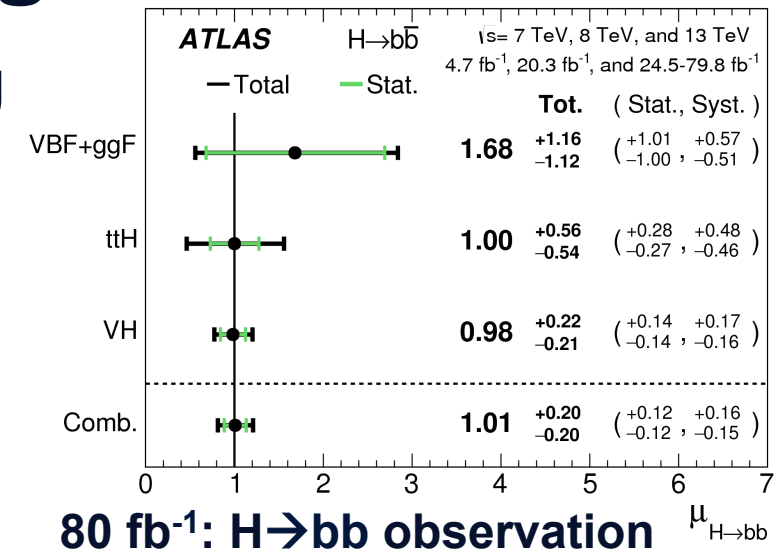
Status and Plans of the H→bb Analyses

Elisabeth Schopf on behalf of the analysis teams

ATLAS Overview Week, October 2019

+ Hbb Group's Analyses

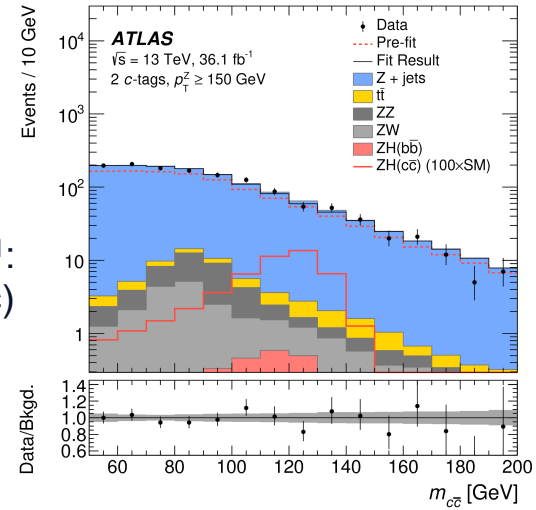
- Currently 6 analyses teams working towards publication in 2019/2020:
  - VH(bb) resolved
  - VH(bb) boosted — New
  - VH(cc) (resolved)
  - VBF H(bb)+γ
  - VBF H(bb)+hadronic
  - Hadronic (ggF) H(bb)+ISR — Previously hosted in Exotics

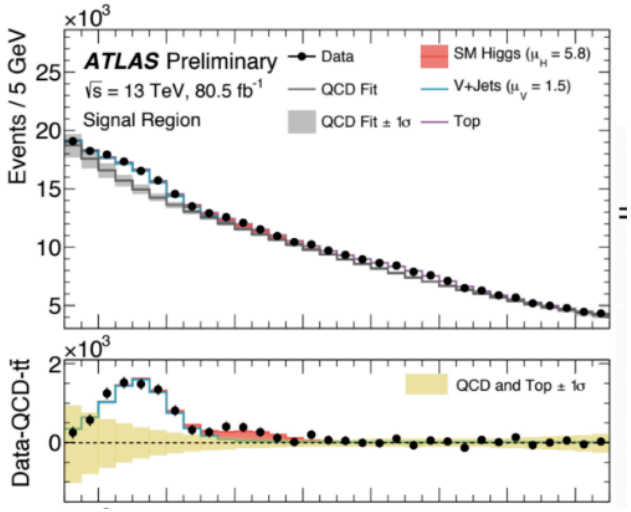


80 fb^{-1}: H→bb observation

https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2018-04/

36 fb^{-1}: VH(cc)





80 fb^{-1}: H(bb)+ISR

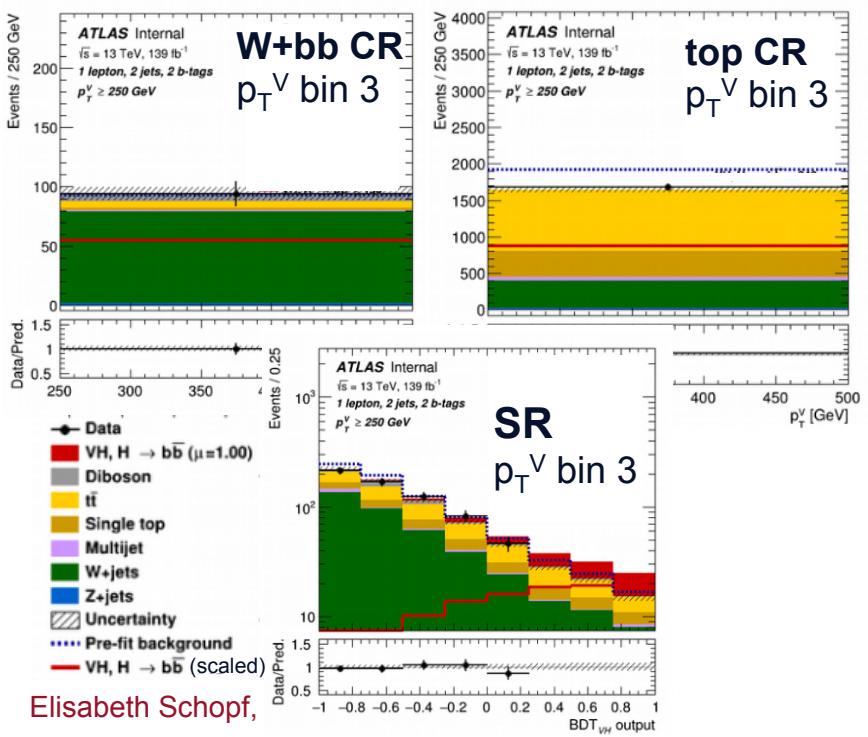
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2017-01 https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-052

VH(bb/cc): Analysis Activities

- → Focus on resolved and boosted VH(bb) analyses
- → VH(cc) working towards publication in early 2020 and currently consolidating fit model and modelling uncertainties
- → Long term: “legacy paper“ in 2021 combining VH(bb) boosted and resolved and VH(cc) in one harmonised analysis

VH(bb) Resolved: Strategy

- 2 electrons/muons or 1 muon/electron+E_T^{miss} or E_T^{miss} from leptonic V decay + 2 b-tagged R=0.4 jets from Higgs decay
  - Split phase space in N(leptons), N(jets) and p_T^V categories
  - Signal and control regions defined in $\Delta R(b,b)$-p_T^V plane **New**
  - Multivariate discriminant (BDT) in signal regions
    - Di-jet mass analysis as cross-check

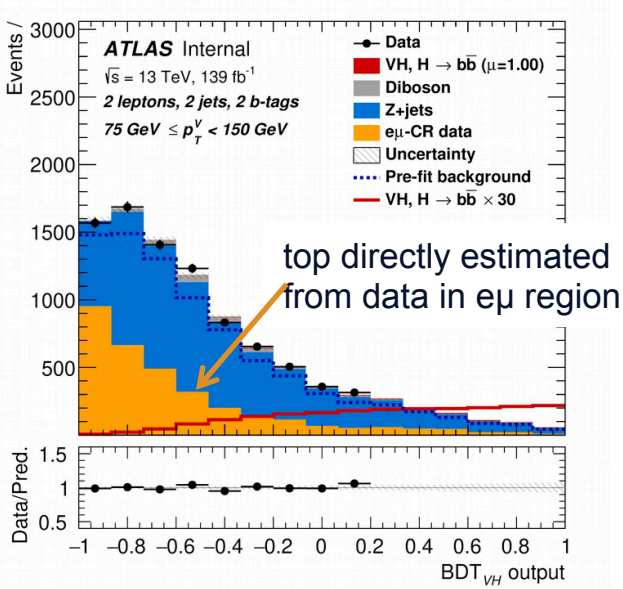


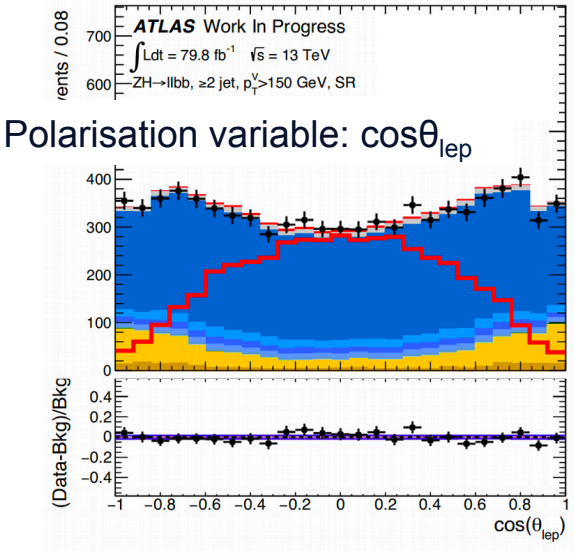
- **Analysis scope with 140 fb^{-1}**:
  - Inclusive analysis (μ)
  - → Improve background control and reduce systematics
  - Differential measurement (STXS) in 3 p_T^V analysis bins: [75,150], [150,250], [250,∞]

Target: ATLAS circulation in 2019

VH(bb) Resolved: Improvements

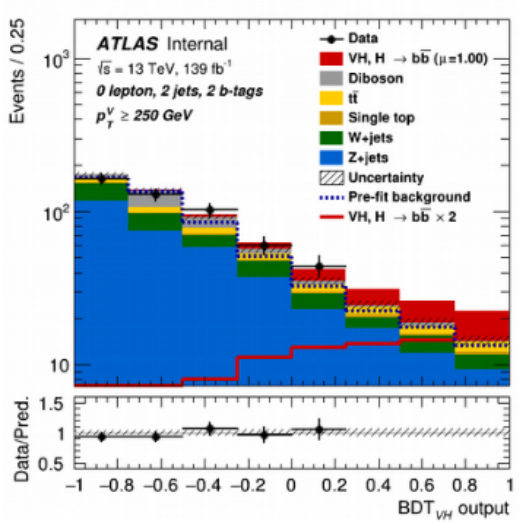
- **Background control and modelling:**
  - New control regions in 3 p_T^V bins
  - Data driven estimate for di-leptonic top events
  - Updated modelling uncertainties including latest samples and recommendations
  - Under scrutiny: multivariate reweighting for modelling uncertainties
- **Sensitivity improvements:**
  - Z-polarisation info in BDT of di-lepton ZH channel
  - Truth tagging for all c- and light-jets in MC
  - Under scrutiny: pseudo-continuous b-tagging
- **STXS bins:**
  - Analysis p_T^V regions in line with STXS bins
  - Under scrutiny: 75 GeV<p_T^V<150 GeV WH region



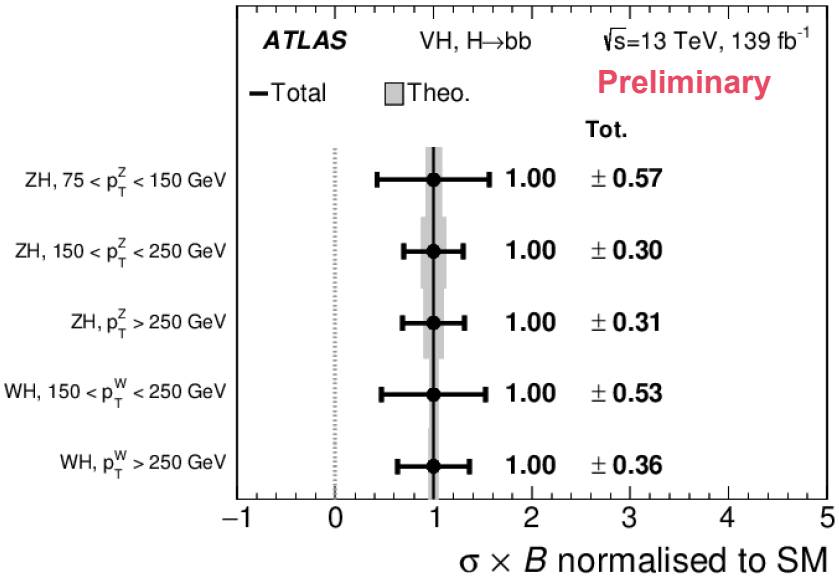


VH(bb) Resolved: Results

Expected significances (full fit)	WH	ZH
80 fb^{-1} (observation paper)	2.3 σ	3.5 σ
Observation paper scaled to 140 fb^{-1}	3.2 σ	4.6 σ
140 fb^{-1} (current baseline)	3.6 σ	5.4 σ



- Result reflects improvements:
  - Significances improve beyond expectation from addition of 2018 data
  - Large reduction in impact from ttbar, W+jets and Z+jets uncertainties
- MC stat. impact reduced due to truth tagging and MC extensions (additional extensions to be added soon)
- Largest impact from: VH modelling, b-tagging and jet systematics

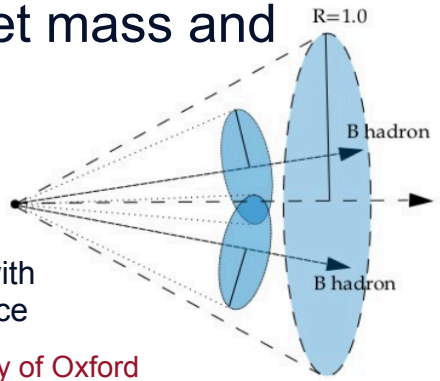


VH(bb) Boosted*: Strategy

- Leptonic V decay (same as resolved) + 1 R=1.0 jet with 2 b-tagged VR track jets from Higgs decay
  - Split phase space in N(leptons), N(jets) outside large-R jet and p_T^V
  - Control regions defined by additional b-tagged jet outside large-R jet
  - Final discriminant: large-R jet mass
- **First boosted VH(bb) analysis**

