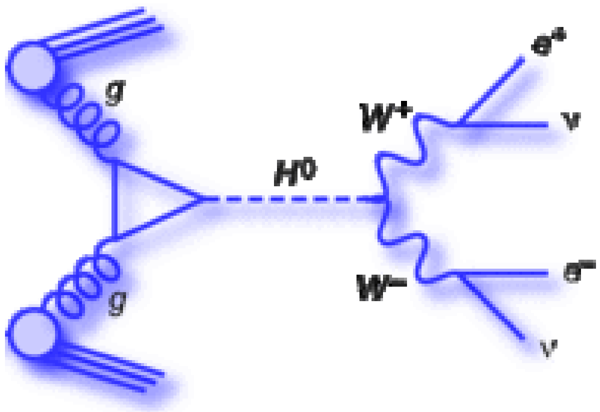
Spin characterization of the 125 GeV boson in the $H \rightarrow WW \rightarrow l\nu l\nu$ channel in ATLAS

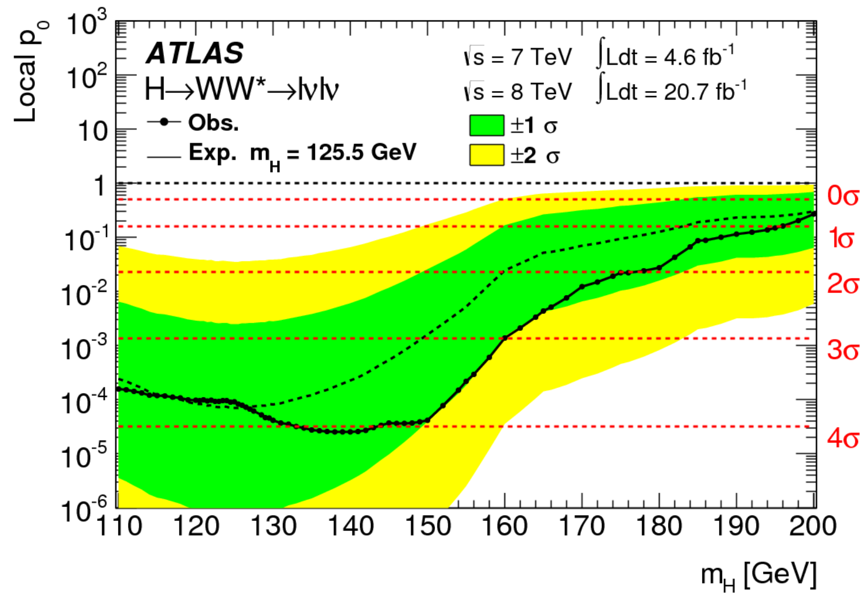
Lashkar Kashif

University of Wisconsin

DPF meeting, Santa Cruz, CA
August 16, 2013

Motivation





- Boson near 125 GeV discovered in 2012 in search for SM Higgs
- In *H* → *WW* → *lνlν* channel, signal significance ~3.8σ at m_H = 125 GeV
  - details in Doug and Ben's talks in this session
- Crucial to establish quantum numbers of this boson
- SM Higgs would have $J^P = 0^+$
  - can test alternate J^P hypotheses against 0^+ and see if they can be eliminated

Spin discrimination in $H \rightarrow WW \rightarrow l\nu l\nu$ channel

publication of 3-channel combination: **arXiv:1307.1432 [hep-ex], accepted by PLB**

Conference note in this channel only: **ATLAS-CONF-2013-031**

- In this channel, $J^P = 0^+$ is tested against 3 hypotheses: $J^P_{alt} = 1^+, 1^-, 2^+$

  - $J^P = 0^-$ not currently tested in this channel owing to low sensitivity

- $J = 1$: disfavored on theoretical grounds since boson decays to $\gamma\gamma$

  - nevertheless, important to test experimentally

  - produced via $q\bar{q}$ annihilation

- $J^P = 2^+$: graviton-like tensor with minimal couplings to SM particles

  - can be produced via gluon fusion (ggF) or $q\bar{q}$ annihilation, with different polarization states

  → experimental observables sensitive to polarization states, and hence to production modes

  - analysis done for 5 different $q\bar{q} \rightarrow 2^+$ production fractions f_{qq} in range [0, 1]: 0, 0.25, 0.5, 0.75, 1

Analysis strategy

- Shape-based analysis using multivariate method (BDT)

