vertices – final reconstruction is 1% within the true energy

- Tau energy scale (TES) uncertainty:
  - Low p and |η|<2.5: in-situ E/P measurements
  - High p and |η|<0.8: combined test beam (CBT) measurements
  - High p and |η|>0.8: MC simulation, but confirmed with the Z→ττ data analysis
  - Neutral pions: EM response from Z→ee and muons in Tile
  - Shower shape models, dead material, underlying event and pileup

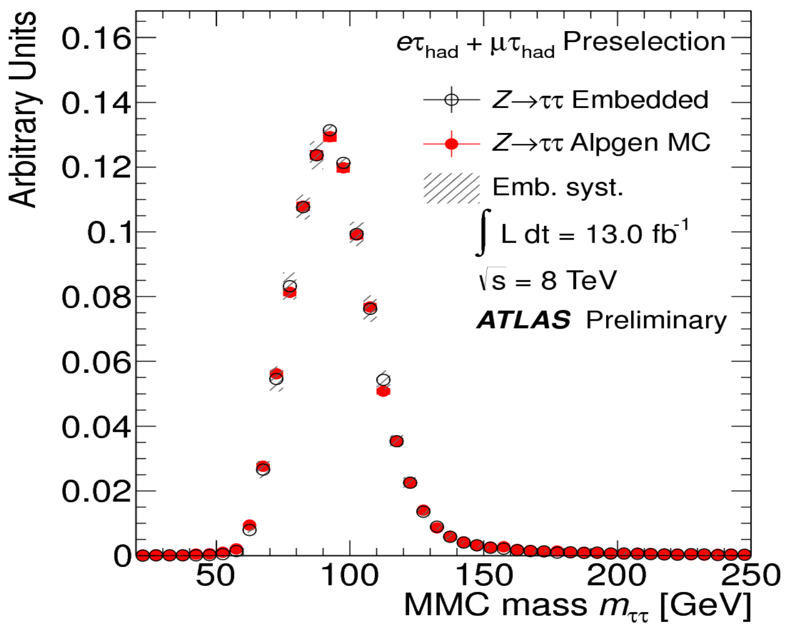


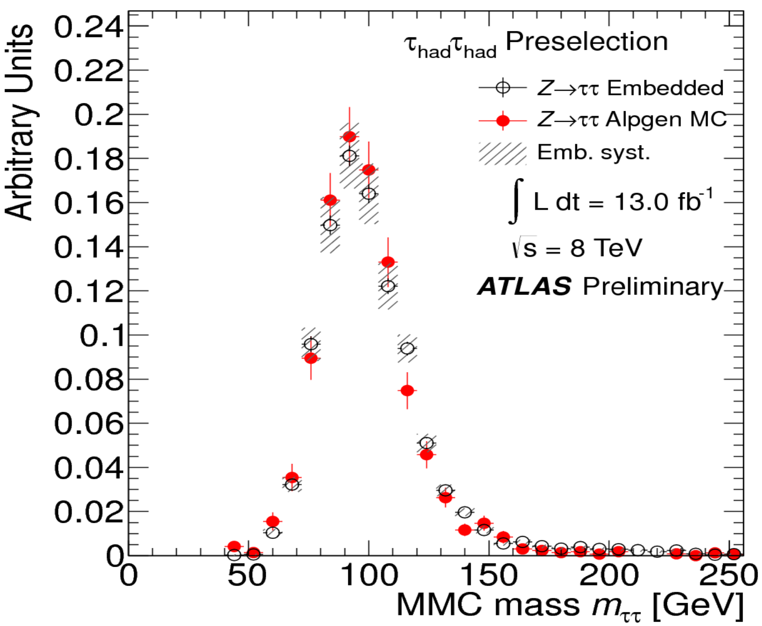
- Final TES uncertainty of ~3% is provided by the Tau Working Group

Embedding sample to model $Z \rightarrow \tau\tau$ background



Good agreement in the MMC mass is observed between embedding and MC in all 3 channels





Event display of a candidate VBF $\tau\tau \rightarrow e\mu$ event