→ Large effort to understand the boosted objects (with CP groups)

- b-tagging
- VR track jet overlap removal
- Large-R jet mass and resolution



*currently there is overlap with resolved VH(bb) phase space

- **Analysis scope with 140 fb^{-1}**:
  - Inclusive analysis (μ measurement)
  - Investigate STXS fit in the 2 p_T^V analysis bins: [250,400], [400,∞]

Target: ATLAS circulation in 2019

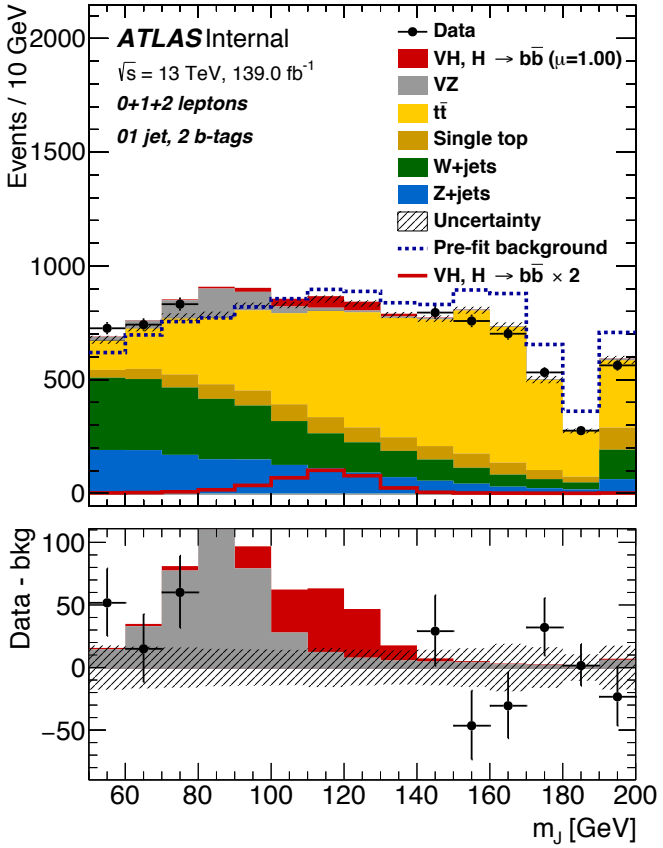
+ VH(bb) boosted: Optimisation and Status

- Optimisation studies to improve sensitivity (being) finalised
  - Investigated **selection requirements**: triggers, Z-mass window, Z-polarisation*, ΔY(V,H), substructure variables (C2)
  - Implementation of **truth tagging**
  - Optimised SR and CR **categorisation**

*$(p_T^{lep1}-p_T^{lep2})/p_T^Z$

→ **Focus now**: modelling and fit studies

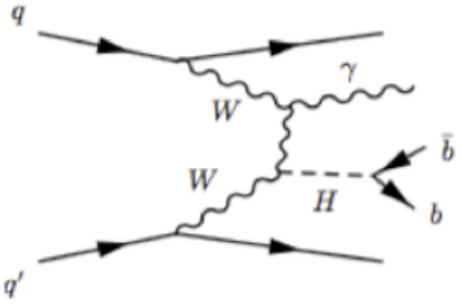
- **Expected (preliminary) results:**
  - Signal significance: 3.2 σ
  - Uncertainties with largest impact: data statistics, large-R jet, MC modelling, MC stat. (MC extensions on their way)



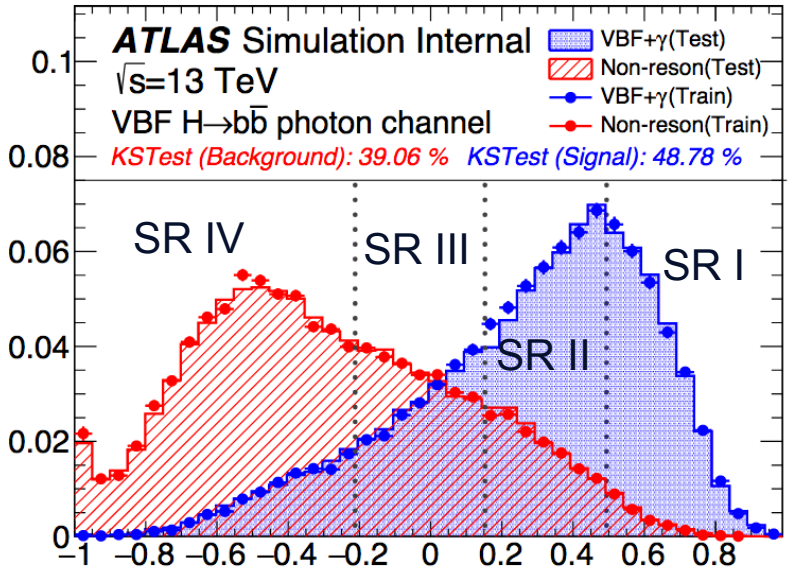
VBF H(bb): Analysis Activities

- → Focus on VBF H(bb)+γ
- → Hadronic VBF H(bb) working toward publication in summer 2020
  - → Currently focussing on optimisation and validation
  - → Optimisation of multivariate algorithm for categorisation (investigating adversarial neural nets)

VBF H(bb)+γ: Strategy



- 1 photon, 2 R=0.4 jets from VBF signature + 2 b-tagged R=0.4 jets from Higgs decay
  - Multivariate algorithm (BDT) to define 4 signal regions
  - Final discriminant: invariant mass of di-b-jet system
  - Multi-jet background extrapolated from fit to m_{bb} sidebands



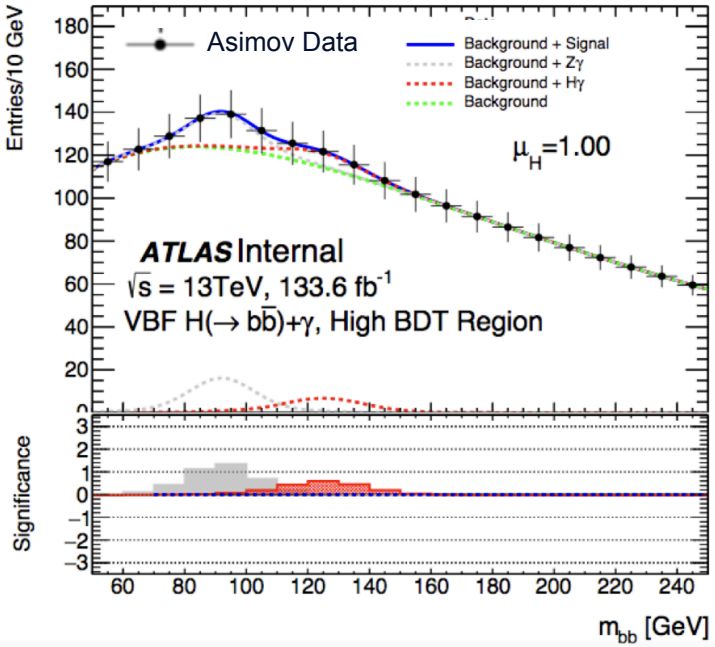
- **Analysis scope with 140 fb^{-1}:** Inclusive analysis (μ measurement)

Target: publication in 2019

+ VBF H(bb)+γ: Activities and Status

- BDT training and spurious signal estimate relies on MC
  - New high stat. NLO signal samples: changes observed in crucial BDT input variables with known modelling problems
  - → 25% improvement in signal-background separation in BDT
  - Large effort to get robust spurious signal estimate
  - → Difficult due to limited MC statistics but converging now

- **Expected (preliminary) results:**
  - Stat. only signal significance: **1.2 σ**
    - Inclusion of spurious signal: <1 σ
  - Analysis limited by data statistics and spurious signal



Hadronic (ggF) H(bb)+ISR: Analysis Activities

→ New analysis in Hbb group

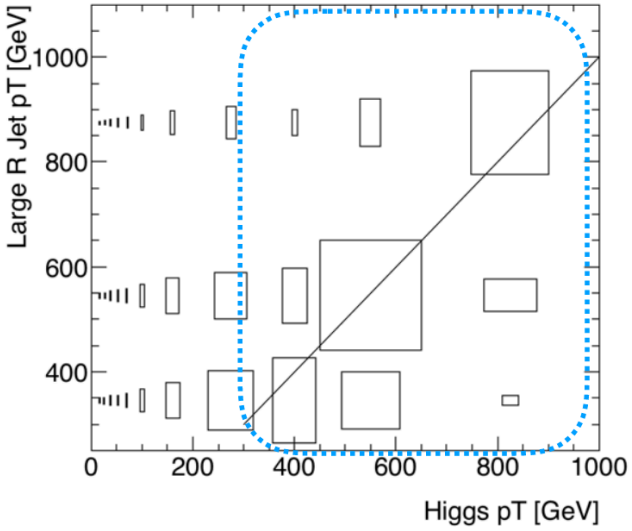
→ Previous publications within Exotics group

Hadronic H(bb)+ISR: Strategy

- 1 R=1.0 ISR recoil jet + 1 R=1.0 jet with 2 b-tagged VR track jets from Higgs decay:
  - Split phase space separating events where leading R=1.0 jet has b-tags and sub-leading R=1.0 jet has b-tags
  - Final discriminant: large-R jet mass of Higgs candidate jet
  - Validation region for data driven QCD estimate: 2 large-R jets with 0 b-tags

- **Analysis scope with 140 fb^{-1}**:
  - Inclusive analysis (μ+fiducial x-section)
  - Differential analysis (STXS): bins in p_T^H [300,450],[450,650],[650,1000]

Target: Moriond 2020



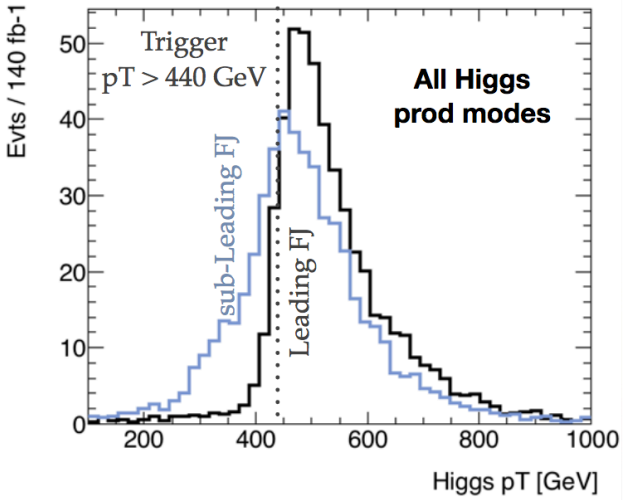
Hadronic H(bb)+ISR: Activities and Status

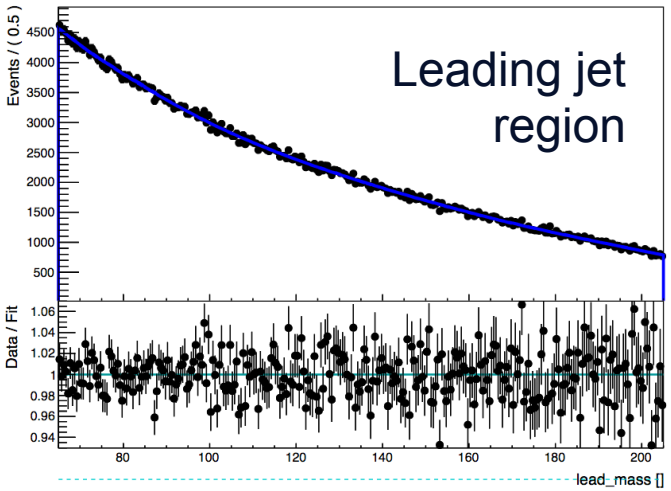
- Optimised analysis selection to be able to probe lower p_T
  - Updated **trigger** strategy
  - Including events where **sub-leading large-R jet** has 2 b-tags

  → Gain in sensitivity: **~20%**

→ Finalisation of analysis strategy

- Main focus now: study QCD background in analysis regions and p_T^H bins and consolidation of fit model





Analysis would benefit significantly from new b-taggers and improved Xbb tagging but probably beyond current analysis time scale

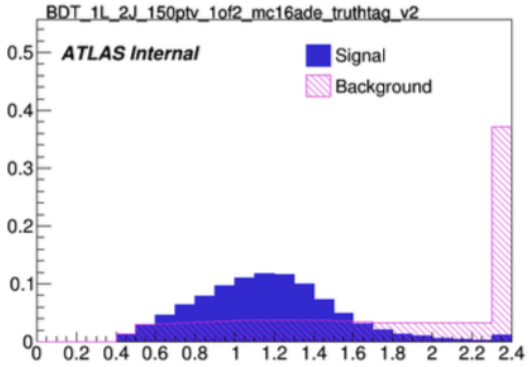
Summary

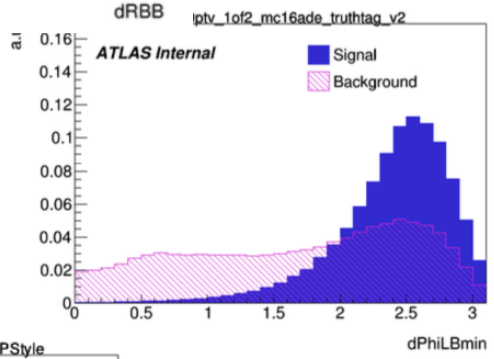
- VH(bb) analyses working towards publication before Moriond 2020
  - Resolved analysis included several analysis improvements → WH: evidence (exp.), ZH: observation (exp.)
  - First analysis in boosted phase space: 3.2 σ (exp.)
- VBF H(bb)+γ working towards publication before Moriond 2020
  - Large effort going into signal MC validation and evaluation of systematics
  - Stat. only significance (exp.): 1.2 σ
- Hadronic H(bb)+ISR working towards publication on Moriond 2020 time scale
  - Improved analysis strategy and investigation of binning for STXS analysis
- Other analyses working towards publication in 2020:
  - (resolved) VH(cc): Moriond 2020
  - Hadronic VBF H(bb): summer 2020