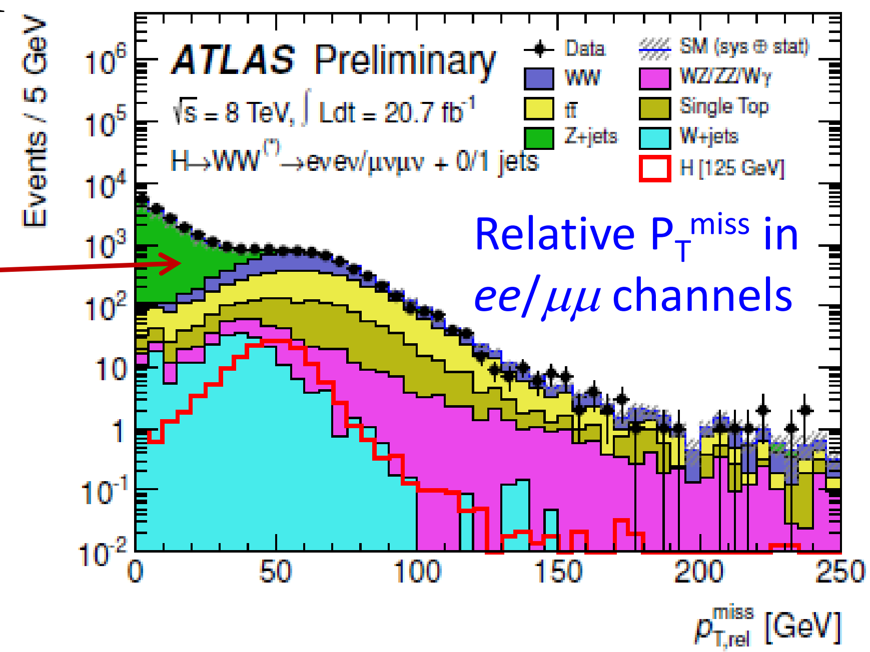
➢ For each hypothesis test, two separate BDTs trained:
one to separate $J^P = 0^+$ from bkg, and one to separate J^P_{alt} from bkg

○ BDT output in 2D plane used in likelihood fit to extract results

- Full 8 TeV dataset used in *lvlv* channel
- Large pile-up in 8 TeV data leads to significant mis-modeling of E_T^{miss}

  → increased bkg from *Z+jets*

  - only opposite lepton flavor channels ($e\mu$) are used
- Analysis in 0-jet final state only



Object and event selection

Lepton selection

- Both leptons in $|\eta| < 2.5$
- $p_T(\ell 1) > 25$ GeV, $p_T(\ell 2) > 15$ GeV

Jet selection for jet veto

- p_T (j) > 25 GeV for $|\eta| < 2.5$
- p_T (j) > 30 GeV for $2.5 < |\eta| < 4.5$

- Exactly two oppositely charged leptons in event
- Relative E_T^{miss} : $MET_{rel} > 20$ GeV

$$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi \geq \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} \cdot \sin\Delta\phi & \text{if } \Delta\phi < \pi/2 \end{cases}$$

- reduces *Z+jets* and multijet bkg

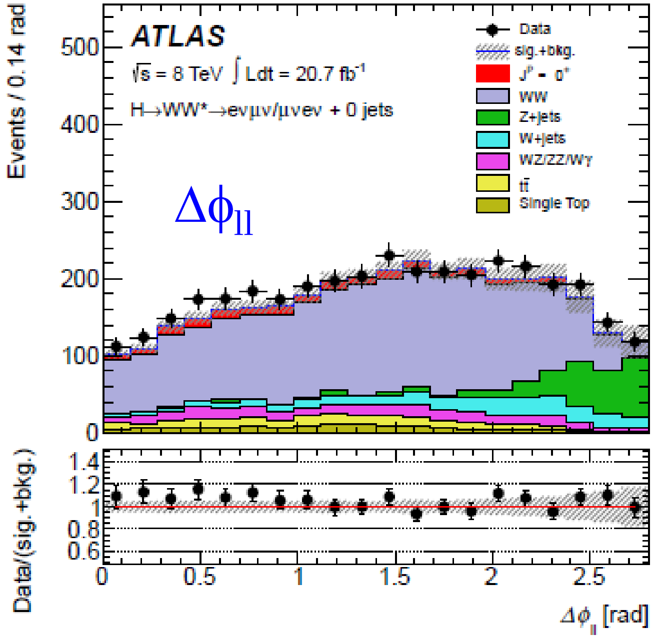
- Transverse momentum of dilepton system: $p_T^{ll} > 20$ GeV

