Search for Higgs pair-production

decaying to pairs of b quarks and τ leptons

EMILY GRAHAM

UNIVERSITY OF LIVERPOOL, ATLAS COLLABORATION

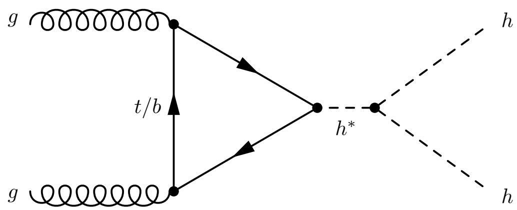
IOP JOINT HEPP AND APP ANNUAL CONFERENCE

27TH MARCH, 2018

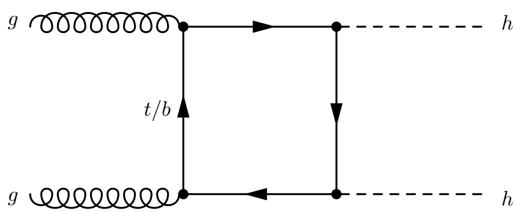
Overview

1. Introduction to di-Higgs → 2. Analysis Strategy → 3. Expected sensitivity

HIGGS PAIR-PRODUCTION



1. Higgs self-interaction



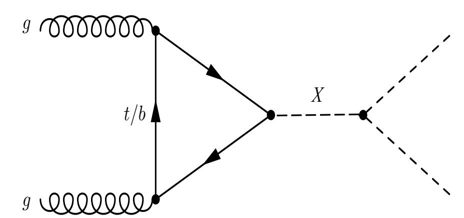
2. Virtual fermion (top) loops

NON-RESONANT

- Two *gg*-fusion processes in the Standard Model (SM)
- Diagram 1 provides a probe of the Higgs trilinear self-coupling λ_{hhh}
  - Destructive interference between two diagrams → small cross-section
    - $\sigma_{hh} \sim 33.49$ fb @ $\sqrt{s} = 13$ TeV
  - Possible BSM enhancements include composite Higgs models

RESONANT

- Randall-Sundrum graviton (RSG), *G*
  - Spin-2 Kaluza-Klein excitations in the bulk RS model
  - $c = \frac{k}{\bar{M}_{\mathrm{Pl}}} = 1, 2$
- Two Higgs Doublet Model (2HDM) Higgs, *H*
  - Heavy CP-even (spin 0) neutral Higgs, *H*
- 260 - 1000 GeV



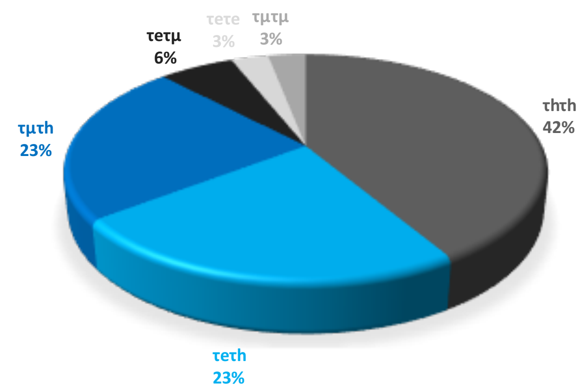
3. Resonant production

Higgs Pair-Production

	bb	WW	$\tau\tau$	ZZ	$\gamma\gamma$
bb	33%				
WW	25%	4.6%			
$\tau\tau$	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
$\gamma\gamma$	0.26%	0.10%	0.029%	0.013%	0.0053%

Decay Channels

- $hh \rightarrow b\bar{b}\tau^{+}\tau^{-}$ decay channel
  - Third-largest observable branching ratio, $\mathrm{BR} \sim 7.4\%$
  - Relatively low QCD background compared to *bbbb*
- Two τ decay channels are considered:
  - $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ (lep = e, μ): BR = 45.8%
  - $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$: BR = 41.9%
- This talk will focus on $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$. $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ follows same general strategy.



Analysis Strategy

Event Preselection

- Lowest un-prescaled single lepton triggers with $p_T^{\text{lep}} > 25 - 27$ GeV (period dependent)
  - Exactly **1 e/μ** and **1 "medium" hadronic τ** of opposite sign and $\geq$ **2 central jets**
    - $p_T^{\tau} > 20$ GeV, $p_T^{\text{jet}} > 45, 20$ GeV

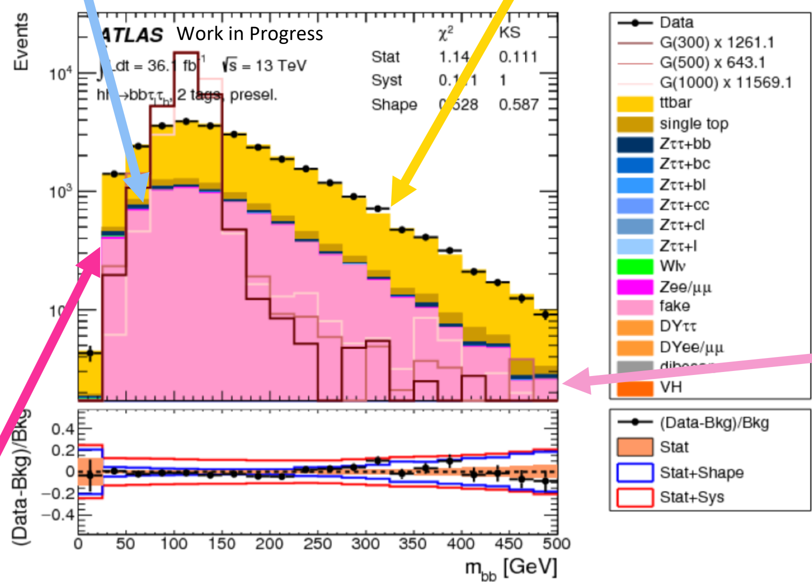
Strategy

- 2015+2016 data
- BDT used to discriminate between signal and background (giving up to a 50% improvement in sensitivity)
- 2 b-tag SR used for training and analysing BDT
- 1 b-tag, high-m_T and SS 2 b-tag CRs used to validate modelling
- Blinding: SR BDT score is blinded from the right-hand-side (higher score region) towards the left until 90% of signal is covered

BACKGROUNDS

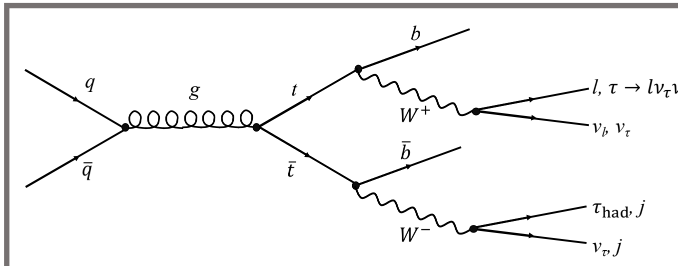
$Z \to \tau\tau$ (+jets)*

$t\bar{t}$ (real τ_{had})*



All fake τ_{had}

$Z \to ee/\mu\mu$ (+jets)



- In 2 b-tag signal region, $\boldsymbol{t\bar{t}}$ is $> 90\%$ of background (**real** and **fake** τ_{had})

- Fake τ_{had} = all backgrounds where jet misidentified as τ_{had}
- Use data-driven **combined fake-factor** method for fake backgrounds (all others from MC)

