Boosted $H \to bb$ tagging searches

F. Parodi on behalf of the ATLAS Collaboration

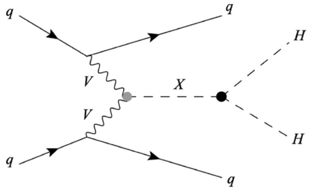
Università di Genova/INFN Genova

Why $H \to b\bar{b}$ (boosted) ?

- Search for new particles decaying to Higgs bosons (i.e. $X \to HH$)
- Not a single golden channel but $b\bar{b}$ final state has the highest BR (often selected as decay channel of at least one of the two Higgs)

	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

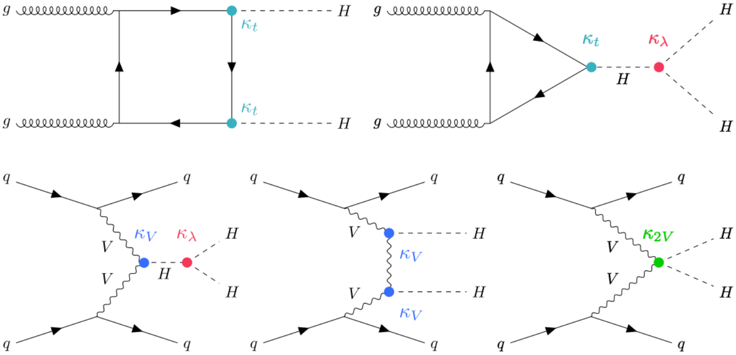
- High HH-mass systems produce boosted Higgs ($m_{HH} > 1$ TeV)



- The analysis presented in this talk exploits vector-boson fusion production to enhance $S/\sqrt{B}$

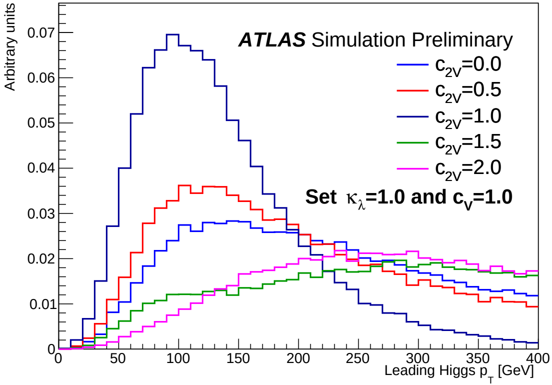
Non resonant HH

- Non resonant HH production relevant for Higgs self-coupling ($\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$) and Higgs coupling to vector bosons ($\kappa_{2V} = \lambda_{VVHH}/\lambda_{VVHH}^{SM}$). In particular the ggF production mode is sensitive to κ_λ while VBF production mode is sensitive to both κ_λ and κ_{2V}.



Higgs boson p_T vs κ_{2V}

- Boosted topologies very sensitivity to κ_{2V}

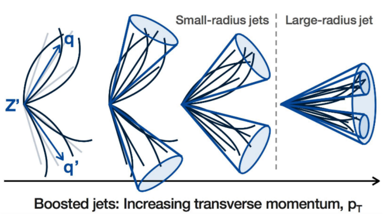


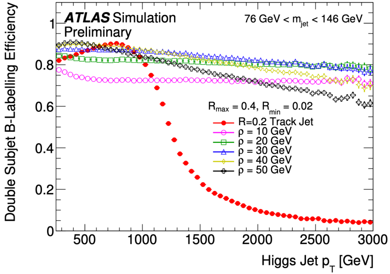
Leading Higgs boson pT when varying κ_{2V}

Jet reconstruction for boosted jets

Large-R jet:

- *b*-jets from decay high p_T particles are merged in a single large-radius jet
  - Dedicated ATLAS calorimeter jet reconstruction with a (fixed) radius of R=1 (anti k_T algorithm).