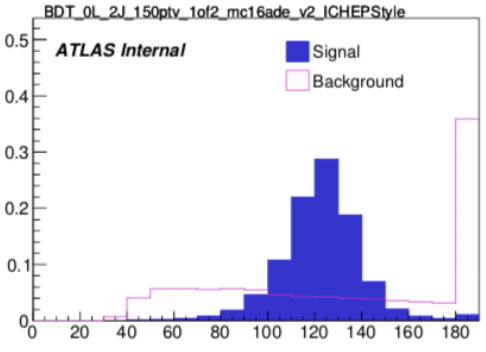
Backup

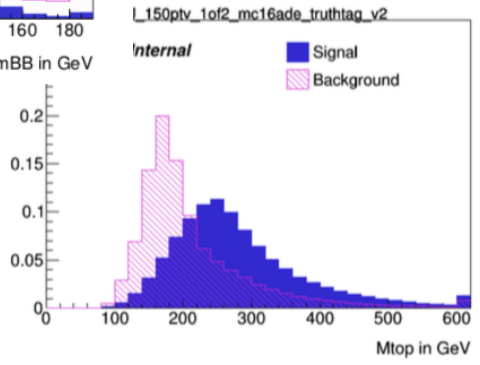
+ Resolved VHbb: BDT Variables

Variable	Name	0-lepton	1-lepton	2-lepton
m_{jj}	mBB	✓	✓	✓
$\Delta R(jet_1, jet_2)$	dRBB	✓	✓	✓
p_T^{jet1}	pTB1	✓	✓	✓
p_T^{jet2}	pTB2	✓	✓	✓
p_T^V	pTV	✓	✓	✓
$\Delta\phi(V, H)$	dPhiVBB	✓	✓	✓
$\lvert\Delta\eta(jet_1, jet_2)\rvert$	dEtaBB	✓		
M_{eff}	MEff	✓		
$E_{\text{T}}^{\text{miss}}$	MET	$\equiv p_T^V$	✓	
$\min(\Delta\phi(\ell, jet))$	dPhiLBmin		✓	
m_{T}^W	mTW		✓	
$\Delta Y(W, H)$	dYWH		✓	
m_{top}	mTop		✓	
$E_{\text{T}}^{\text{miss}}$ significance	METSig			✓
$\Delta\eta(V, H)$	dEtaVBB			✓
$m_{\ell\ell}$	mLL			✓
$\cos\theta$	cosTheta			✓
	Only in 3 Jet Events			
$p_T^{\text{jet}_3}$	pTJ3	✓	✓	✓
m_{jjj}	mBBJ	✓	✓	✓



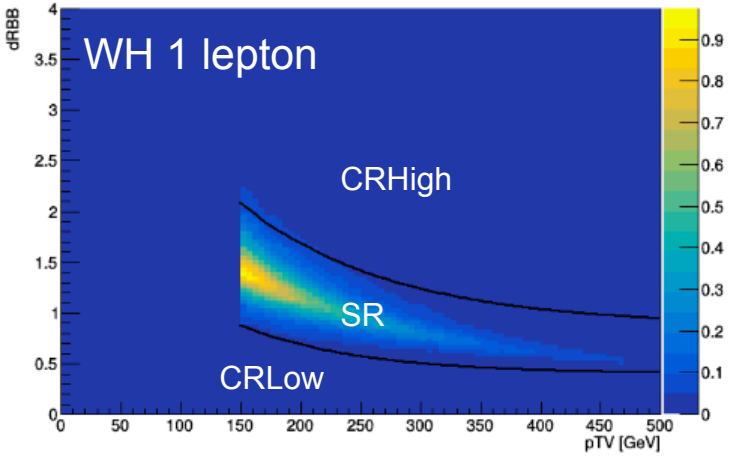


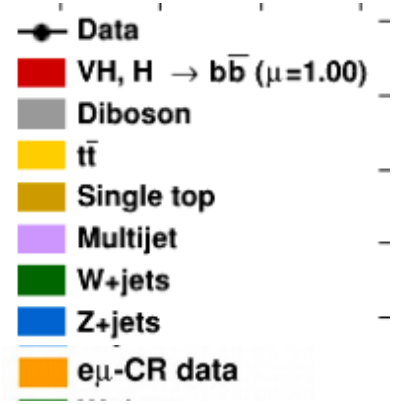


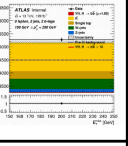
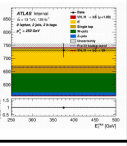
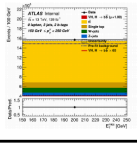
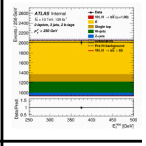
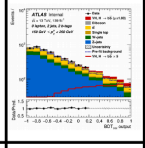
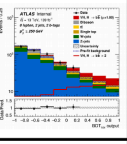
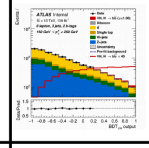
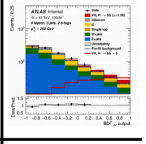
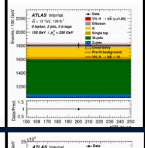
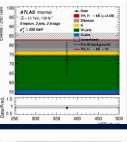
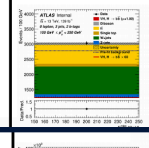
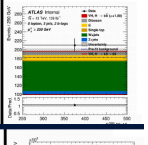
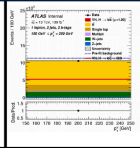
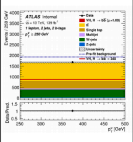
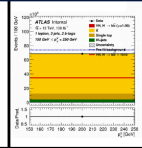
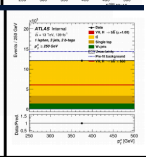
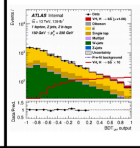
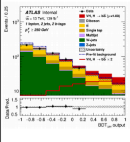
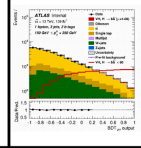
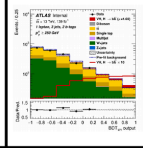
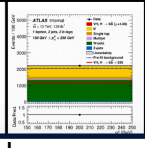
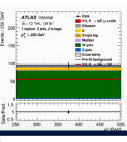
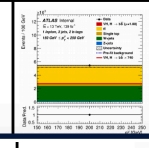
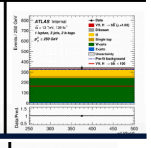
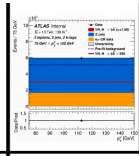
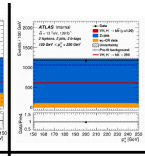
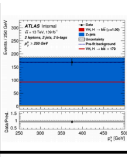
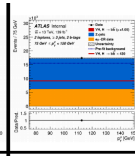
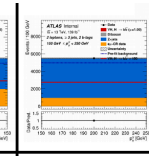
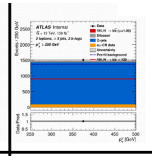
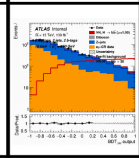
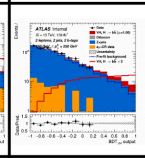
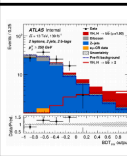
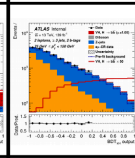
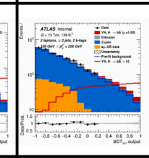
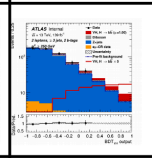
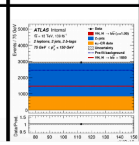
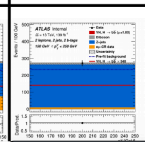


+ Resolved VHbb: SR-CR

Visualisation of CR-SR region selection





		N(jet) == 2			N(jet) >= 3		
		$p_T(V)$: 75-150 GeV	$p_T(V)$: 150-250 GeV	$p_T(V)$ > 250 GeV	$p_T(V)$: 75-150 GeV	$p_T(V)$: 150-250 GeV	$p_T(V)$ > 250 GeV
0 lepton	CR – high						
	SR						
	CR – low						
1 lepton	CR – high	Under scrutiny			Under scrutiny		
	SR	Under scrutiny			Under scrutiny		
	CR – low	Under scrutiny			Under scrutiny		
2 lepton	CR – high						
	SR						
	CR – low						

Resolved: VHbb NP Breakdown

80 fb-1

NP set	STXS measu signed
Total	+0.272 / -0.254
DataStat	+0.164 / -0.162
FullSyst	+0.217 / -0.196
Floating normalizations	+0.033 / -0.041
Multi Jet	+0.005 / -0.006
Modelling: single top	+0.029 / -0.026
Modelling: ttbar	+0.049 / -0.051
Modelling: W+jets	+0.068 / -0.066
Modelling: Z+jets	+0.053 / -0.055
Modelling: Diboson	+0.055 / -0.053
Modelling: VH	+0.110 / -0.074
Detector: lepton	+0.009 / -0.008
Detector: MET	+0.019 / -0.018
Detector: JET	+0.045 / -0.031
Detector: FTAG (b-jet)	+0.062 / -0.062
Detector: FTAG (c-jet)	+0.044 / -0.037
Detector: FTAG (l-jet)	+0.004 / -0.004
Detector: FTAG (extrap)	+0.009 / -0.008
Detector: PU	+0.010 / -0.007
Lumi	+0.026 / -0.017
MC stat	+0.072 / -0.074