* These backgrounds are normalised to data in fit

Fake Backgrounds

- Any background where a jet is misidentified as a hadronic τ
- This is poorly modelled by Monte Carlo → use data-driven method for estimation
- Calculate a 'fake factor' separately for each fake-dominated process for 1- and 3-prong τs, parameterised in terms of $p_T(\tau_h)$

Fake Factor Method

	SR cuts (eg isol lep)	CR cuts (eg !isol)
ID τ	SR	CR (ID τ)
!ID τ	Template	CR (!ID τ)

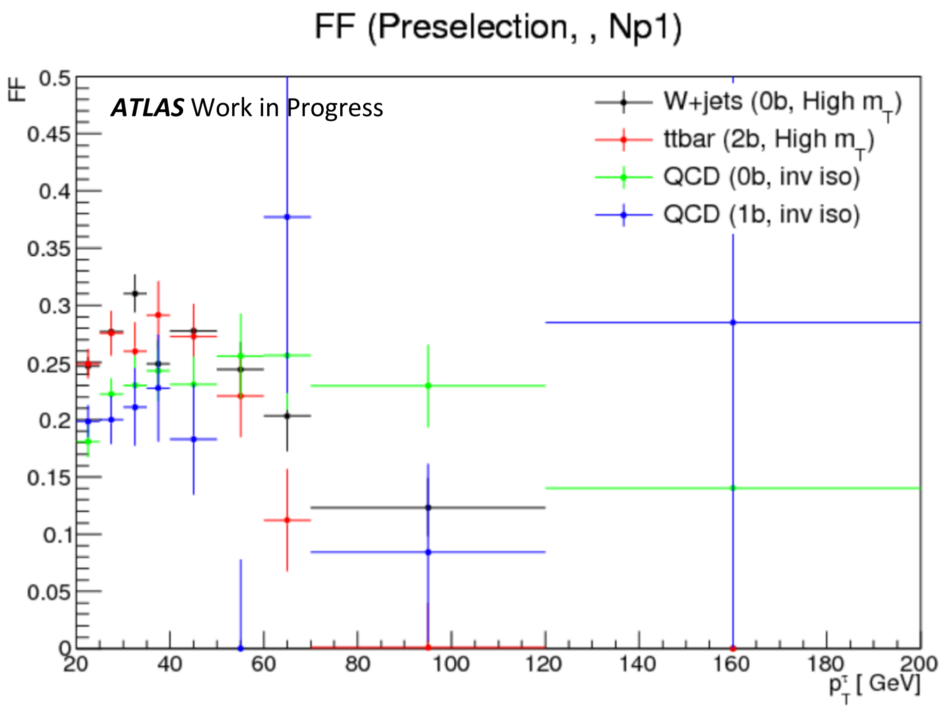
- Define control regions for each process:
  - ***$t\bar{t}$*:** $m_T^{l\nu} > 40$ GeV, 2 *b*-tag
  - ***W*+jets:** $m_T^{l\nu} > 40$ GeV, 0 *b*-tag
  - **QCD multi-jet:** inverse lepton isolation, 1 *b*-tag

Number of anti-τ events (those which fail τ ID but pass looser selection) with the real τ contribution (MC) subtracted

$$N_{\mathrm{Bkg}}^{\mathrm{Est.}} = \left(N_{\mathrm{data,SR}}^{\mathrm{anti}-\tau} - N_{\mathrm{true,SR}}^{\mathrm{anti}-\tau}\right) \times \mathrm{FF}_{\mathrm{comb}}$$

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

Fake Backgrounds



- Combined FF is derived by:

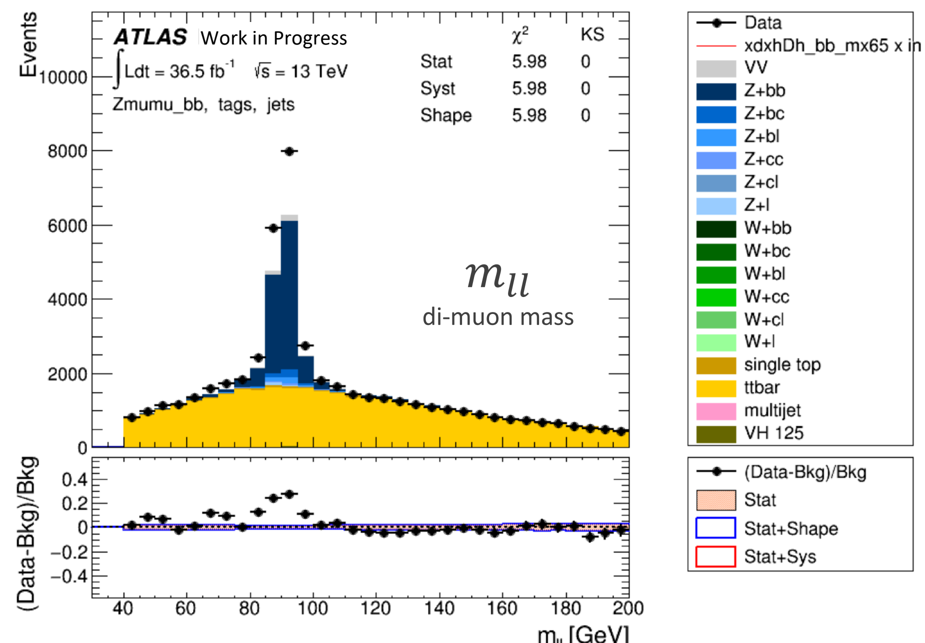
$$FF(\text{comb}) = FF(\text{QCD}) \times r_{\text{QCD}} + FF(W/t\bar{t}) \times \left(1 - r_{\text{QCD}}\right)$$

- r_{QCD} (also a function of p_T^τ) is defined as the fraction of fakes from multi-jets in the anti-τ region

$$r_{\text{QCD}} = \frac{N(\text{multijet,}\quad \text{data})}{N(\text{data}) \; - N(\text{true}\ \tau_{\text{had}},\quad \text{MC})}$$

Z+HF and True $t\bar{t}$ Normalisation

- *Z*+HF production cross-section not well described by MC:
  - Normalise these processes to data in a CR
  - Use $Z \rightarrow \mu\mu$ + HF data events
  - Include in final fit in order to get normalisation SF
  - m_{ll} distribution re-binned to one bin to fit only normalisation (not shape)
  - Z+HF SF ~ 1.4
  - Modelling uncertainties
    - ME variation
    - Vary factorisation/renormalisation



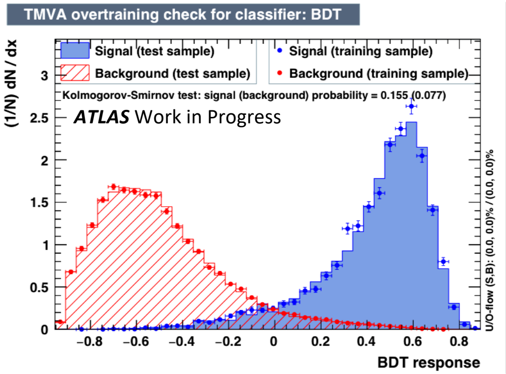
Before normalisation

- True $t\bar{t}$ normalisation allowed to float in fit and constrained by low BDT:
  - $t\bar{t}$ SF ~1
  - Modelling uncertainties from varying matrix element, parton shower and ISR/FSR