Jet substructure

- Dedicated track jet collection with variable size : $R = \rho/p_T$ (jet)
- Defined by three parameters: $\rho = 30$ (dimensionless constant), R_{min}=0.02 (minimal size) and R_{max}=0.4 (maximal size)
- Exploit tracking-only info: variable-radius track jets (VRTrack-jet)

ATL-PHYS-PUB-2017-010 [http://cds.cern.ch/record/2268678/files/ATL-PHYS-PUB-2017-010.pdf]

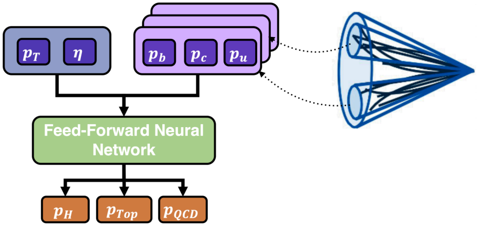
Boosted bb tagger

Xbb tagger:

- feed-forward NN that combined flavour tagging discriminants from subjets
- calibrated on ATLAS Run 2 data
- used in most of the Run 2 results

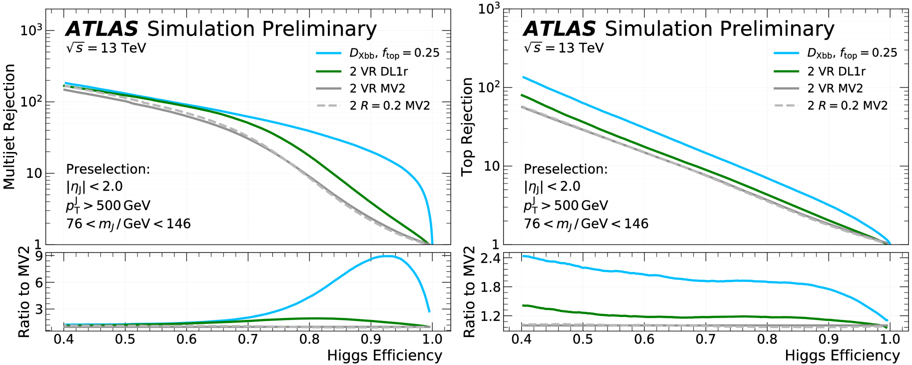
Tagger definition

- Individual subjets (VRTrack jets) are tagged using single btagging DL1r algorithm (DNN) optimised for VR jets
- Jet information + DL1r output nodes of max 3 subjets are fed to the net



Performance

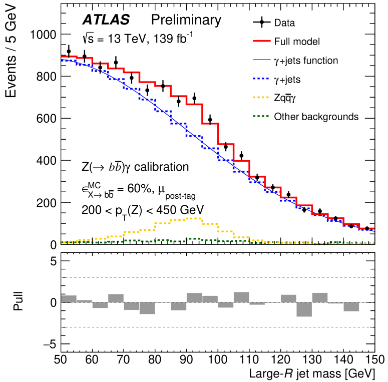
- $H \to b\bar{b}$ vs $t\bar{t}$ and multijet:

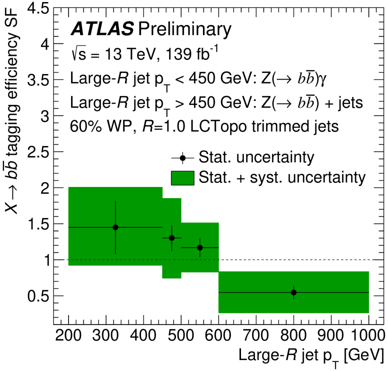


ATL-PHYS-PUB-2020-019 [http://cds.cern.ch/record/2724739/files/ATL-PHYS-PUB-2020-019.pdf]

Xbb calibration

- Signal calibration on: Z(bb)+γ (low p_T), Z(bb)+jets (high p_T)