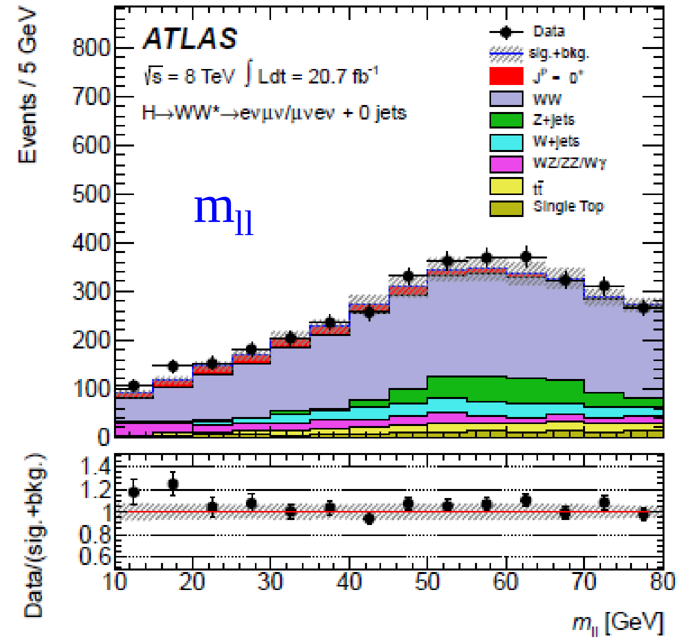
- reduces *Z+jets* bkg further

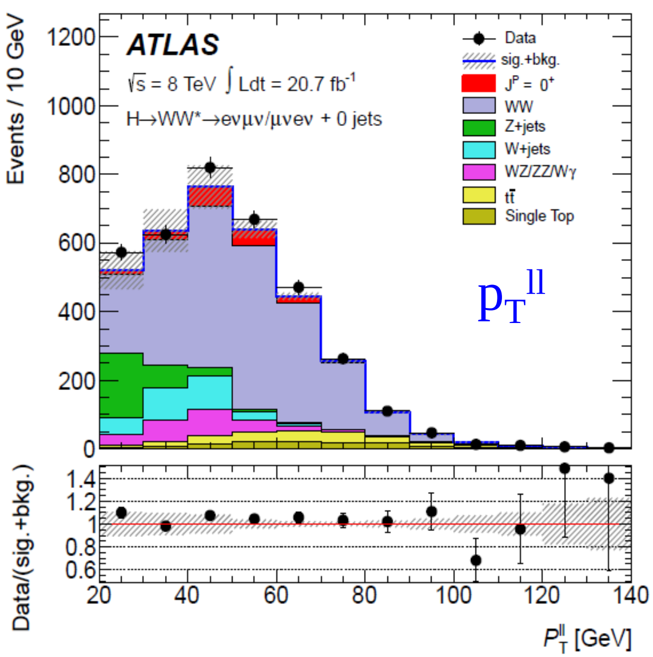
- Dilepton invariant mass : $m_{ll} < 80$ GeV
- Azimuthal separation of leptons : $\Delta\phi_{ll} < 2.8$