-50%

-60%

-50%

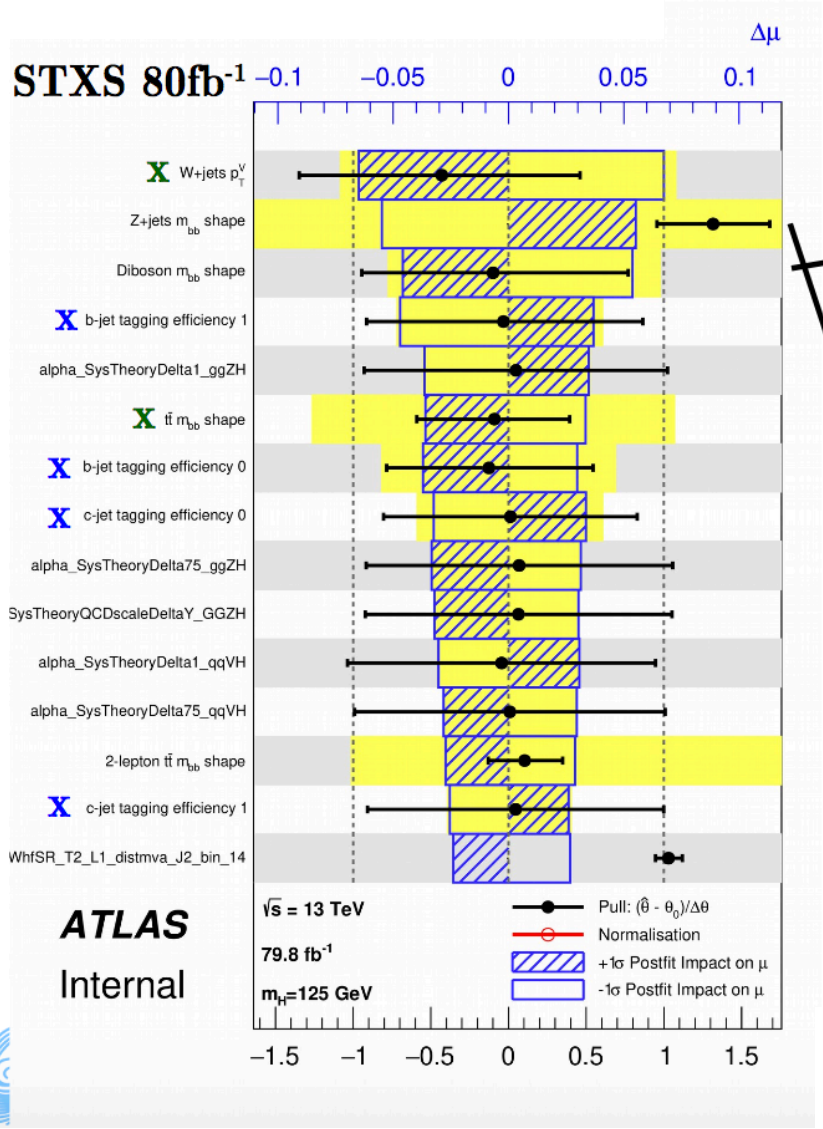
-30%

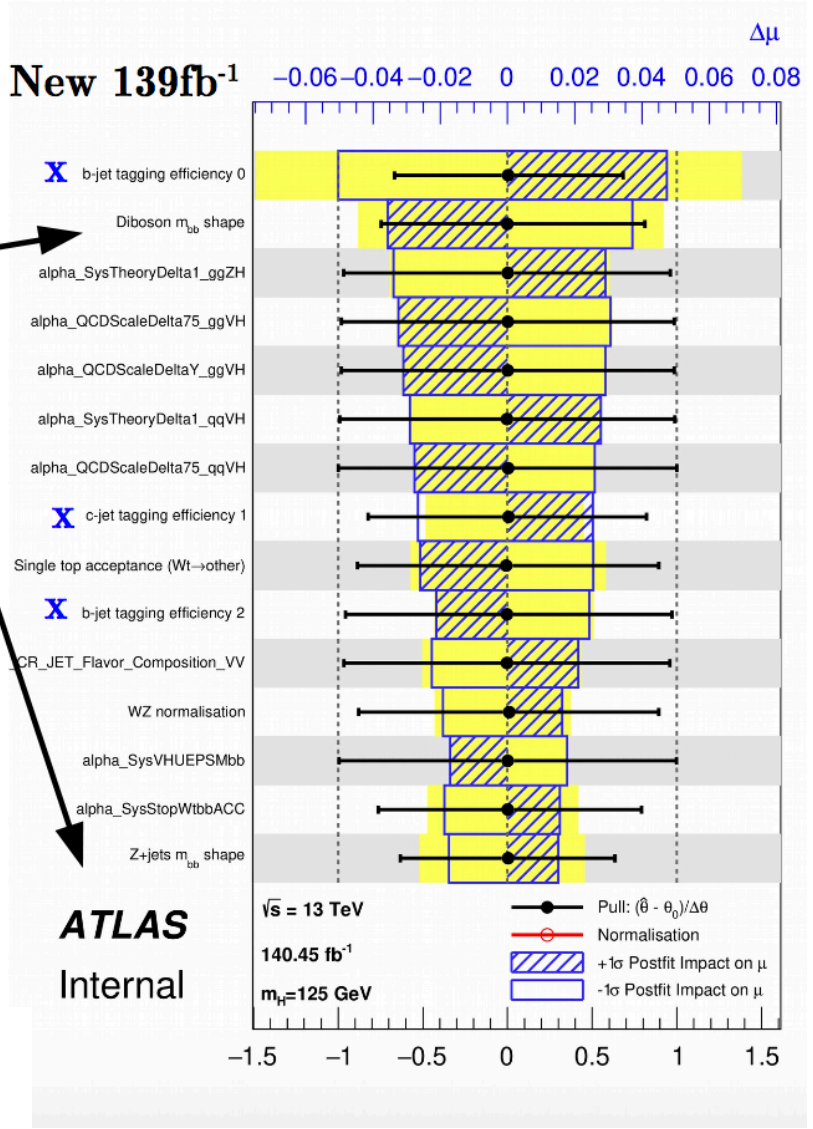
140 fb-1

Set of nuisance parameters	Impact on error Signed impact
Total	+0.188 / -0.172
DataStat	+0.121 / -0.119
Data stat only	+0.114 / -0.113
Floating normalizations	+0.019 / -0.020
FullSyst	+0.144 / -0.124
Modelling: VH	+0.064 / -0.048
Modelling: Background	+0.074 / -0.071
Multi Jet	+0.005 / -0.004
Modelling: single top	+0.032 / -0.030
Modelling: ttbar	+0.019 / -0.018
Modelling: W+jets	+0.018 / -0.018
Modelling: Z+jets	+0.020 / -0.021
Modelling: Diboson	+0.047 / -0.043
MC stat	+0.037 / -0.037
Experimental Syst	+0.077 / -0.069
Detector: lepton	+0.004 / -0.005
Detector: MET	+0.006 / -0.006
Detector: JET	+0.044 / -0.035
Detector: FTAG (b-jet)	+0.052 / -0.048
Detector: FTAG (c-jet)	+0.023 / -0.020
Detector: FTAG (l-jet)	+0.006 / -0.006
Detector: FTAG (extrap)	+0.002 / -0.002
Detector: PU	+0.005 / -0.003
Lumi	+0.017 / -0.012

The relevant information is the size of the NP group w.r.t. total (not the absolute size)

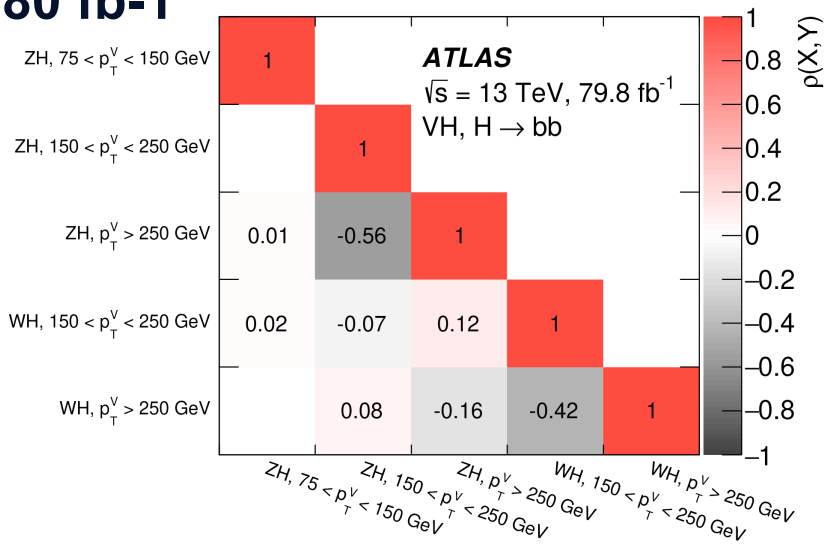
Resolved VHbb: Ranking



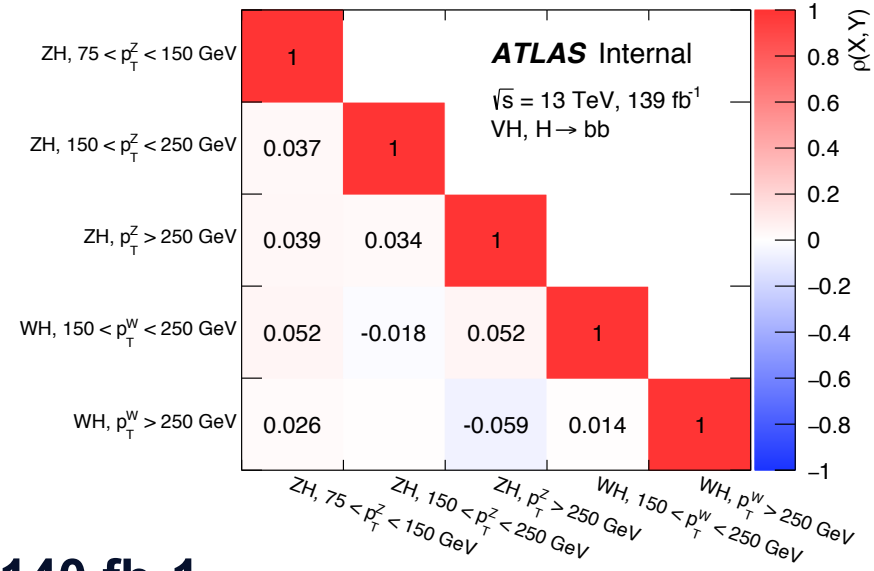


Resolved VHbb: STXS Fit Results

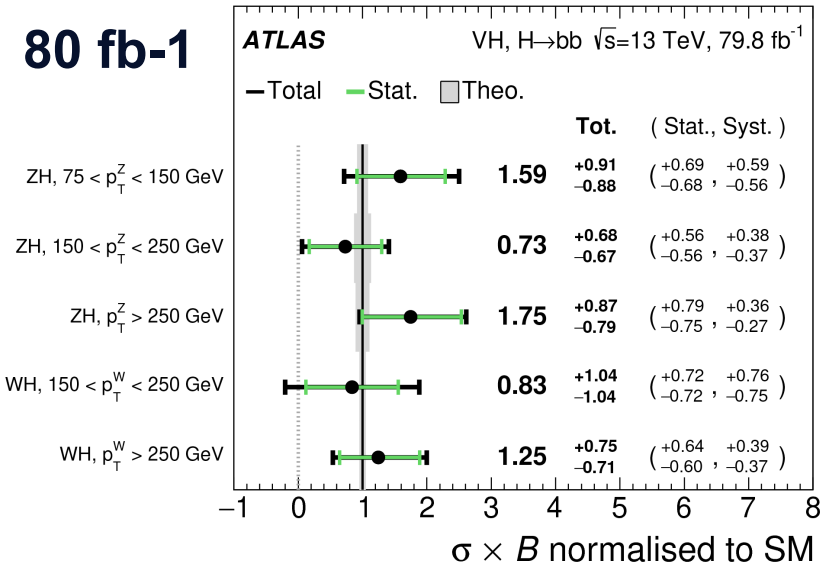
80 fb-1



140 fb-1



80 fb-1



140 fb-1

Region	Exp. Significance
WH, 150 GeV < p_T^W < 250 GeV	1.9 σ
WH, p_T^W > 250 GeV	3.0 σ
ZH, 75 GeV < p_T^Z < 150 GeV	1.8 σ
ZH, 150 GeV < p_T^Z < 250 GeV	3.6 σ
ZH, p_T^Z > 250 GeV	3.7 σ

Boosted VHbb: Event Selection

- single lepton (2L, 1e) or MET (0L, 1μ) triggers
  MET trigger also for 2μ? (+4% (10)% acceptance gain in total (μμ))

0L:

- $E_T^{miss} > 200$ GeV

Anti multijet cuts:

- $\Delta\varphi(E_T^{miss}, \text{fat jet}) > 120°$
- $\Delta\varphi(E_T^{miss}, p_T^{miss}) < 90°$
- $\min[\Delta\varphi(E_T^{miss}, \text{calo jets*})] > 30°$

*calo jets: $p_T > 70$ GeV, $|\eta| < 4.5$, not matched to the fat jet

1L:

- $E_T^{miss} > 30$ GeV
- $\Delta y(W, H) < 1.4$

recently adopted, not used in the fitting yet

2L:

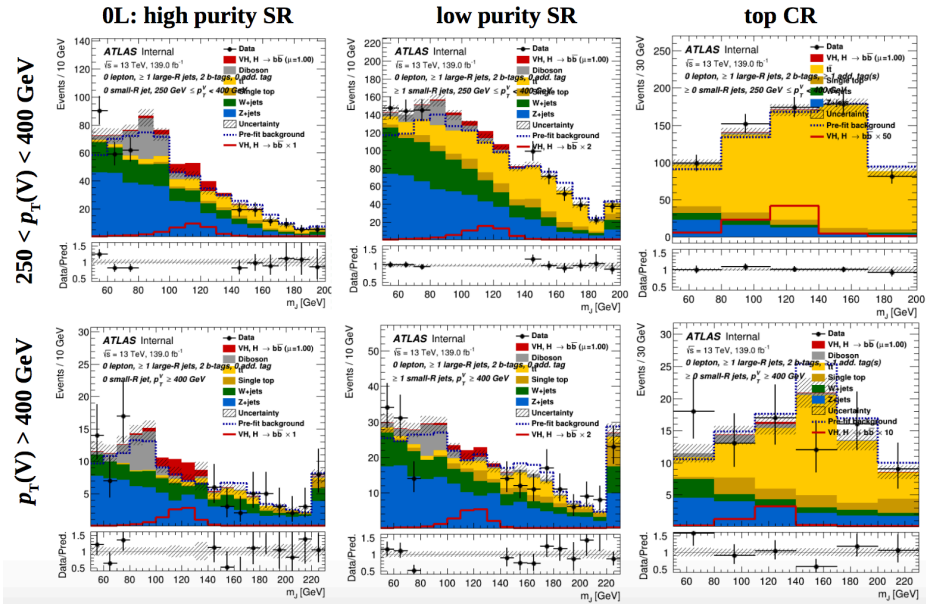
- Loose Z-mass window (66-116 GeV)

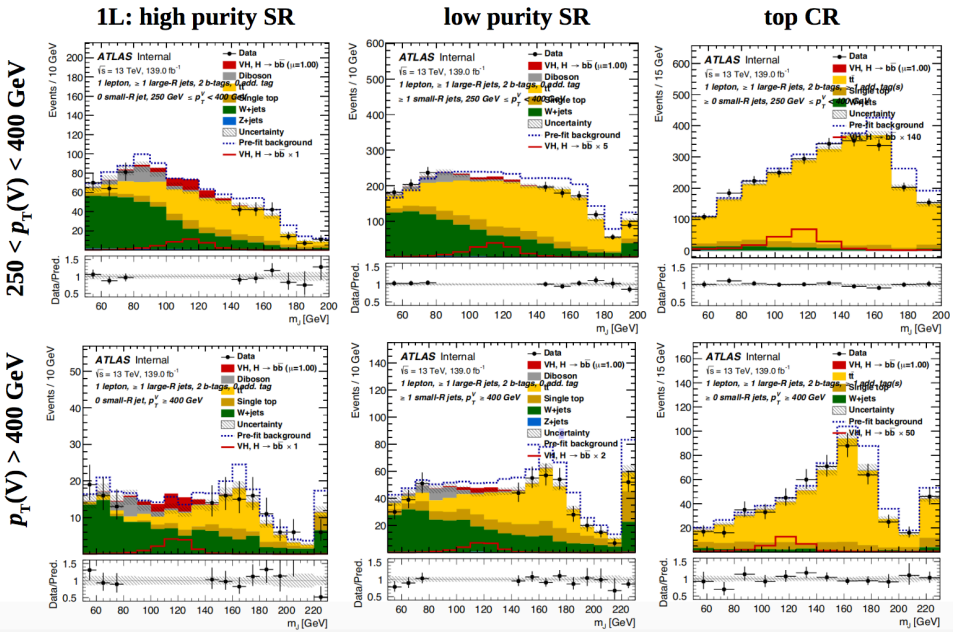
Anti Z+jets cuts:

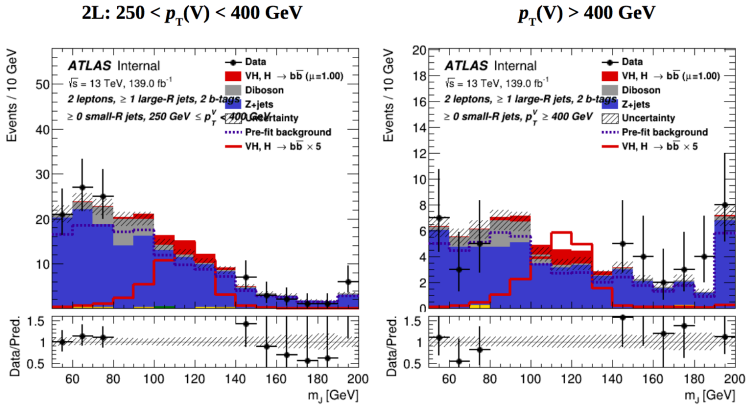
- Z polarisation → $[p_T(L_1)-p_T(L_2)]/p_T(Z) < 0.8$
- $\Delta y(Z, H) < 1.4$

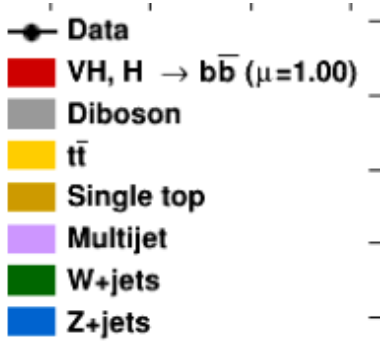
Boosted VHbb: Analysis Regions

Fitting range	0L Add. Calo jets 0	0L Add. Calo jets ≥1	1L Add. Calo jets 0	1L Add. Calo jets ≥1	2L
$250 < p_T(V) < 400$ GeV	HP SR	LP SR	HP SR	LP SR	SR
	top CR		top CR		
$p_T(V) > 400$ GeV	HP SR	LP SR	HP SR	LP SR	SR
	top CR		top CR		







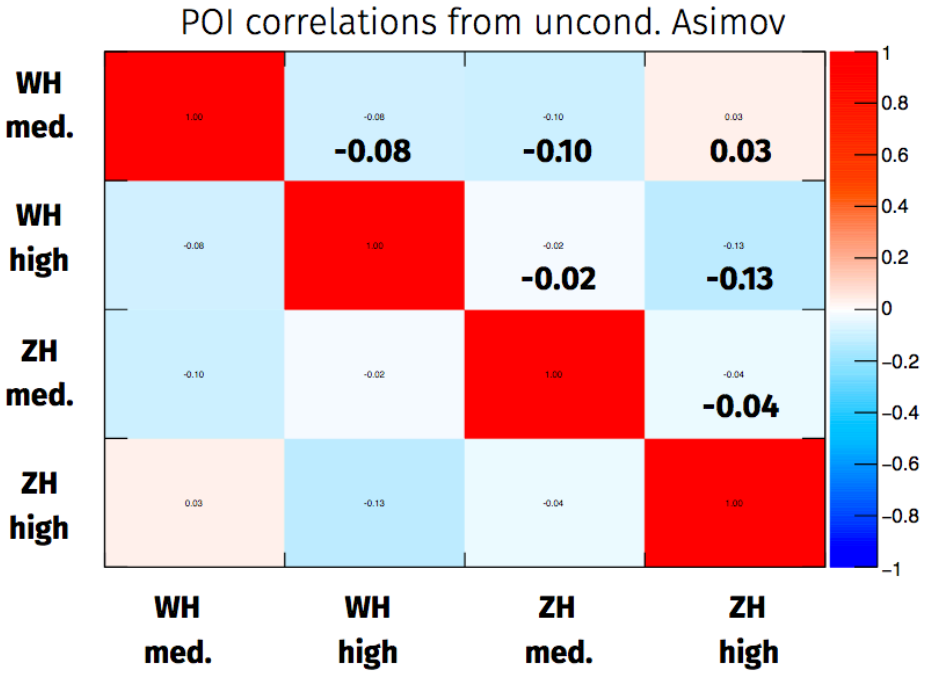


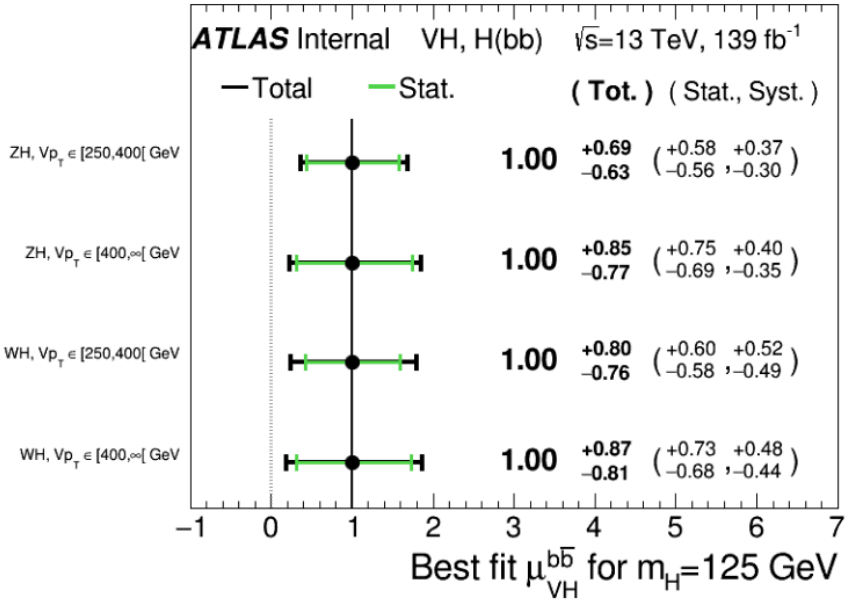
Boosted VHbb: First Look at STXS Fit

Z_{exp}	[250, 400[	[400, ∞[
ZH	1.60 (1.38)	1.31 (1.19)
WH	1.32 (1.38)	1.24 (1.43)

1 POI: Z_{exp} = 3.18 (3.06)

Adding in quadrature: Z_{exp} = 2.86 (2.70)



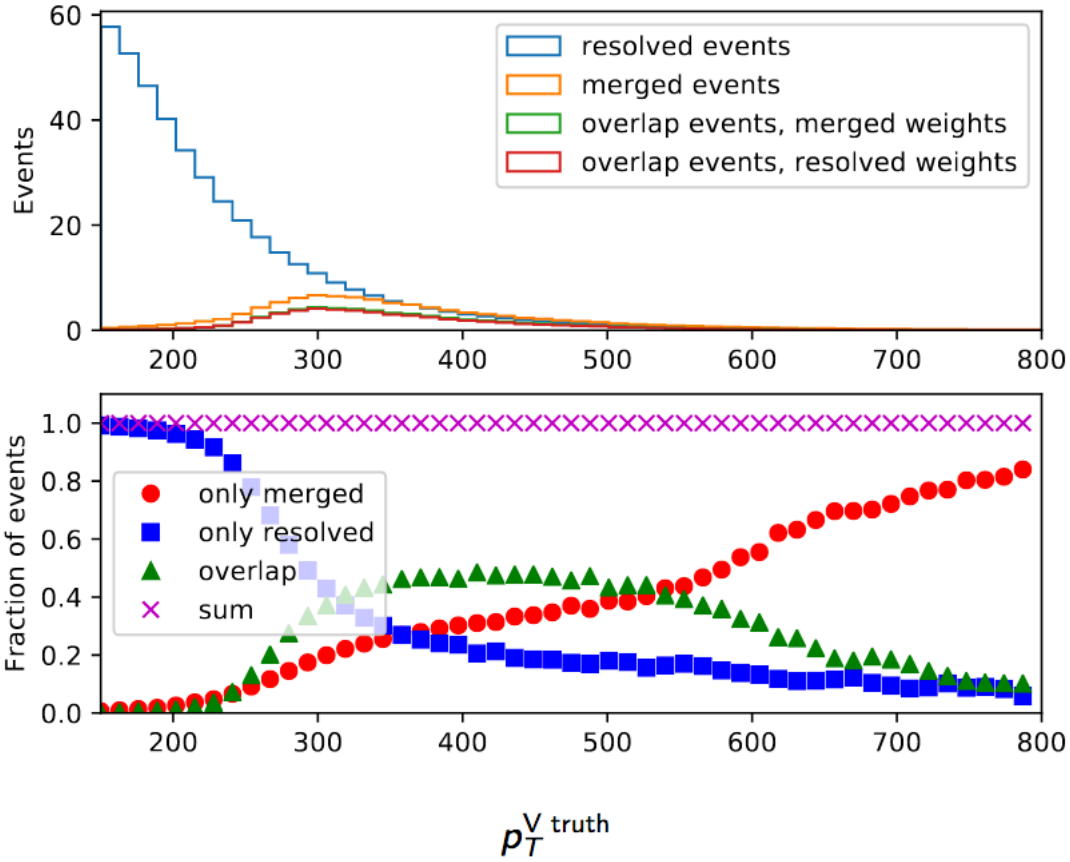


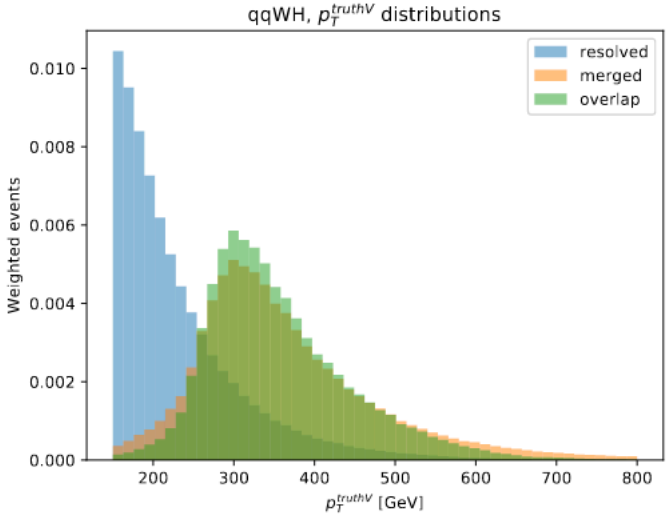
Important:

- Statistical uncertainty includes floating normalizations
- Numbers are Pre-Fit (Post-fit) Asimov

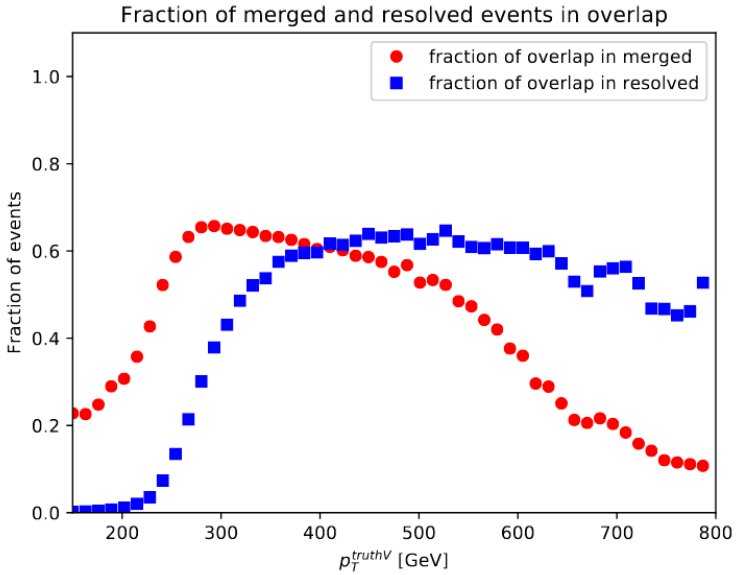
VHbb: Overlap Between Boosted and Resolved Analysis

Study in 1 lepton channel



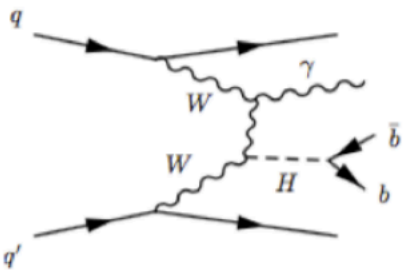


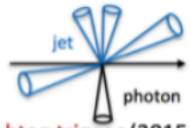
$p_T^{V\ truth}$ histograms normalised to 1



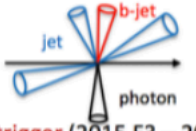
VBF+γ: Event Selection

VBF+gamma: preselection



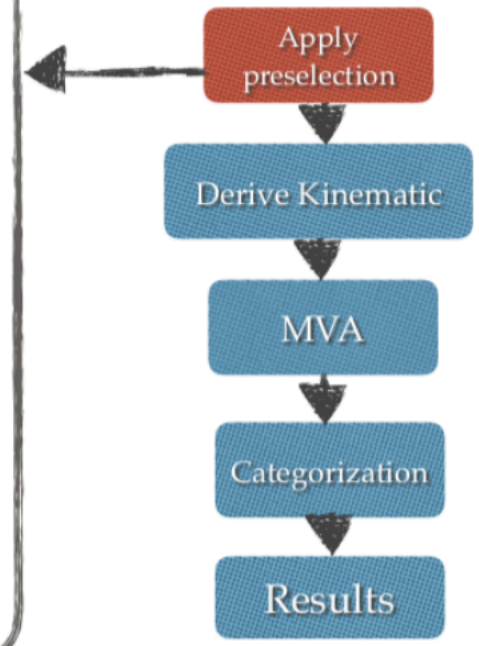


0-btag trigger (2015 F3 – 2016 F)
- 1 photon: EM22 and g25
- 4 jets: 35 GeV
- dijet mass: 700 GeV