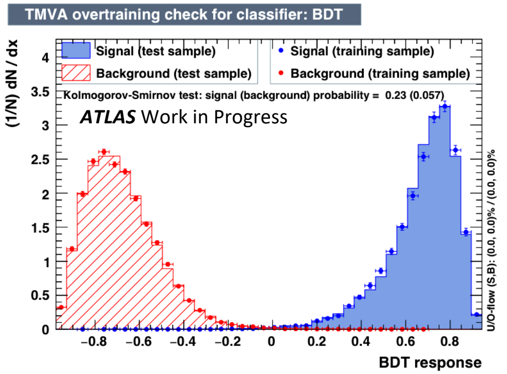
Boosted Decision Tree

Strategy

- BDT used to separate signal from background
- Trained on MC after preselection cuts and 2 *b*-tag requirement
- Signal trained against $t\bar{t}$ background only (true and fake τ components from MC)
- Resonant analysis:
  - Separate training for each mass point and signal model
  - Combine signal sample for each mass with signal masses either side (e.g. for 300 GeV use 275, 300, 325 GeV)
- Used as final discriminant for fit and limit setting



1. 2HDM 'even' BDT training, 260 GeV



2. 2HDM 'even' BDT training, 1000 GeV

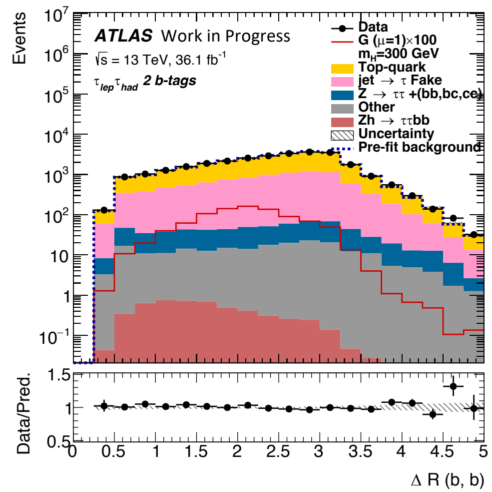
BDT Training

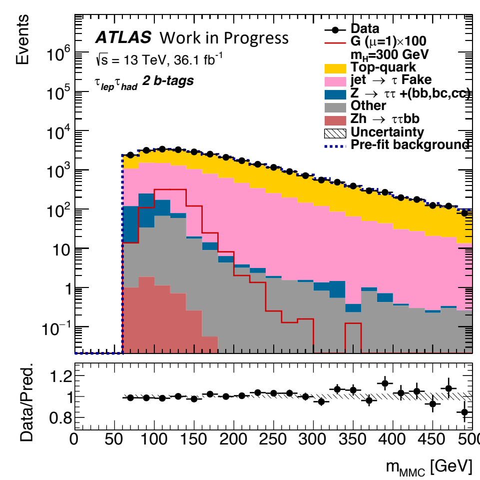
- MC is split by odd/even event numbers
- This allows for training on one half of MC and testing on other
- Then apply 'odd' training to 'even' data sample and vice versa
- Plots show training and test samples for signal and background

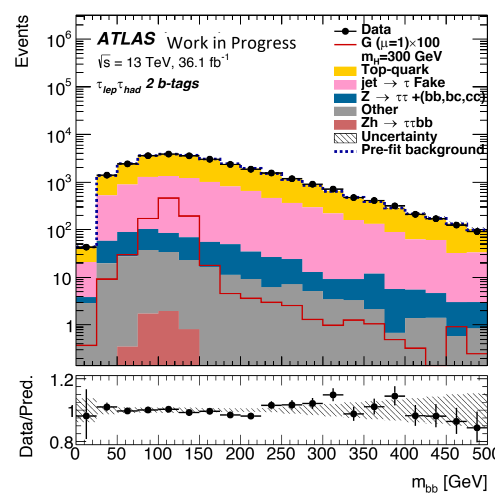
Boosted Decision Tree

- BDT input variables shown to be well modelled
- Non-resonant variable list is a subset of resonant list
- The least discriminating of any pairs of variables with high correlation co-efficient was removed from list

Variable	Lep-had (res)	Lep-had (non-res)
m_{HH}	✓	✓
m_{MMC}	✓	✓
m_{bb}	✓	✓
$\Delta R(\tau,\tau)$	✓	✓
$\Delta R(b,b)$	✓	✓
E_T^{miss}	✓	
E_T^{miss} ϕ cent	✓	
m_T^{lv}	✓	✓
$\Delta\phi(H,H)$	✓	
$\Delta p_T(l,\tau)$	✓	
p_T^{B2}	✓	





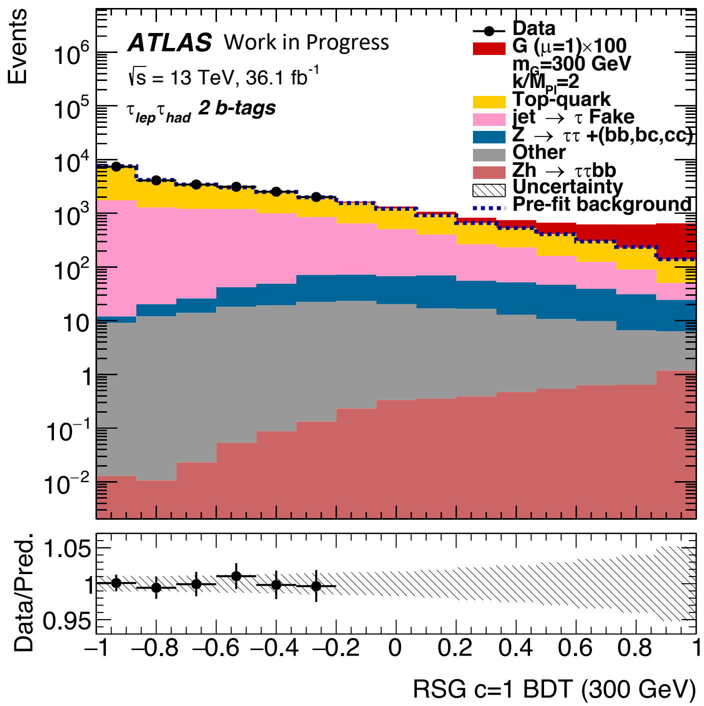


Input variables used in the BDT. The red line shows the RSG signal with $m_G = 300$ GeV (scaled by 100).