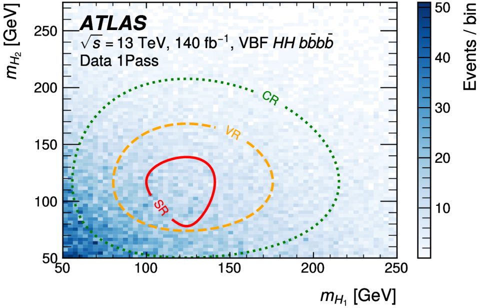


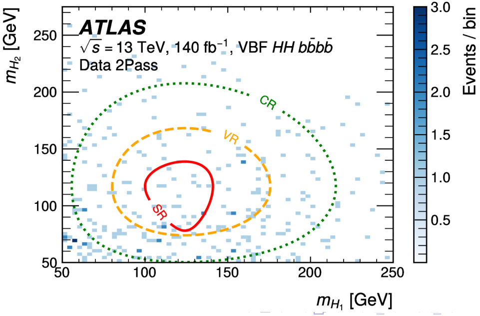
- Background calibration on $t\bar{t}$ events
- Validation on $g \to b\bar{b}$ events

ATL-PHYS-PUB-2021-035 [http://cds.cern.ch/record/2777811/files/ATL-PHYS-PUB-2021-035.pdf]

$HH \to b\bar{b}b\bar{b}$ analysis: event selection

- Trigger: at least 1 Large-R jet
- VBF topology: two Higgs candidates + two VBF jets
- Higgs candidates: Large R-jet, tagged with Xbb. At least two candidates: H_1 leading ($p_T > 450$ GeV), H_2 sub-leading ($p_T > 250$ GeV).
- Signal/control regions:
  - **2Pass**: both candidates pass the 60% Xbb working point, **1Pass** only one of the two candidates satisfies the selection
  - **Signal Region**/**Validation Region**/**Control Region**: defined in the $m_{H_1} - m_{H_2}$ plane





$HH \to b\bar{b}b\bar{b}$ analysis: multivariate discriminant

- Boosted decision tree to separate signal/background in the SR

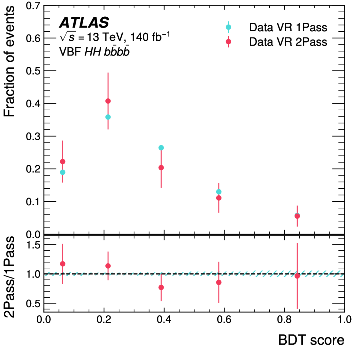
Physics objects	BDT input variables
Higgs Boson Candidate (H_i, $i = 1, 2$)	$p_T^{H_i}$, η_{H_i}
Di-Higgs System (HH)	p_T^{HH}, η_{HH}, m_{HH}
VBF Jets (j_i, $i = 1, 2$)	$p_T^{j_i}$, η_{j_i}, E_{j_i}

- Non-resonant analysis: training with $k_{2V} = 0$
- Resonant analysis: mass-parameterized BDT (pBDT) to accommodate multiple resonant signals

- Background processes in the SR predominantly originate from non-resonant multijet production (b or lighter quarks + 10% $t\bar{t}$)
- Multijet background estimated using a data-driven method and 1Pass events:
  - BDT shapes compatible between 1Pass and 2Pass
  - Derive multijets in 2Pass/SR using 1Pass/SR events (signal "contamination" negligible) applying a normalization factor

$$w = N_{1Pass\,CR}/N_{2Pass\,CR}$$

  - Uncertainty from the independent w estimate in VR



- Sistematic uncertainties: analysis is statically limited. Main systematic uncertainty comes from Xbb calibration uncertainty (20-30%).