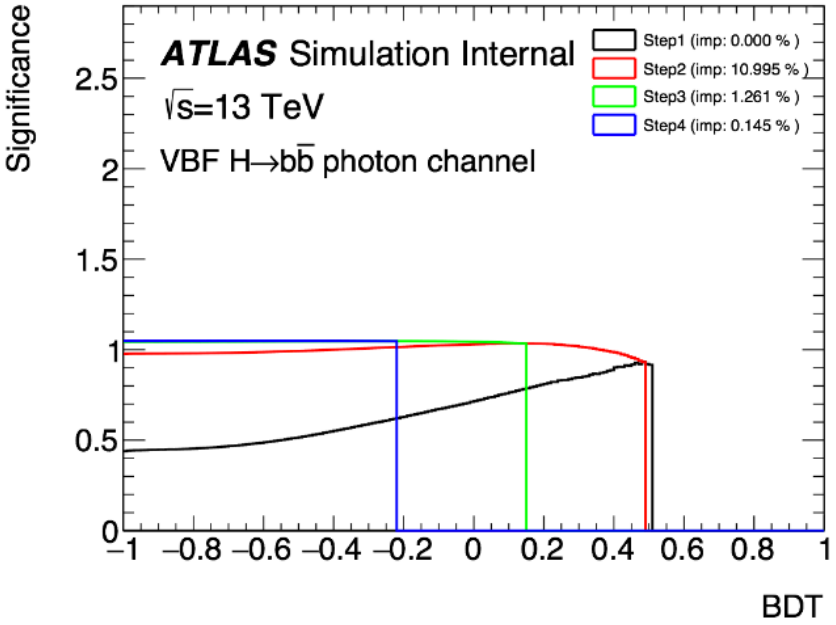
1-btag trigger (2015 F3 – 2018)
- 1 photon: EM22 and g25
- 4 jets: 35 GeV (1 bjet: 35 GeV, 77%)
- dijet mass: 700 GeV

Trigger	Dedicated VBF+gamma triggers
2b-jets	p_T>40 GeV $\|\eta\|$<2.5
2VBF jets	p_T>40 GeV $\|\eta\|$<4.4
Photon	E_T>30 GeV
Event Topology	$p_T(bb)$>80 GeV m_{jj}>800 GeV Lepton veto

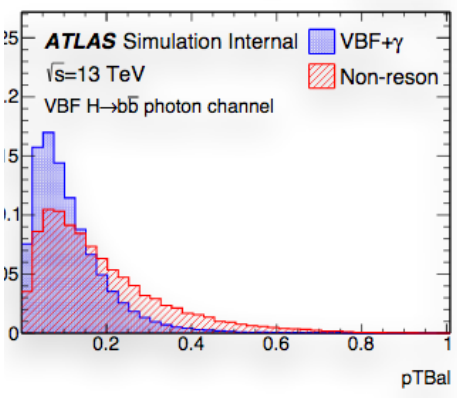


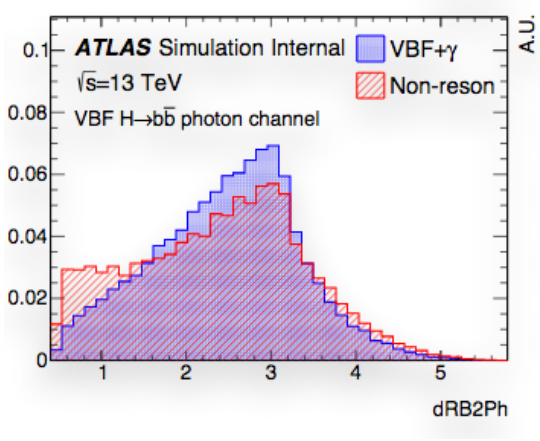
+ VBF+γ: Categorisation BDT

- Sequential optimization (only partial BDT bins):
  - Step1: only count SRI
  - Step2: only count SRI+SRII
  - ….

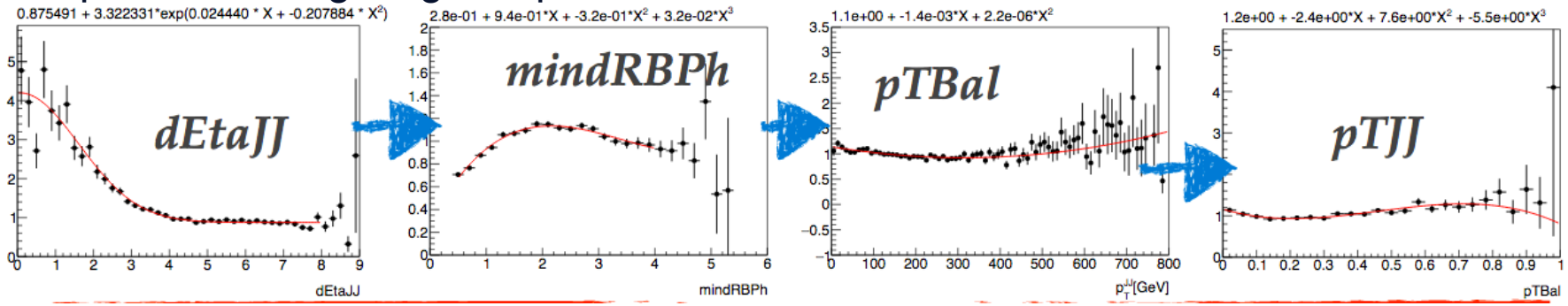


10 variables: '**mJJ**', '**pTBal**', '**dEtaJJ**', '**cenPhJJ**', '**dRB2Ph**', '**dPhiBBJJ**', '**dRB1Ph**', '**cosThetaC**', '**pTJJ**', '***dRB1J1***'



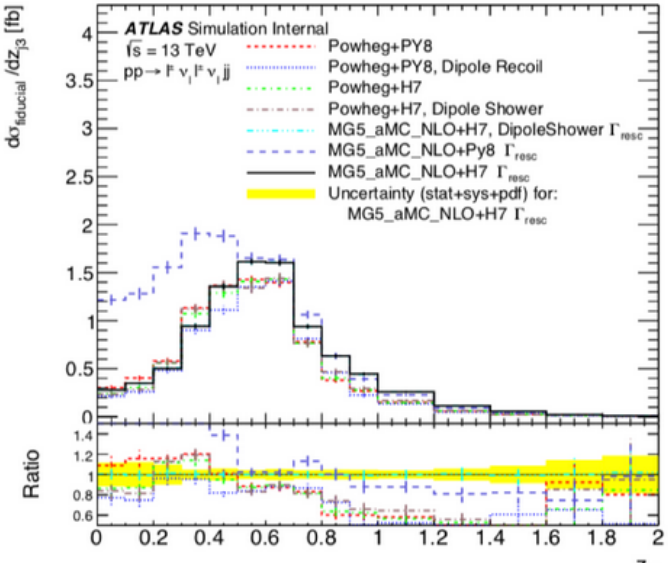


Sequential reweighting of input variables:

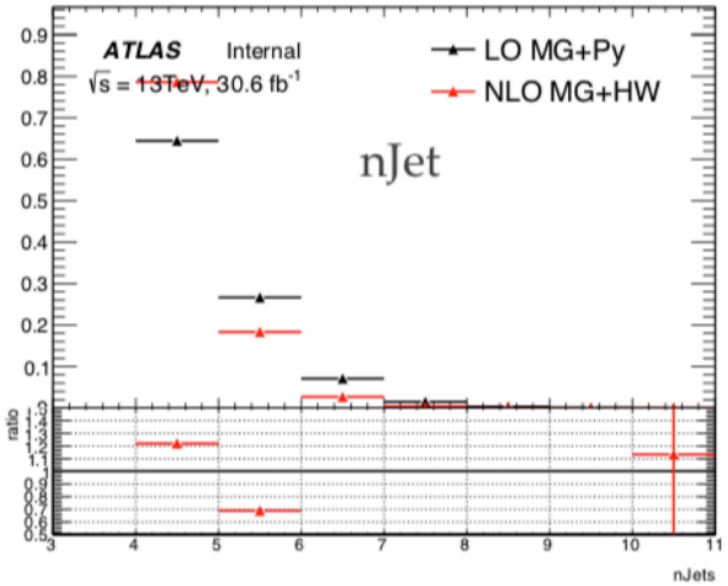


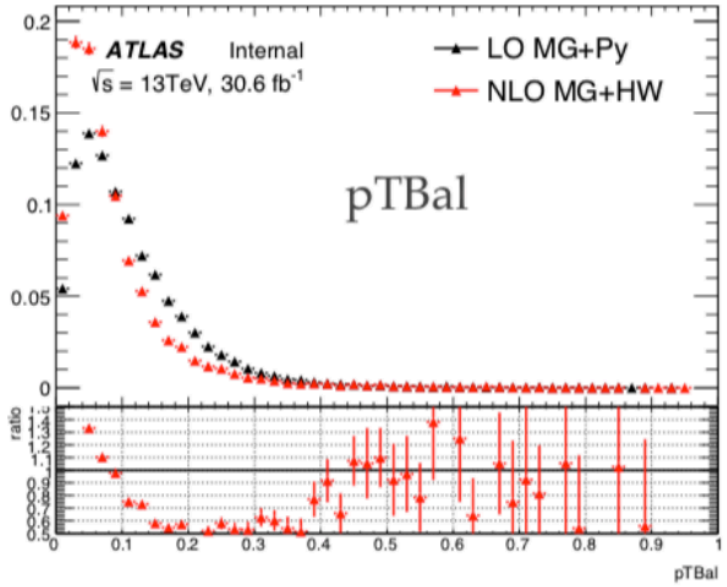
VBF+γ: NLO Signal Sample

From PMG

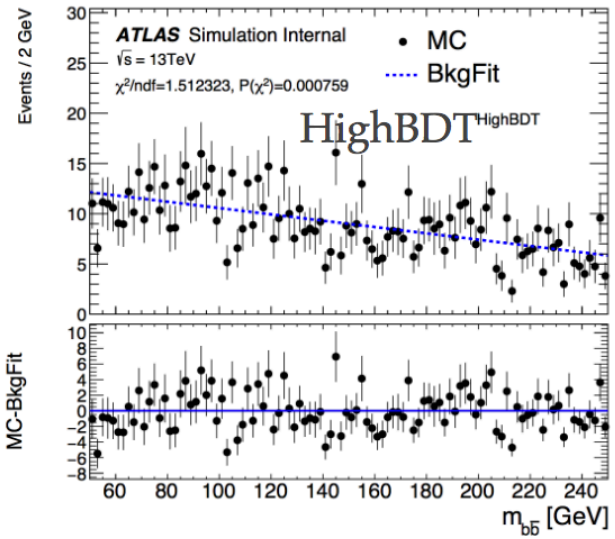


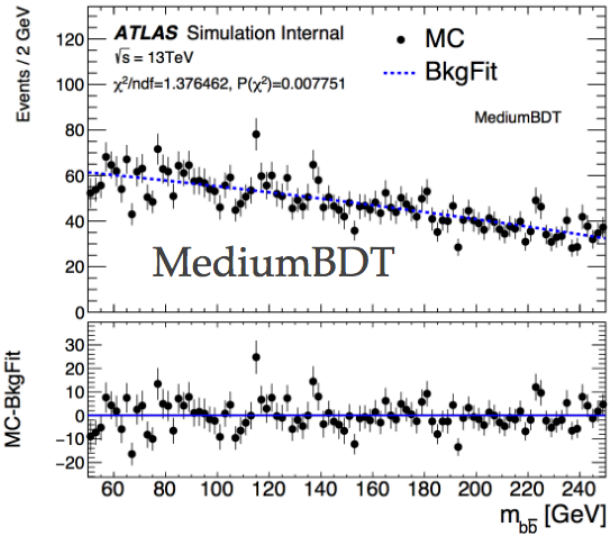
From VBF+γ:

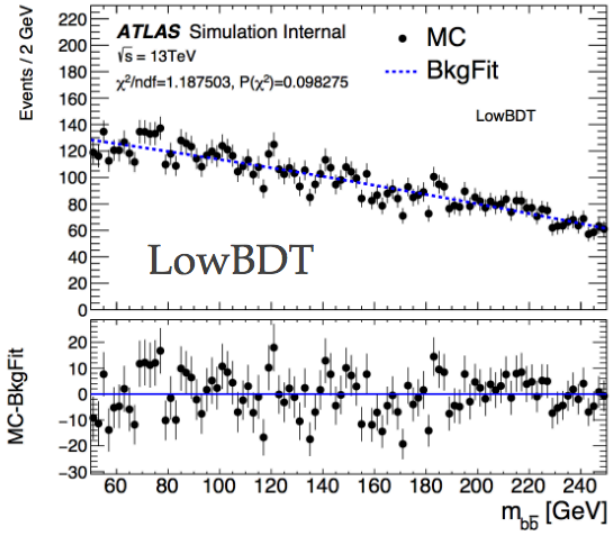


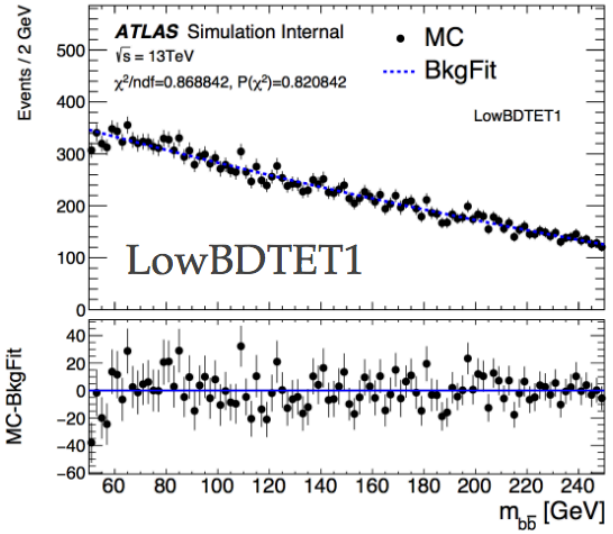


+ VBF+γ: Background Modelling





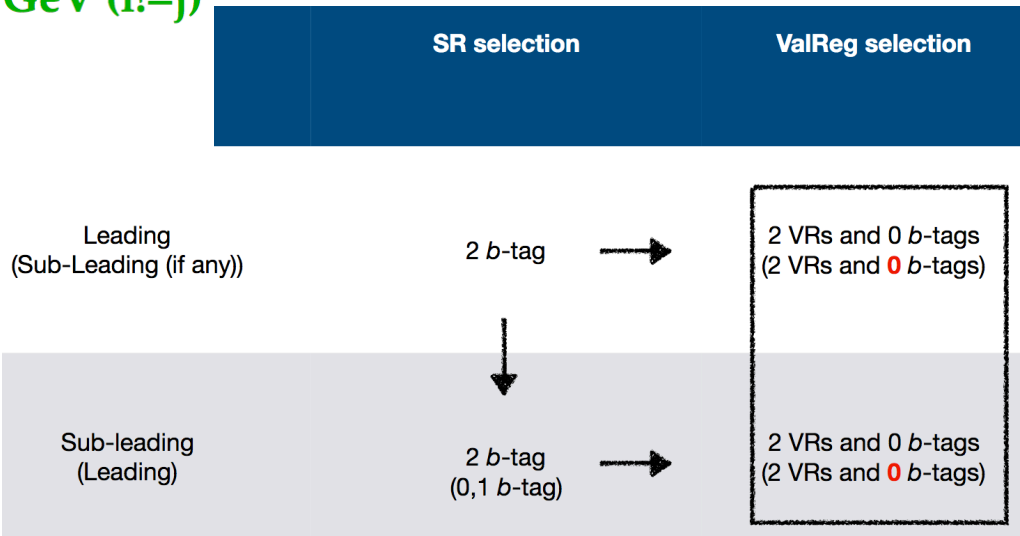




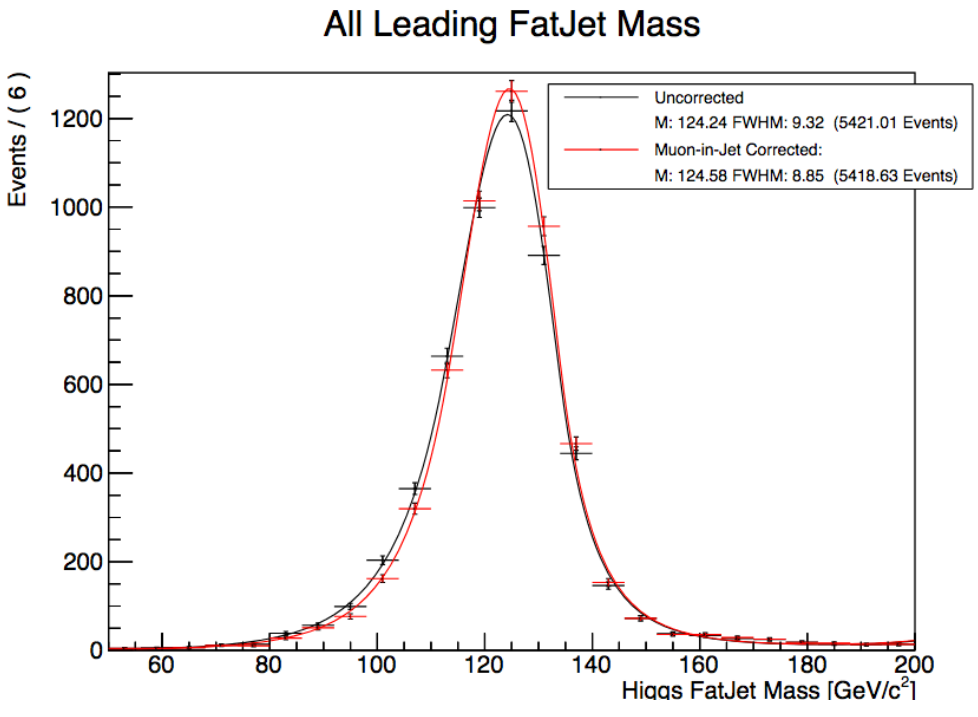
+ Hadronic Hbb+ISR: Event Selection and Categorisation

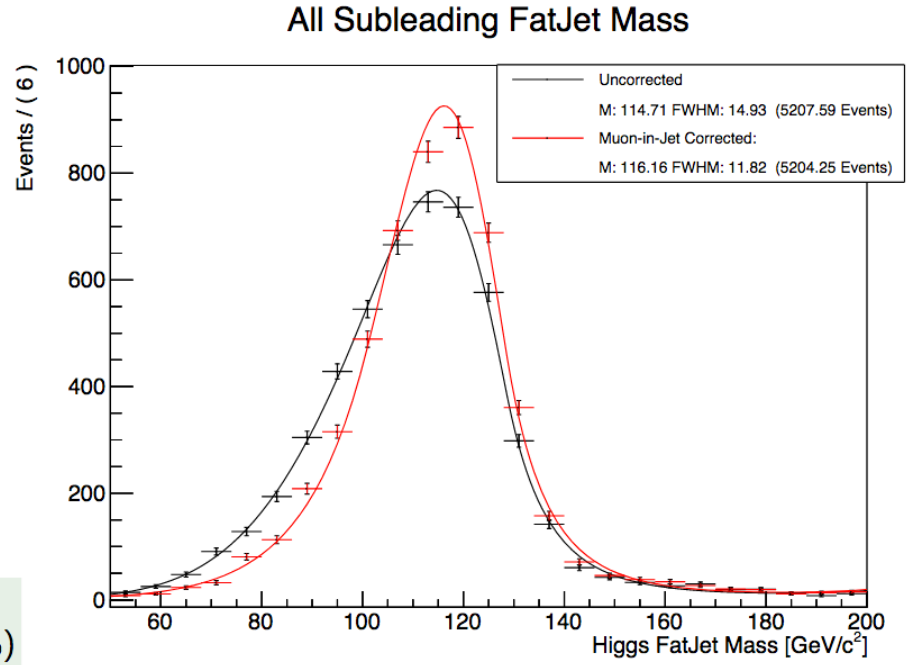
New

- **Trigger strategy and offline cuts: pT >440 GeV && M>45 GeV**
- At least two fat-jets (FJ) with pT > 250 GeV
- FJ and variable-R jet (VR) selection, list of candidates with **p_T >250 GeV && M>60 GeV**:
  - 2M/pT<1
  - at least two 10-GeV VRs
  - **ΔR(VR_i,VR_j) > min(R_VR_i, R_VR_j), with i=two leading VR and j=any VR above 5 GeV (i!=j)**
- **If leading has 2 b-tag (@77%) -> SRL;**
  If subleading has 2 b-tag (@77%) -> SRS;



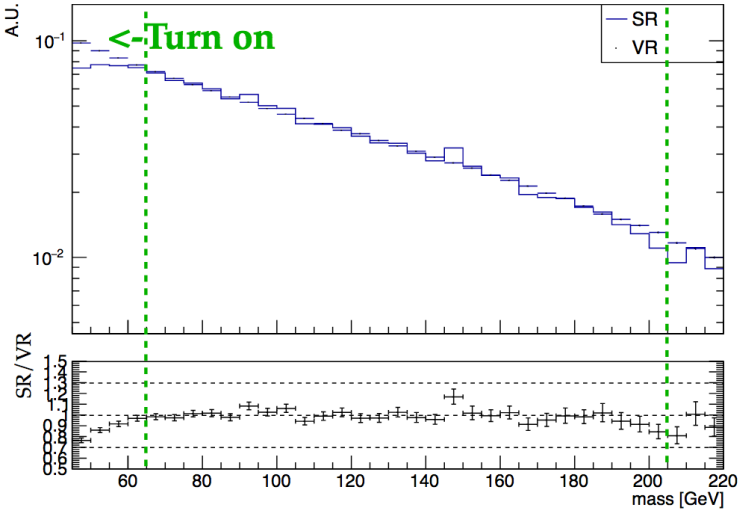
Hadronic Hbb+ISR: Muon-in-Jet Correction



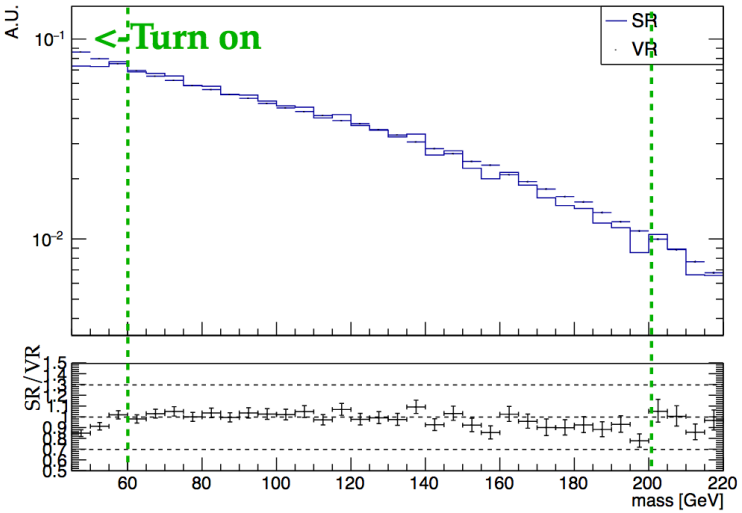


Hadronic Hbb+FSR: Background Modelling

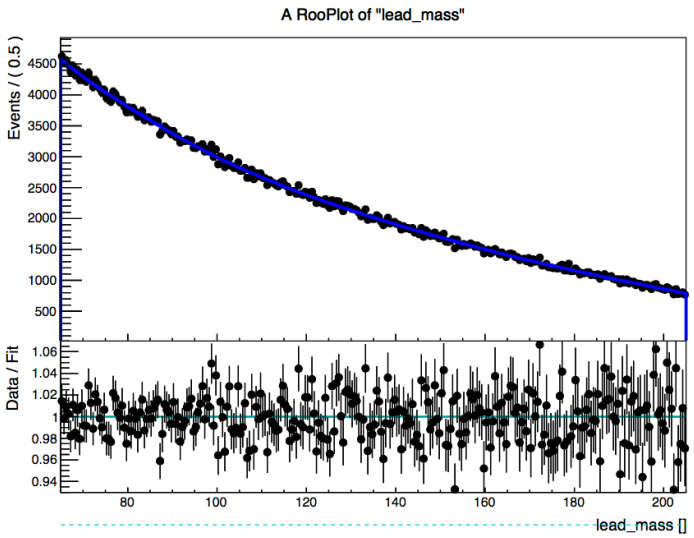
Leading



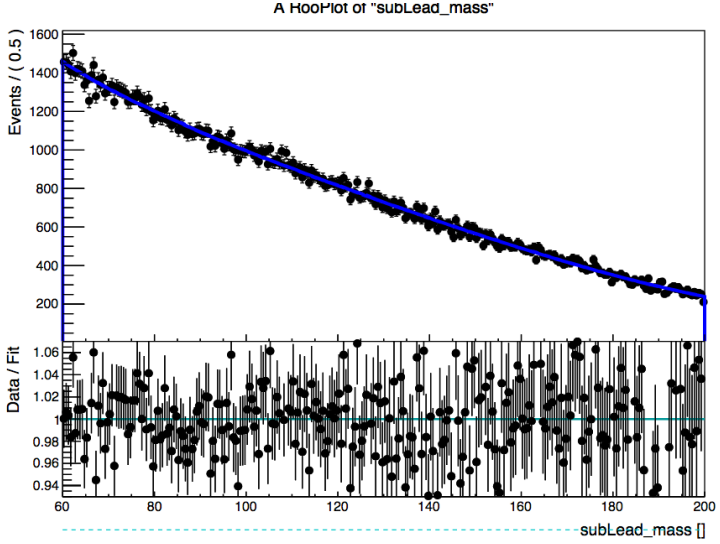
Subleading



$Mass_{lead}$ [65,205]

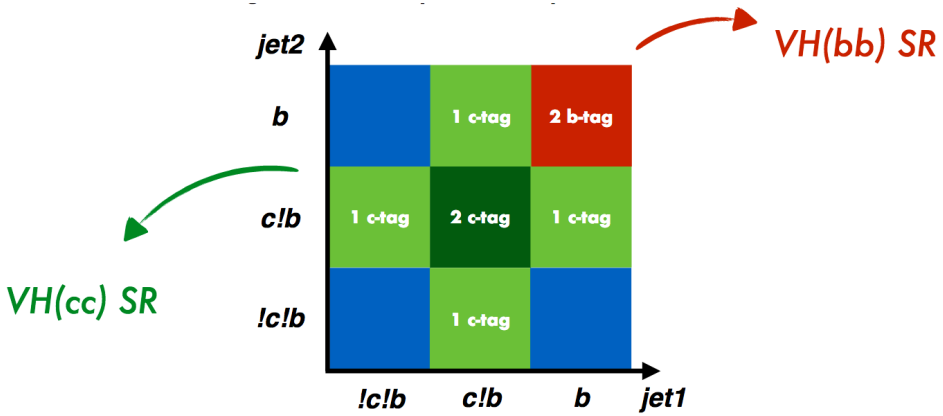


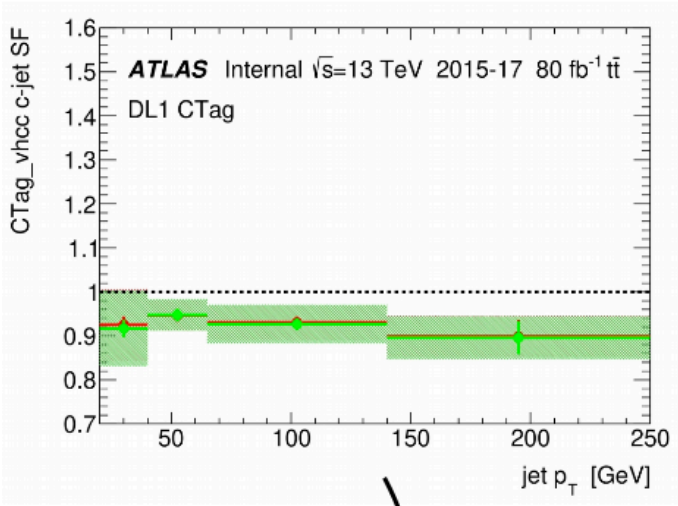
$Mass_{sublead}$ [60,200]