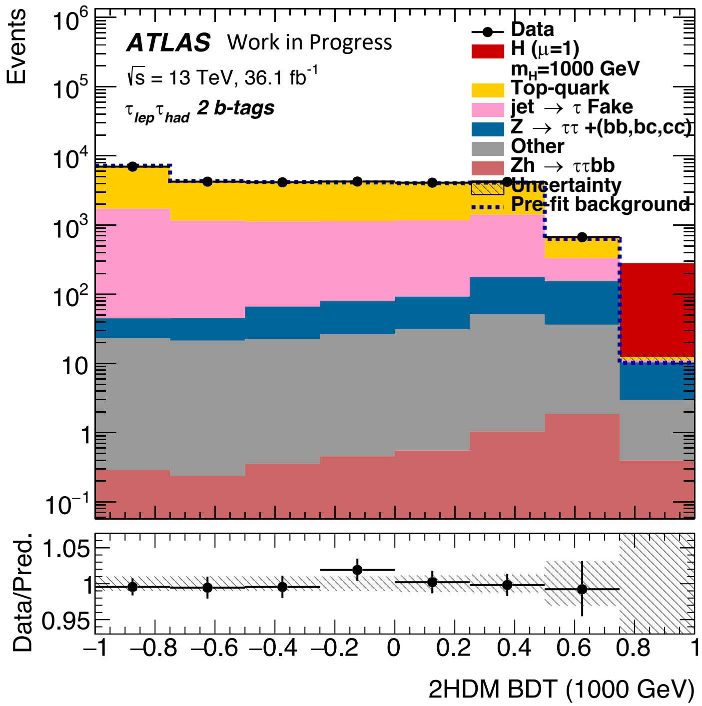
BOOSTED DECISION TREE

RESONANT BDT SCORES

- BDT scores used as final discriminant for limit setting
- Data here is blinded
- RSG signal scaled by 100



1. BDT Score for RSG c=1, $m_G = 300$ GeV

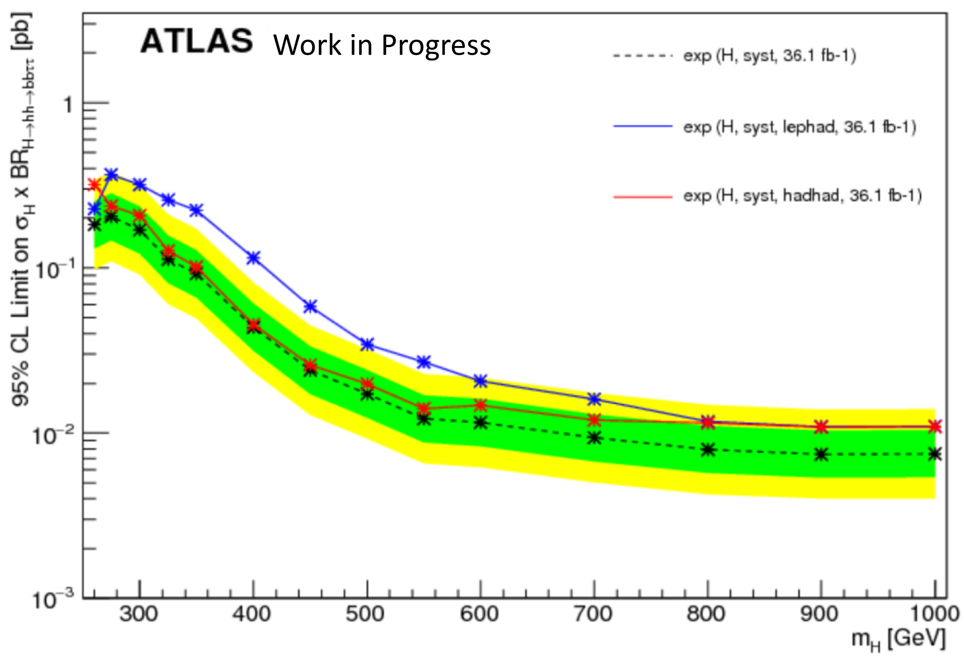


2. BDT Score for 2HDM, $m_H = 1$ TeV

EXPECTED SENSITIVITY

RESONANT EXPECTED LIMITS

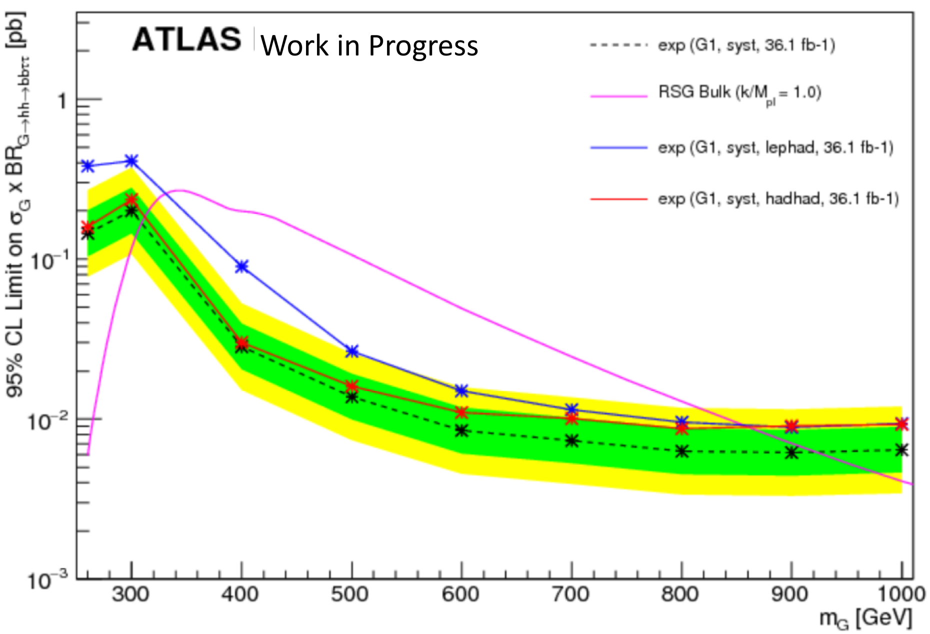
- Limits on $\sigma_H \times \mathrm{BR}_{H\to hh\to bb\tau\tau}$ where *H* denotes a heavy scalar in the 2HDM
- Plots show the **combined** limits for the two τ decay channels, $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$
- As well as the preselection + 2 *b*-tag signal region, a control region of $Z \to \mu\mu$ data events is included as a single bin in the fit in order to derive a normalisation factor for $Z + \mathrm{HF}$ background processes

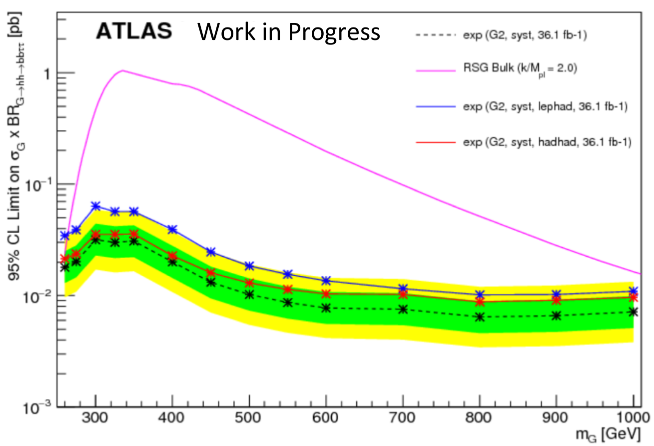


Expected Sensitivity

Resonant Expected Limits

- Limits on $\sigma_G \times \mathrm{BR}_{G\to hh\to bb\tau\tau}$ where G denotes a RS graviton with $\mathrm{c} = \frac{k}{\bar{M}_{Pl}} = 1.0$ (left) and 2.0 (right)
- For the RSG, this is shown alongside the theory predictions for $\sigma\times\mathrm{BR}$
- Plots show the **combined** limits for the two τ decay channels, $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$