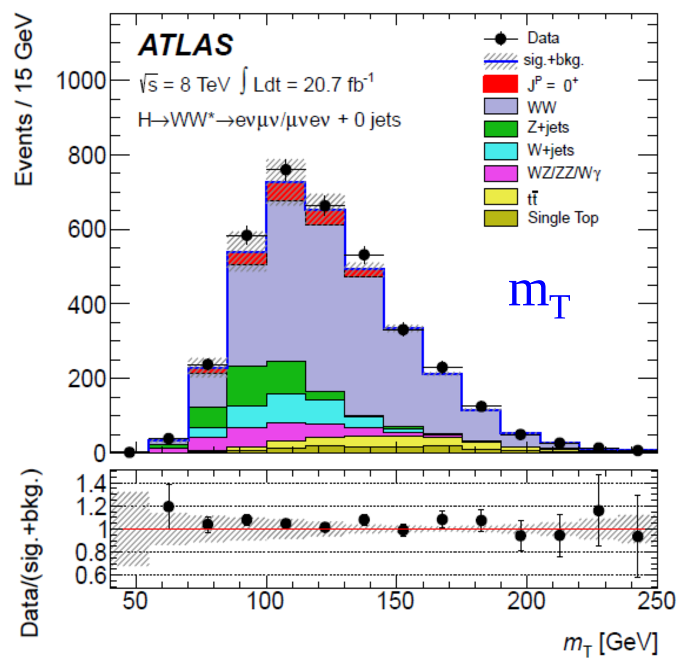
} Definition of signal region

Kinematic distributions after selection



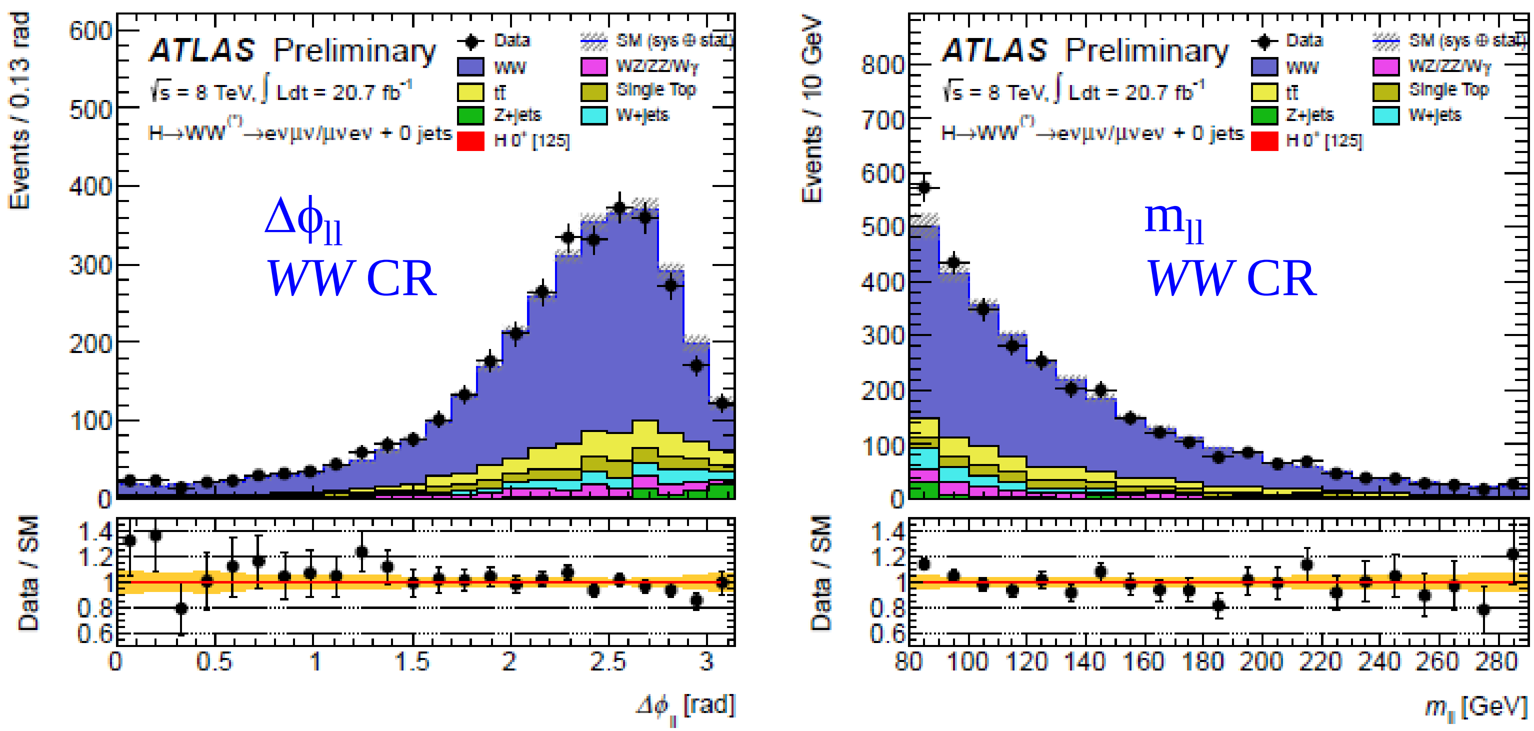






Background estimation: *WW* and *Z+jets*

- *WW* : data control region (CR) defined using m_{ll} > 80 GeV
  - no $\Delta\phi_{ll}$ requirement
  - non-*WW* contributions in CR subtracted off



- *Z+jets*: data CR defined using m_{ll} < 80 GeV and $\Delta\phi_{ll}$ > 2.8
  - no p_T^{ll} requirement

Background estimation: others

- Top: estimated using two CRs

  - one CR is all events after MET_{rel} selection

  - the other uses events after MET_{rel} selection that have at least one *b*-jet

- *W+jets*: fully data-driven; data CR defined by loosening selection criteria on one lepton (‘anti-identified’ lepton)

  - CR → SR transfer factor derived as ratio of fully identified to anti-identified leptons in inclusive dijet sample

- *Non-WW diboson*: small contributions, estimated from MC

Event yield in signal region

(Uncertainties are statistical only)