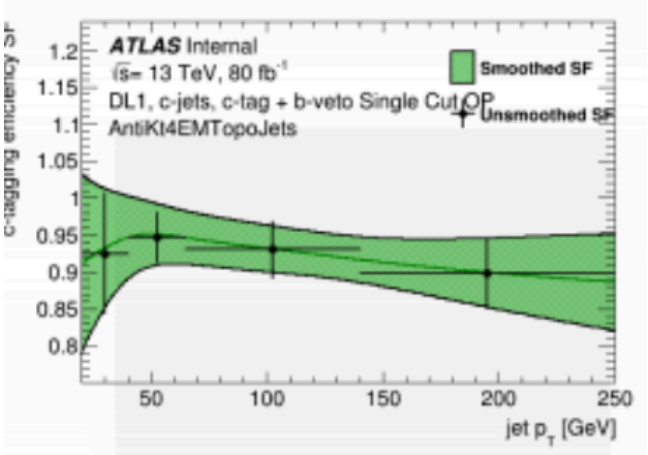
VHcc: Flavour Tagging

- Combine c-tagging with DL1 and b-veto with MV2c10 (orthogonality with VHbb)
- Custom flavour tagging calibrations performed in analysis group



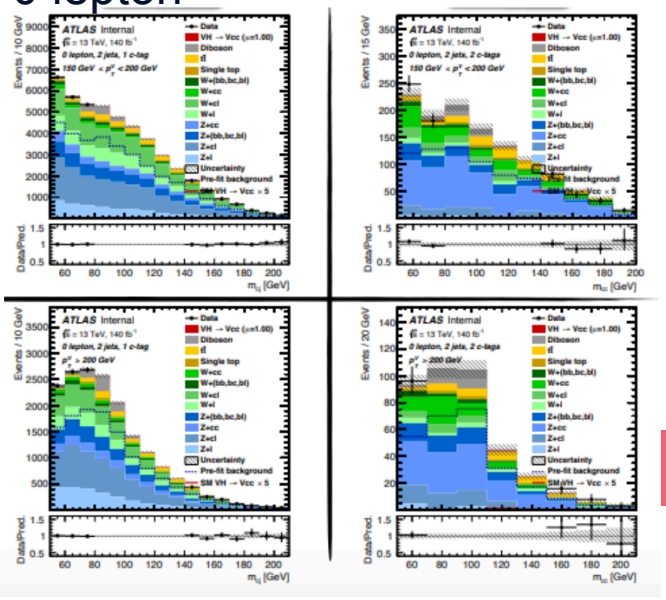


c-jets



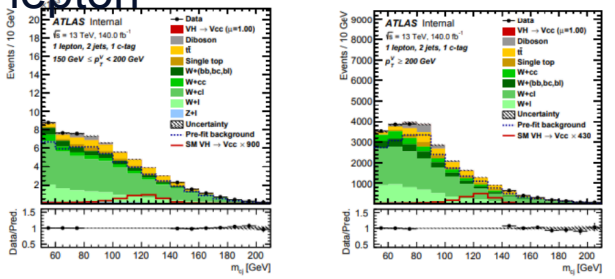
VHcc: Fit and Exp. Results

0 lepton



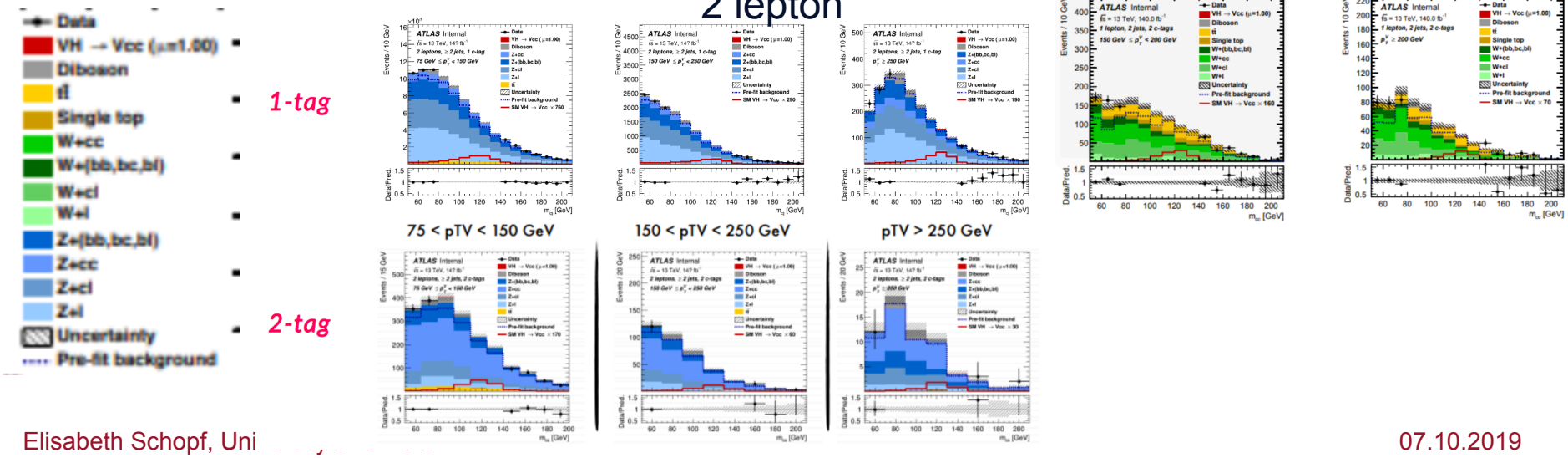
Current limit	0-lep	1-lep	2-lep
Systematics	33 X SM	43 X SM	40 X SM

1 lepton



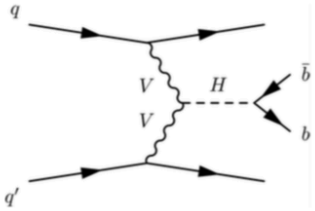
Very preliminary!

2 lepton

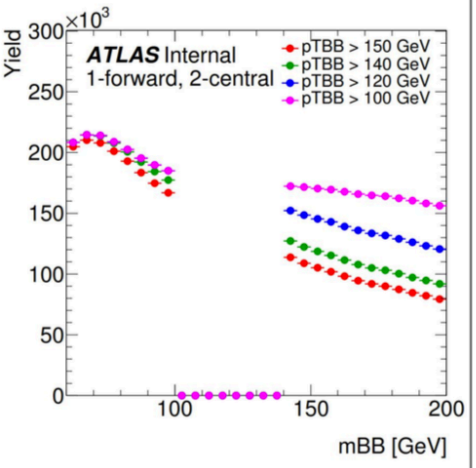


All-hadronic VBFHbb: Analysis Strategy

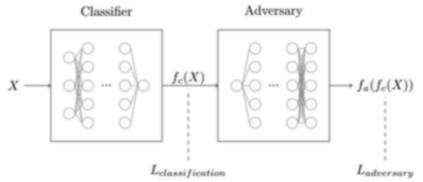
Step 1



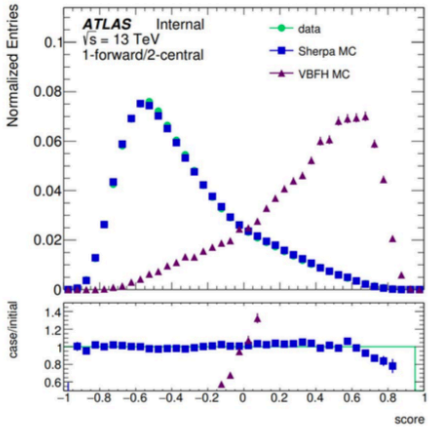
Loose preselection, based on trigger, orthogonality requirements, and pTBB



Step 2

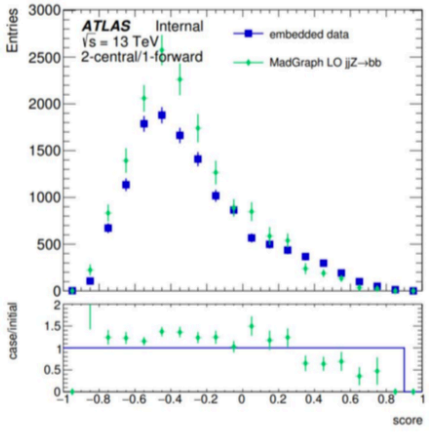


MVA-based categorization, using variables with low mBB-correlation

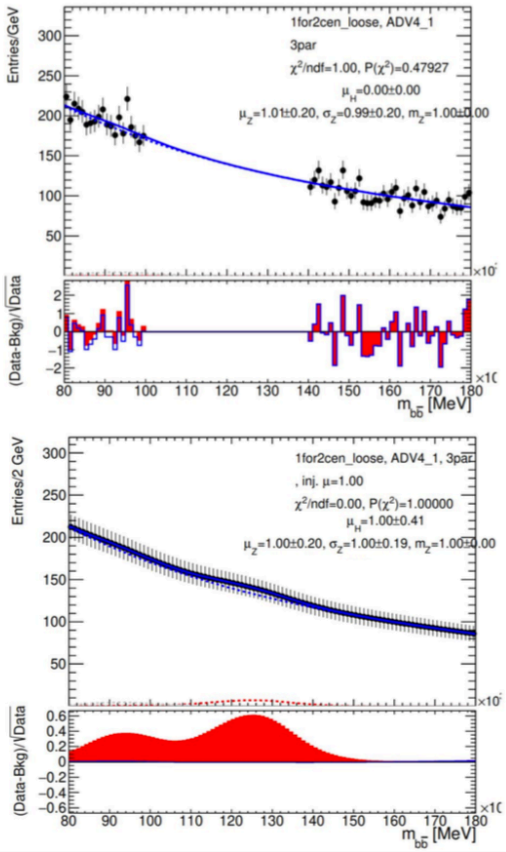


Step 3

Zμμ→Zbb embedding for Zbb estimate

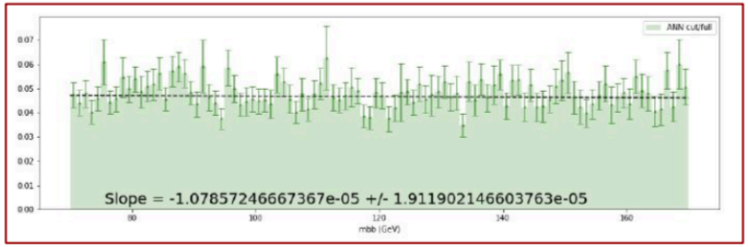


Step 4: Fit mBB distribution in each region



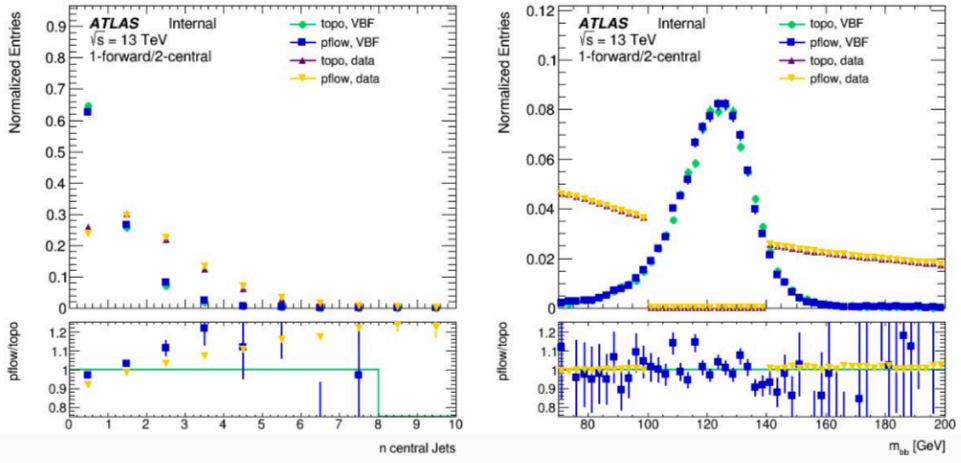
All-hadronic VBFHbb: Developments

Adversarial neural net to decorrelate mBB from MVA classifier-will allow us to fit background in SRs simultaneously



Example performance on multijet MC: shape difference between events with top 5% of classifier score and inclusive distribution

Switched to PFlow jets - slightly better signal/background, with lower systematics



Most pieces of analysis in hand for primary trigger channel, finalizing optimization of the MVA approach and adding uncertainties.

Timeline:

- Finalize first pass at analysis by end of 2019 and request EB
- Circulate to ATLAS by May 2020