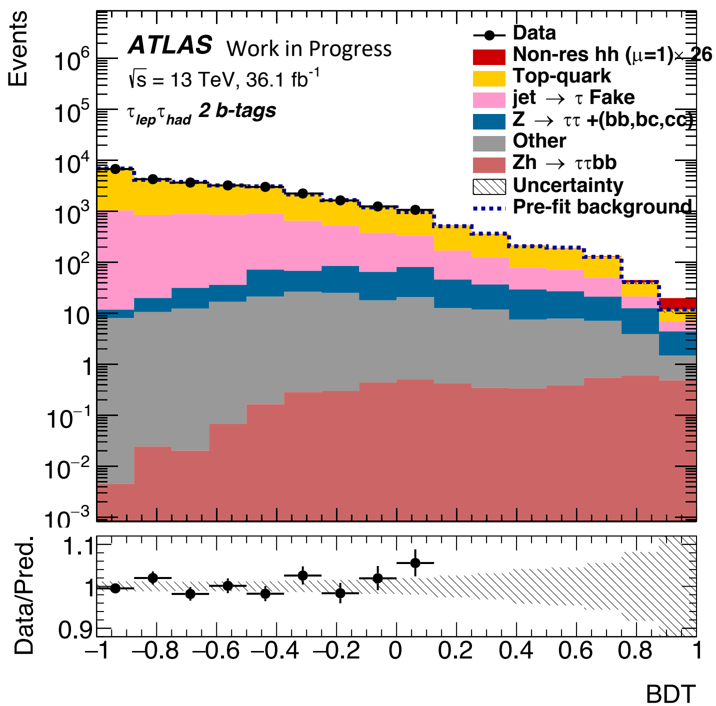




Expected Sensitivity

Non-Resonant Expected Limits

- Non-resonant BDT score is shown blinded for $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ only
- Expected limit for non-resonant di-Higgs production decaying to a $bb\tau\tau$ final state is $13.3\times\sigma_{\mathrm{SM}}$
- This is shown in the table alongside other ATLAS and CMS results



Channel	Obs. (exp.) 95% CL limit on $\sigma/\sigma_{\mathrm{SM}}$ — ATLAS	CMS
bbbb	29 (38) ATLAS-CONF-2016-049	342 (308) arXiv:1606.04782
bbWW	-	410 (227) CMS-PAS-HIG-16-024
bbττ	**(13.3)**	28 (25) CMS-PAS-HIG-17-002
bbγγ	117 (161) ATLAS-CONF-2016-004	19.2 (16.5) CMS-PAS-HIG-17-008
WWγγ	747 (286) ATLAS-CONF-2016-071	-

2.3-3.2 fb^{-1} | 13.3 fb^{-1} | 35.9-36.1 fb^{-1}

SUMMARY

- Observation of non-resonant di-Higgs production will allow measurement of the Higgs self-coupling – a test of the Standard Model
- Resonant di-Higgs production may occur as a result of heavy BSM particles
- Boosted decision tree has increased sensitivity by up to 50% compared to a cut-based analysis
- $bb\tau\tau$ expected sensitivity for non-resonant is competitive with other decay channels

Back-Up

Trigger and Event Selection

Lep-Had: Single Lepton Trigger

- **Single electron triggers**
  - Several different single electron triggers are used to maximise acceptance
  - ≥ 1 electron with $p_T > 24$ GeV satisfying the 'medium' identification criteria and 'loose' isolation requirements (in later data-taking periods the *p*T threshold is raised to 26 GeV and identification changed to 'tight')
  - **or** ≥ 1 electron with $p_T > 60$ GeV that applies the identification criteria but no requirement on the isolation,
  - **or** ≥ 1 electron with $p_T > 120 - 140$ satisfying the 'loose' identification criteria
- **Single muon triggers**
  - 1 muon with $p_T > 24 - 26$ GeV (depending on data-taking period) satisfying a 'loose' isolation criteria
  - **or** ≥ 1 muon with with $p_T > 50$ GeV without other requirements
- Any event that fails the single lepton triggers is tested to see if it passes a lepton-plus-tau trigger (LTT).

Trigger and Event Selection

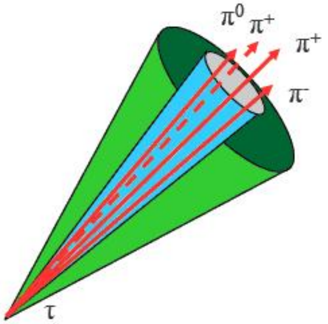
Lep-Had Selection

- $l\tau_{\text{had}}bb + E_T^{\text{miss}}$ final state → BDT
- **Single lepton trigger**
  - Exactly one electron passing the ‘tight’ criteria or exactly one muon passing ‘medium’ criteria (that includes the requirement that the muon must have $\eta < 2.5$), with p_T 1 GeV above the corresponding trigger threshold for the data-taking period.
  - Exactly one hadronic tau with $p_T > 20$ GeV and $\eta < 2.3$
  - At least 2 jets with $p_T > 45, 20$ GeV
- No other electrons and muons in the event
- Opposite sign between the tau and the light lepton

Event Reconstruction

Hadronic Taus

- Jets formed using the anti-k_T algorithm with a radius parameter $R = 0.4$
- Input to the visible hadronic tau decay reconstruction algorithm – provides little rejection of the jet background
- Input the hadronic tau candidates to a BDT
- Trained separately for 1- and 3-track hadronic tau candidates
- Require to pass the “medium” working point, which is a p_T-dependent definition starting at 0.6 of the BDT



Anti-Taus

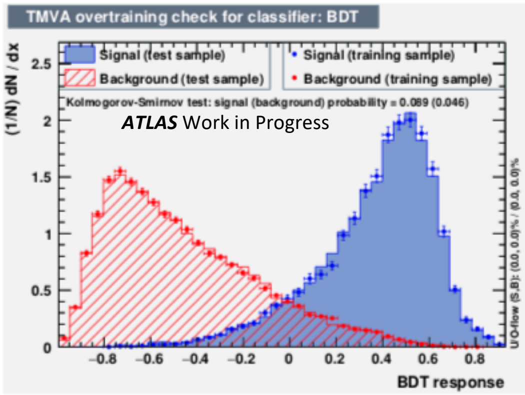
- Anti-tau ID criteria is used in calculating fake factors
- Anti-taus are tau candidates that pass a cut on the tau-ID BDT of 0.35 but fail the ”medium” requirement

EVENT RECONSTRUCTION

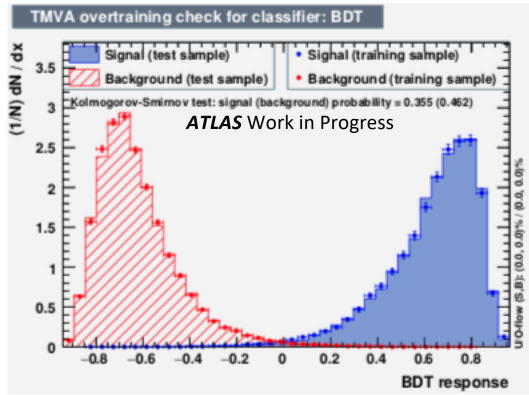
B-TAGGING

- *b*-jets are identified using the MV2 multivariate discriminant
- This combines variables from the basic b-tagging algorithms
- Use the 70% working point (this has an average tagging efficiency of 70% for *b*-jets in ttbar events)
- For small backgrounds not typically produced in association with b-jets, use ‘truth-tagging’ to keep full stats: give each event a random MV2 value and weight by the efficiency of each jet to pass the b-jet selection.