0^+	2^+_m	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	Z+jets	W+jets	Total Bkg.	Observed
170 ± 1	110 ± 1	2190 ± 20	230 ± 10	180 ± 10	120 ± 10	290 ± 20	280 ± 10	3280 ± 20	3615

Systematic uncertainties

➢ Experimental

- Uncertainties on jet energy scale and resolution are dominant
- Uncertainties on lepton energy scales and resolution, and on *W+jets* CR → SR transfer factor also contribute

➢ Theoretical

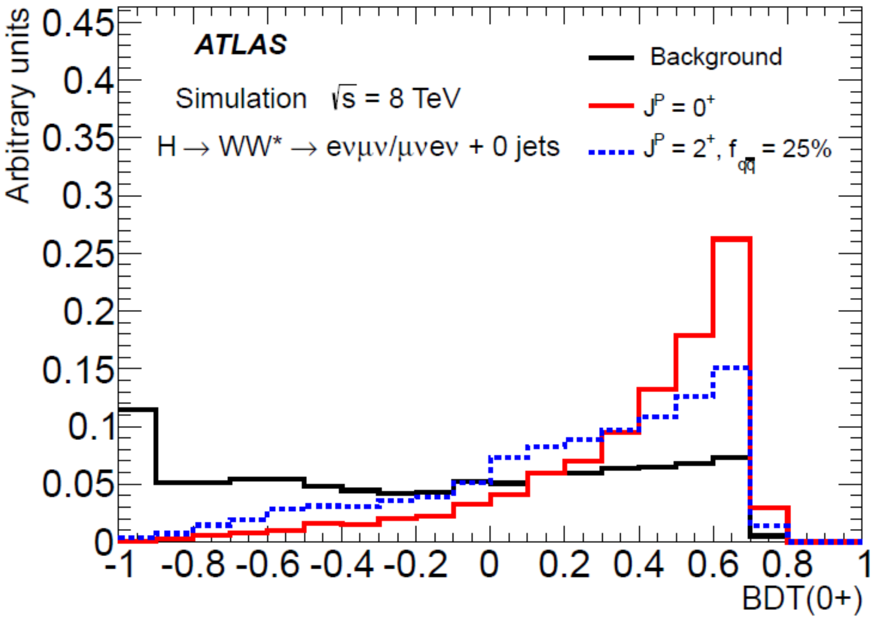
- Errors on shape and normalization of *WW* bkg crucial, since this is the dominant bkg
- For $J^P = 0^+$ vs 2^+ analyses, unknown p_T spectrum of 2^+ can lead to a shape uncertainty

  - studied by re-weighting p_T spectrum of $J^P = 2^+$ to that of $J^P = 0^+$ and repeating analysis → small impact on results

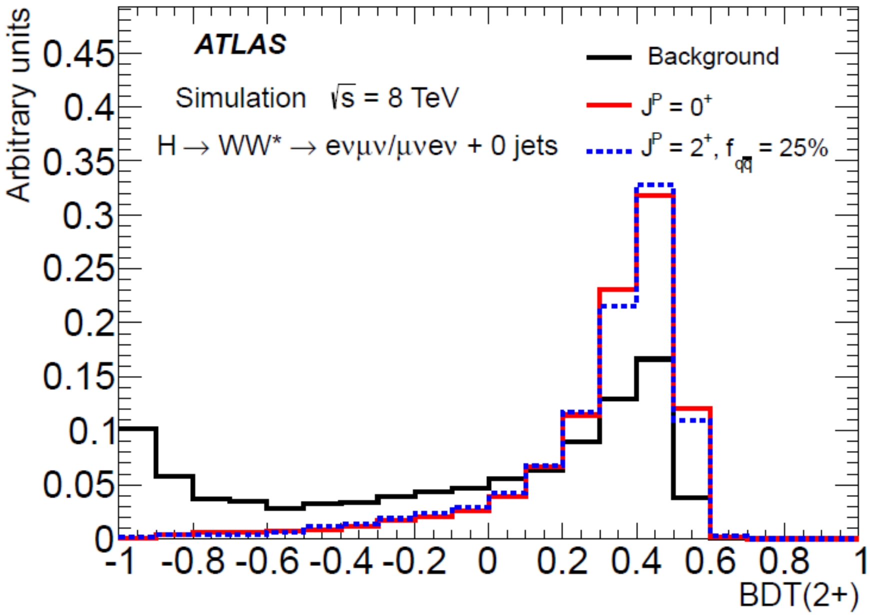
Source	Uncertainty (%)
Jet energy scale & resolution	± 9
WW normalisation, theory	± 9
W+jets fake factor	± 8
Lepton scale & resolution	± 6
Other backgrounds, theory	± 5
Pileup modelling	± 4
PDF model	± 4
E_T^{miss} scale & resolution	± 3