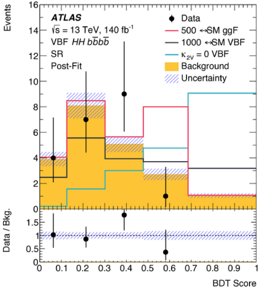
$HH \to b\bar{b}b\bar{b}$: non-resonant analysis

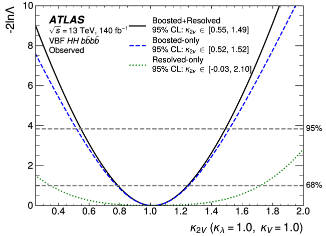
- Binned maximum-likelihood fit to the BDT distributions in the 2Pass/SR

- κ_{2V} (95% CL intervals):
  observed: $0.52 < \kappa_{2V} < 1.52$
  expected: $0.32 < \kappa_{2V} < 1.71$

  Combining with *ggF* and *VBF* categories of the resolved analysis:
  observed: $0.55 < \kappa_{2V} < 1.49$
  expected: $0.37 < \kappa_{2V} < 1.67$

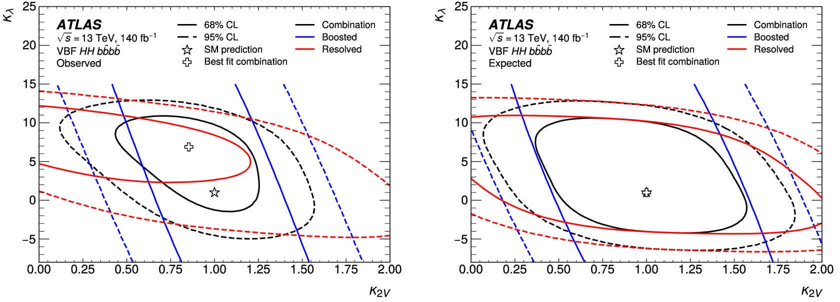
  $\kappa_{2V} = 0$ excluded at 3.4σ.





$HH \to b\bar{b}b\bar{b}$: non-resonant analysis

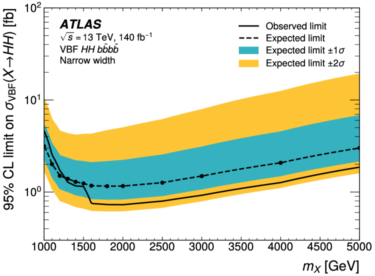
- Boosted analysis is dominant for κ_{2V} sensitivity while κ_λ sensitivity is driven by the resolved analysis

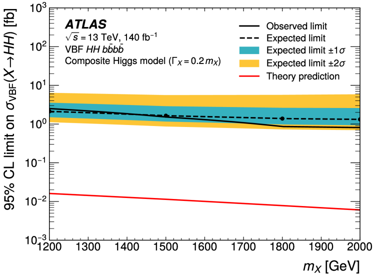


- Results on κ_{2V} analysis are as good as the HL-LHC projections of the previous-best VBF HH analysis (the full Run 2 resolved VBF hh4b).

$HH \to b\bar{b}b\bar{b}$: resonant analysis

- Two resonance-width (Γ_X) hypotheses are considered:
  - narrow-width signal (Γ_X smaller than the detector resolution - 5-6% of the resonance mass)
  - broad-width signal ($\Gamma_X = 20\%$ of the resonance mass, based on the Composite Higgs model





- No excess is found in the region 1-5 TeV. Exclusion limits are set on the production cross-section:
  - Narrow-width [obs. (exp.)]: 4.6 fb (3.1 fb) for $m_X = 1$ TeV to 1.9 fb (3.0 fb) for $m_X = 5$ TeV
  - Broad-width [obs. (exp.)]: 2.5 fb (2.1 fb) for $m_X = 1.2$ TeV to 0.8 fb (1.3 fb) for $m_X = 2$ TeV

Conclusion

- Latest ATLAS results on $HH \to b\bar{b}b\bar{b}$ (VBF channel, both Higgs candidates boosted) are presented
  - Non-resonant analysis: most sensitive constrain to κ_{2V}, evidence of $\kappa_{2V} \neq 0$
  - Resonant analysis: no excess found in the range 1 $TeV < m_{HH} < 5$ TeV, exclusion limits on cross-sections

Outlook

- The analysis is statistically limited, significant improvement expected with:
  - Run3 statistics
  - improved $b\bar{b}$ tagger based on GNN (GN2X, see N. Kumari talk on Monday)