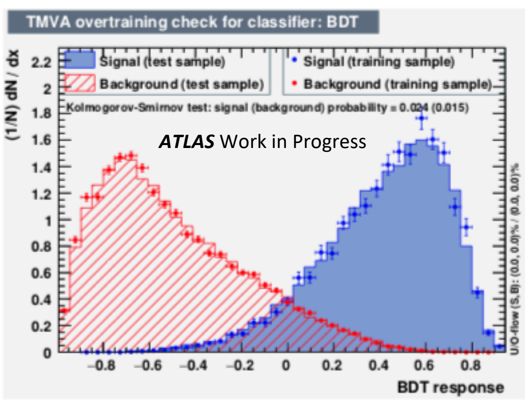
BDT Overtraining Plots



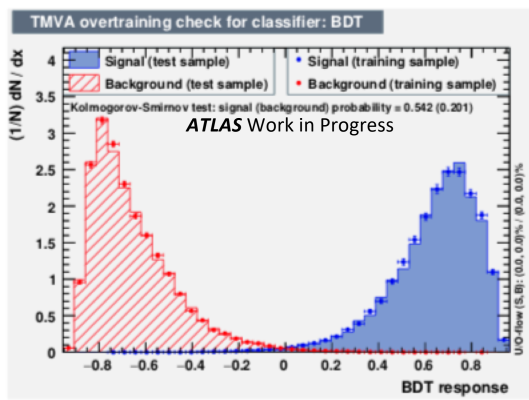
(a) 2HDM (m_H = 300 GeV)



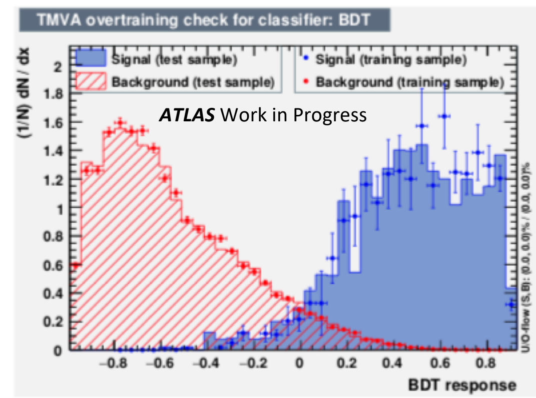
(b) 2HDM (m_H = 700 GeV)



(c) RSG (m_G= 300 GeV)

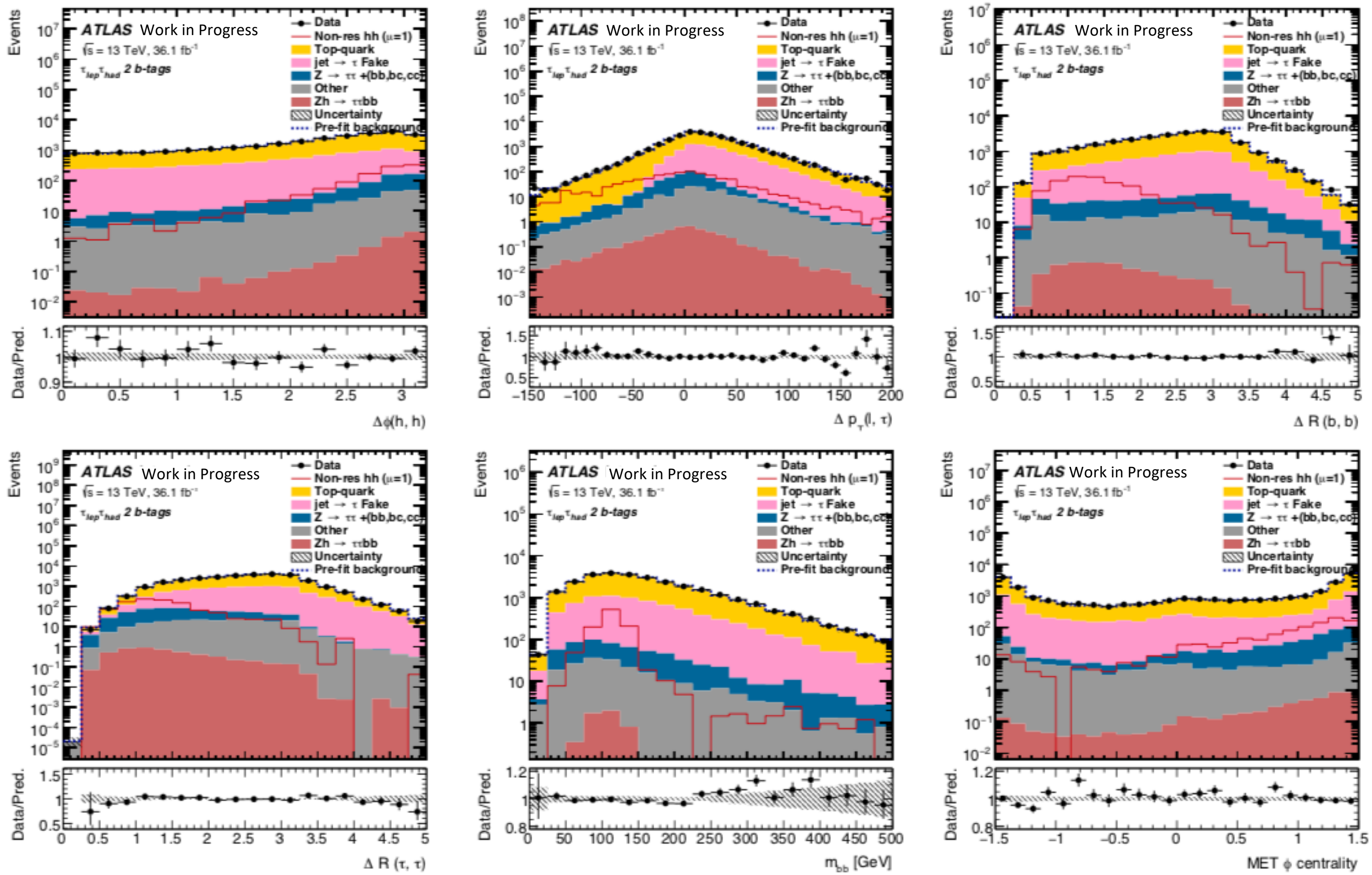


(d) RSG (m_G= 700 GeV)