BDT analysis

- For each $J^p{}_{alt}$ test, the two BDTs are trained using the same 4 variables: $\Delta\phi_{ll}$, m_{ll}, m_T, $p_T{}^{ll}$
- In each training, respective signal hypothesis and all backgrounds used



BDT trained with $J^P = 0^+$ as signal. Output shown for $J^P = 0^+$, $J^P = 2^+$ ($f_{qq} = 0.25$) and total bkg



BDT trained with $J^P = 2^+$ ($f_{qq} = 0.25$) as signal. Output shown for $J^P = 0^+$, $J^P = 2^+$ ($f_{qq} = 0.25$) and total bkg

Statistical treatment

- Likelihood defined with the fraction of $J^P = 0^+$ signal as the parameter of interest ε

$$\mathcal{L}(\epsilon,\theta) = \prod_i^{N_{bins}} P(N_i|\epsilon \cdot S_i^{0^+}(\theta) + (1-\epsilon)S_i^{2_m^+}(\theta) + B_i(\theta)) \times \prod_j^{N_{sys}} \mathcal{A}(\tilde{\theta}_j|\theta_j)$$

- Since have no accurate knowledge of $J^P{}_{alt}$ production cross section, signal event rate is a floating parameter in fit

- The test statistic q is defined as a ratio of likelihoods

$$q = \log \frac{\mathcal{L}(J^P = 0^+, \hat{\hat{\mu}}_{0^+}, \hat{\hat{\theta}}_{0^+})}{\mathcal{L}(J^P_{\text{alt}}, \hat{\hat{\mu}}_{J^P_{\text{alt}}}, \hat{\hat{\theta}}_{J^P_{\text{alt}}})}$$

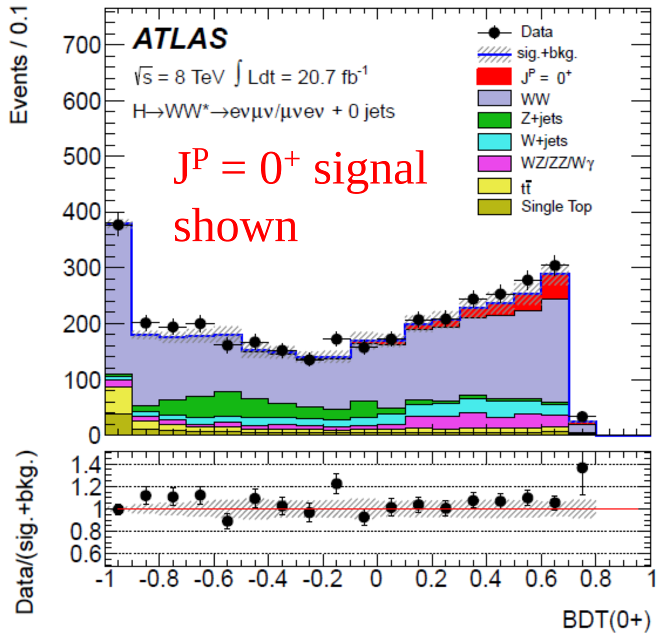
- Distributions of test statistic obtained using toy MC

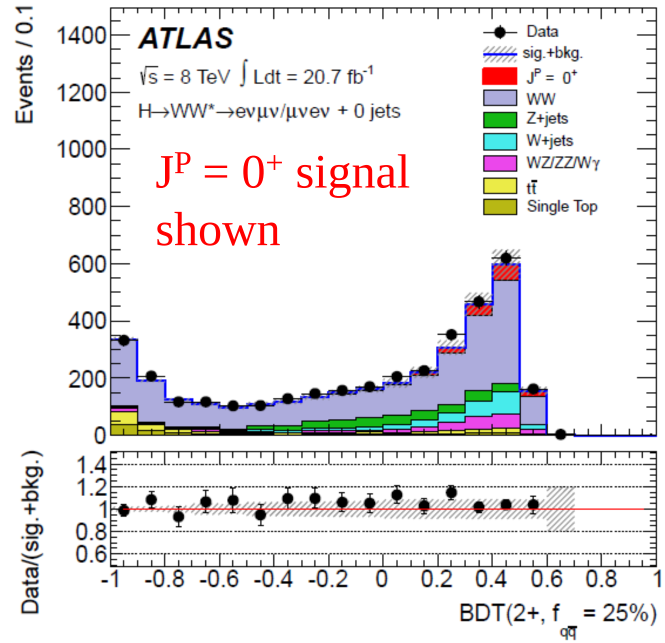
- Rejection of $J^P{}_{Alt}$ with respect to $J^P = 0^+$ evaluated using a CL_s quantity