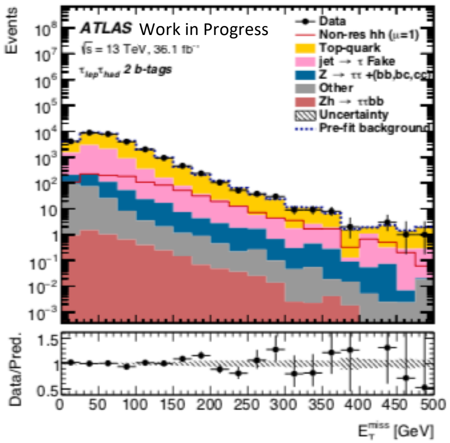


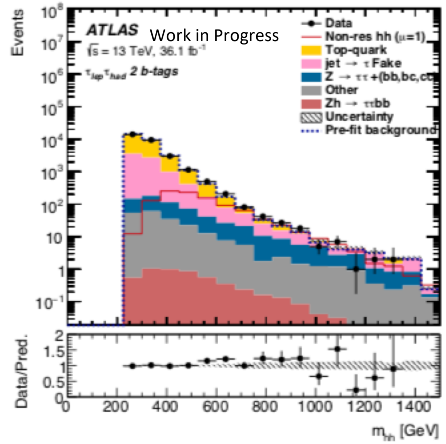
(e) Non-resonant

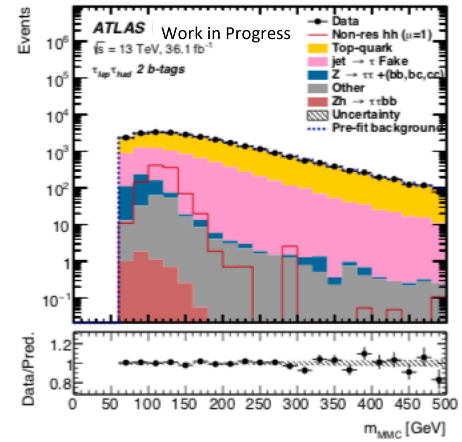
BDT Input Variables

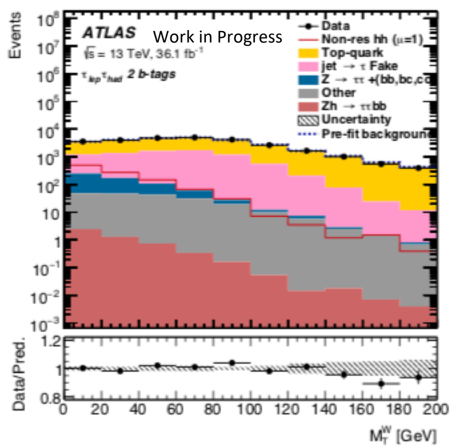


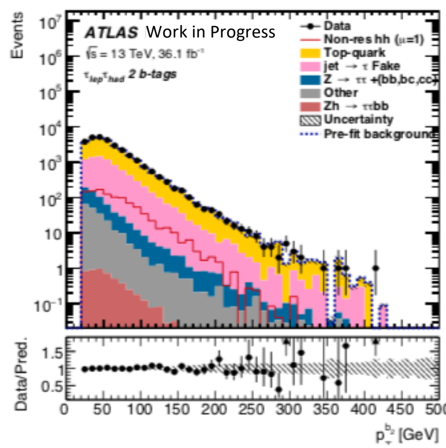
BDT Input Variables











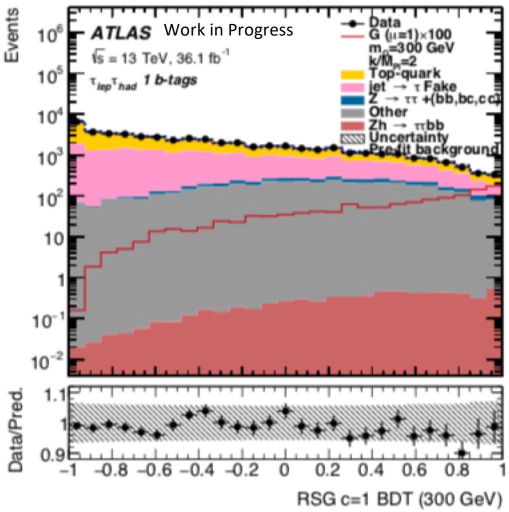
BDT Configuration

- NTrees = 200
- MaxDepth = 4
- MinNodeSize = 5%
- nCuts = 100
- AdaBoostBeta = 0.15
- NegWeightTreatment="InverseBoostNegWeights" (default)

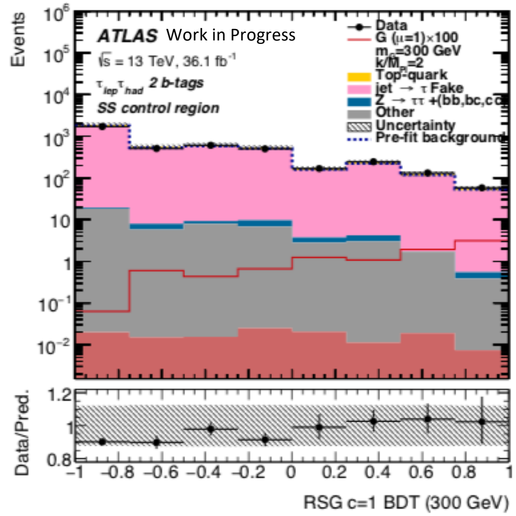
BDT Binning

- Rebin BDT score distributions to balance between the high-stat, low BDT score region and keeping shape information for signal in high BDT score region
- Keep uncertainty on the background less than a certain value (x) multiplied by the fraction of the signal
- Chose value of x by eye by looking at the expected number of events at high-BDT-score and trying to keep it above 10
- Require 10 events in last bin