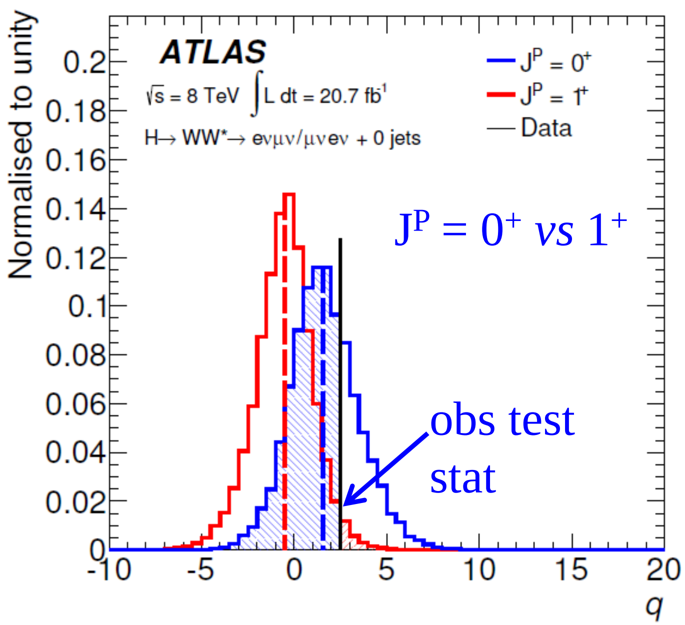
$$\text{CL}_s(J^P_{\text{alt}}) = \frac{p_0(J^P_{\text{alt}})}{1 - p_0(0^+)}$$

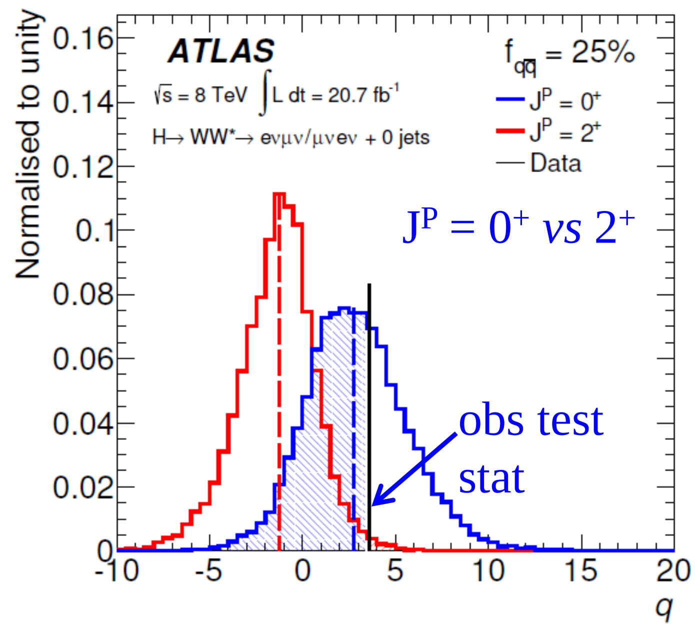
p_0: probability that other signal hypothesis fluctuates to look like hypothesis in question

Results: BDT output and toy distributions









Results: exclusion of J^P_{alt}

$J^P = 0^+$ *vs* 2^+

$f_{q\bar{q}}$	2^+ assumed Exp. $p_0(J^P = 0^+)$	0^+ assumed Exp. $p_0(J^P = 2^+)$	Obs. $p_0(J^P = 0^+)$	Obs. $p_0(J^P = 2^+)$	$CL_s(J^P = 2^+)$
100%	0.013	$3.6 \cdot 10^{-4}$	0.541	$1.7 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$
75%	0.028	0.003	0.586	0.001	0.003
50%	0.042	0.009	0.616	0.003	0.008
25%	0.048	0.019	0.622	0.008	0.020
0%	0.086	0.054	0.731	0.013	0.048

$J^P = 0^+$ *vs* 1^+

Channel	1^+ assumed Exp. $p_0(J^P = 0^+)$	0^+ assumed Exp. $p_0(J^P = 1^+)$	Obs. $p_0(J^P = 0^+)$	Obs. $p_0(J^P = 1^+)$	$CL_s(J^P = 1^+)$
$H \rightarrow WW^*$	0.11	0.08	0.70	0.02	0.08

$J^P = 0^+$ *vs* 1^-

Channel	1^- assumed Exp. $p_0(J^P = 0^+)$	0^+ assumed Exp. $p_0(J^P = 1^-)$	Obs. $p_0(J^P = 0^+)$	Obs. $p_0(J^P = 1^-)$	$CL_s(J^P = 1^-)$
$H \rightarrow WW^*$	0.06	0.02	0.66	0.006	0.017

Conclusion & outlook

- The $J^P = 0^+$ hypothesis for the 125 GeV boson has been tested in ATLAS against 3 alternative hypotheses in the $H \rightarrow WW \rightarrow l\nu l\nu$ decay channel
- With 21 fb^{-1} of 8 TeV data, CL_s exclusion:
  - $J^P = 2^+$: 95% or better for all $q\bar{q} \rightarrow 2^+$ production models
  - $J^P = 1^+$: 92%
  - $J^P = 1^-$: 98.3%
  - $J^P = 0^+$ hypothesis not excluded in any test

    \- boson looks very much a Higgs, thus the shift from 'Higgs-like' to 'SM-like' Higgs
- Coming next: tests for CP admixture in multiple channels
  - Analyses have already started, results expected for next round of publications

Backup

Background-subtracted BDT output in data

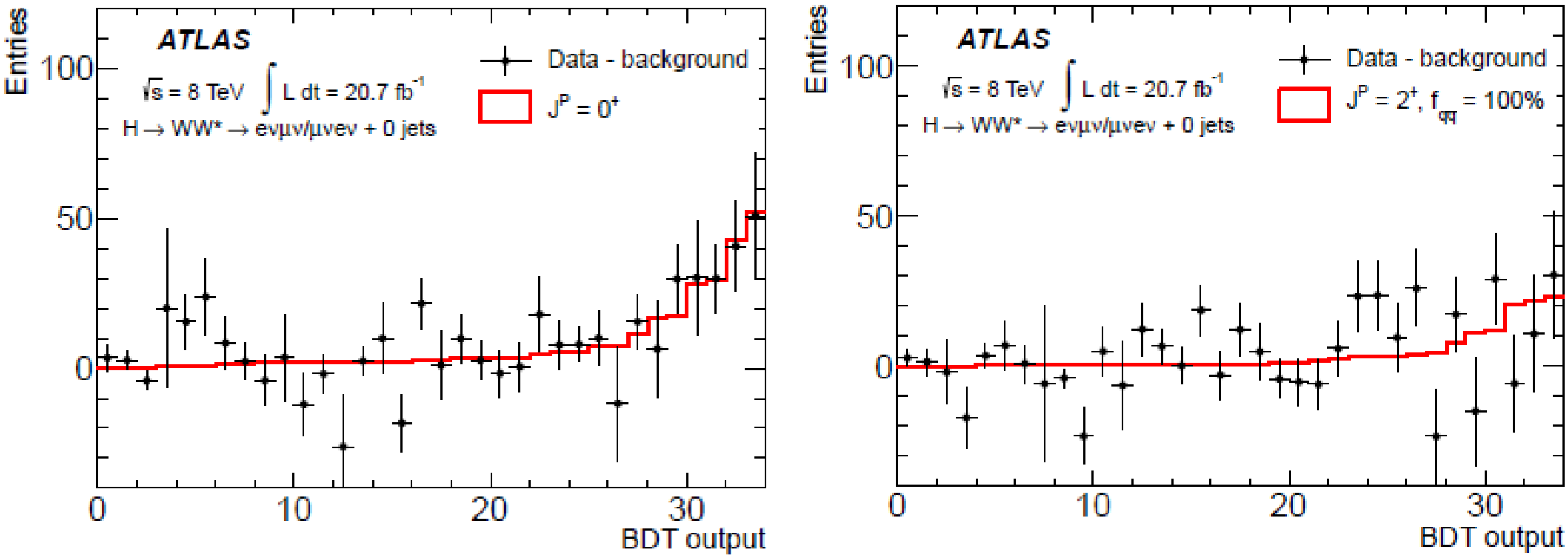


Figure 6: One-dimensional distributions of the outputs of the BDT for the $H \rightarrow WW^*$ channel after background subtraction, using best-fit values for (a) $J^P = 0^+$ and (b) $J^P = 2^+$ with $f_{q\bar{q}} = 100\%$ hypotheses. In each case, the two-dimensional distribution of the two classifiers is remapped into a one-dimensional distribution, with the bins reordered in increasing number of expected signal events. Empty bins, defined as bins with expected content below 0.1, are removed.