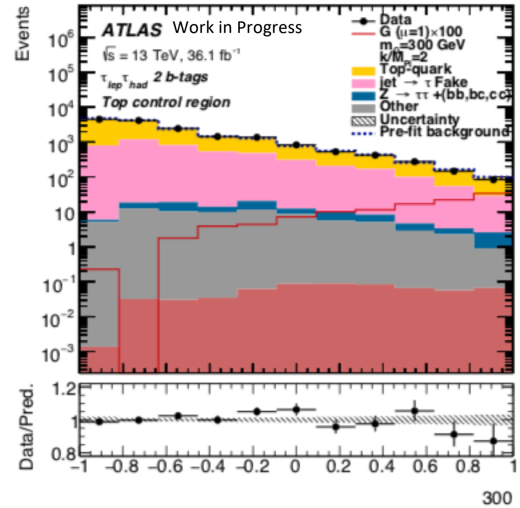
BDT Distributions in CRs



1-tag OS CR

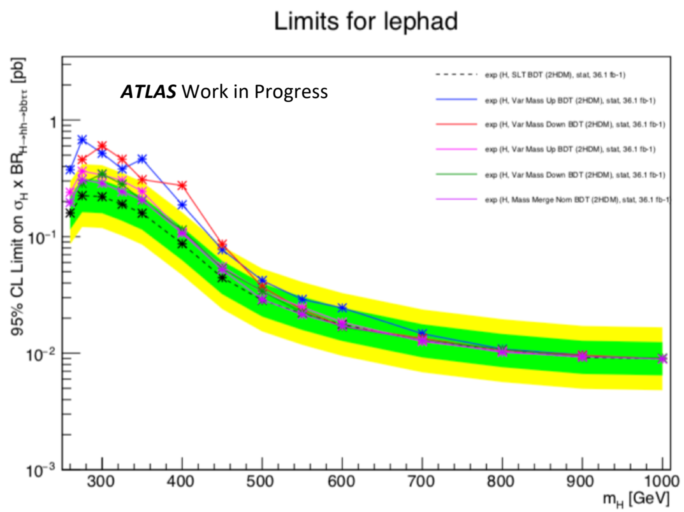


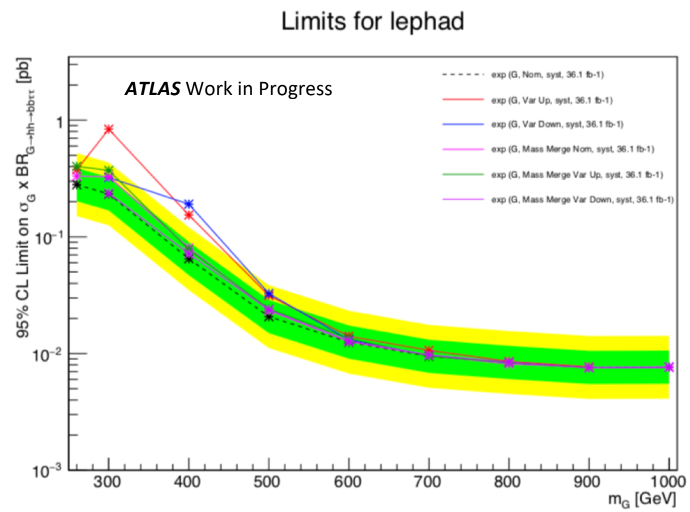
2-tag SS CR

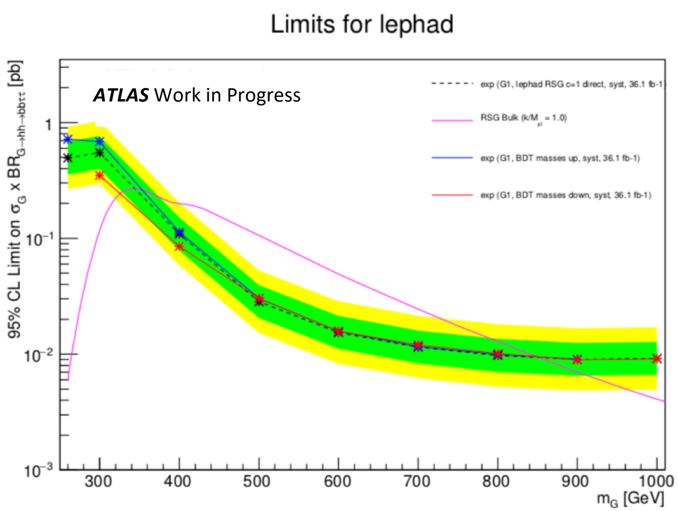


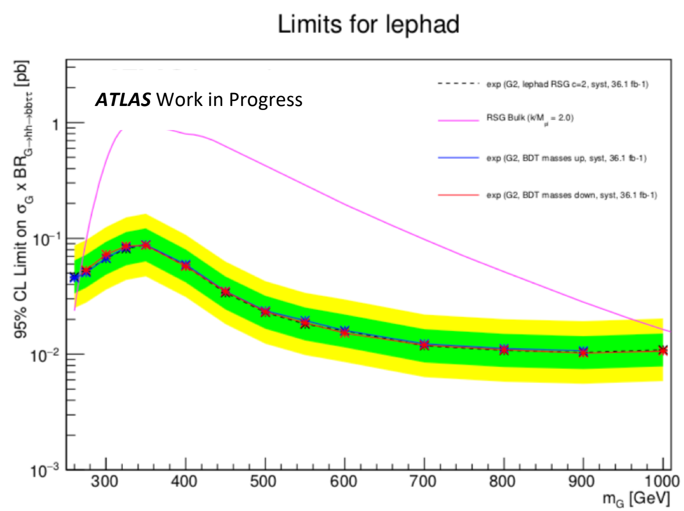
Top CR

BDT Sensitivity

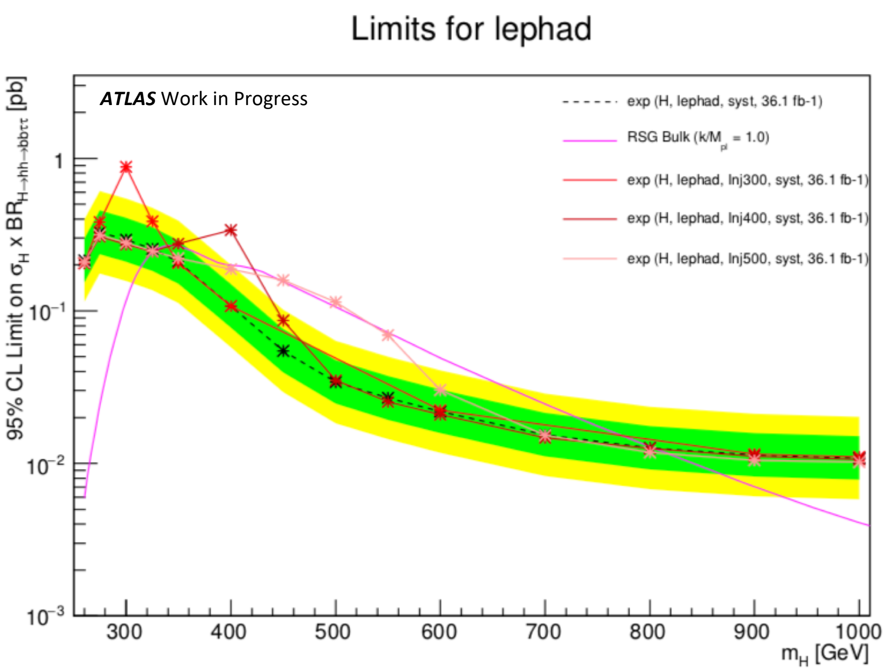


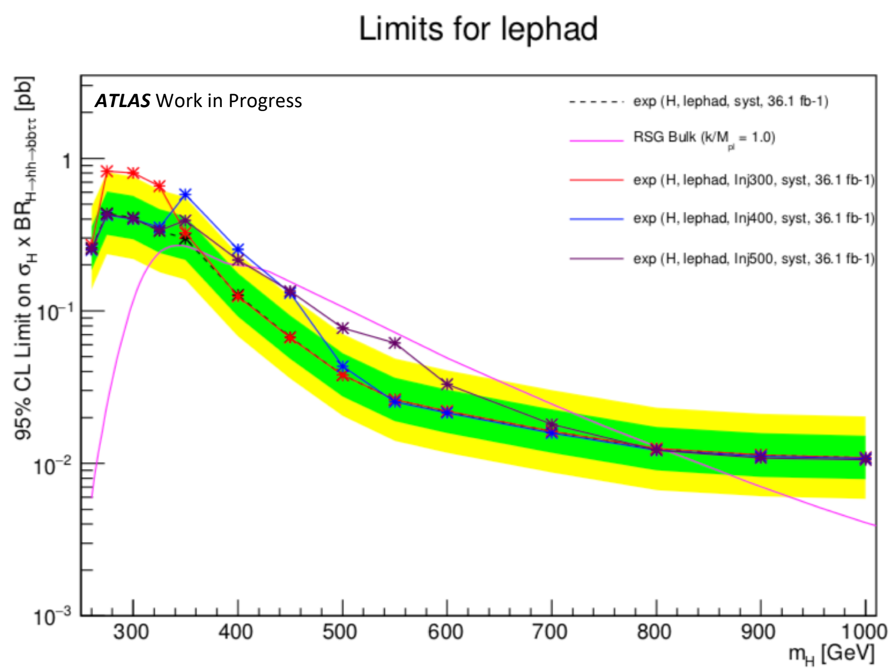






BDT SENSITIVITY





Signal injections: 1 mass point training (left) and 3 (right)

Improvements for Full Run 2

- VBF production mode analysis underway
- Could add leplep final state
- 2-stage BDT selection: add separate BDT training for Z+bb background
- Try different multivariate methods: want this to be mass-independent to improve on current method
- Varied-lambda analysis: Current analysis is not really sensitive to the triangle diagram (i.e. lambda_hhh). Have samples with various values of lambda between -20 and 20. Will be able to produce limits on lambda