New physics searches with boosted objects tagging at the ATLAS experiment

C-LHCP workshop

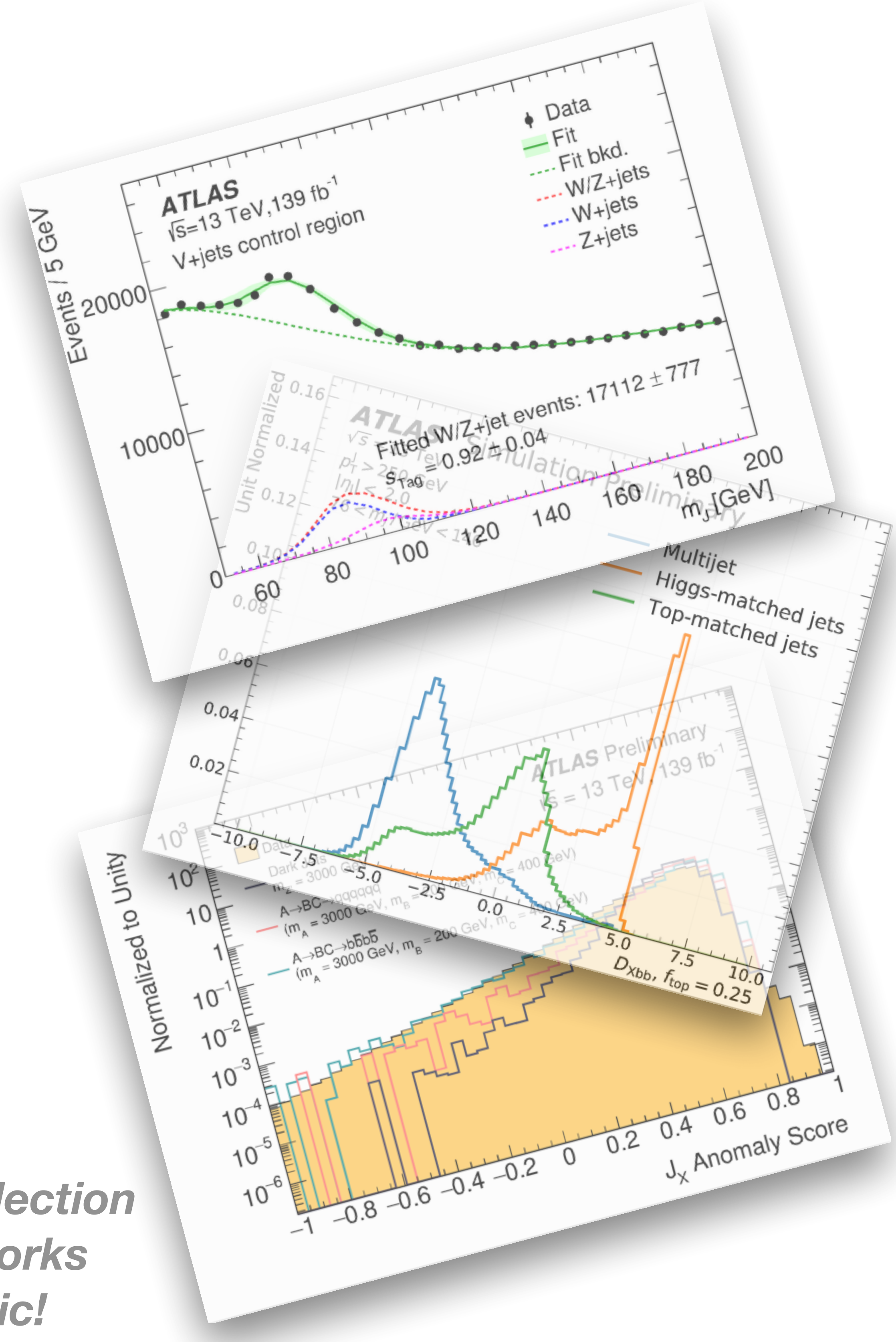
24th November 2021

Antonio Giannini
University of Science and Technology of China

Outline

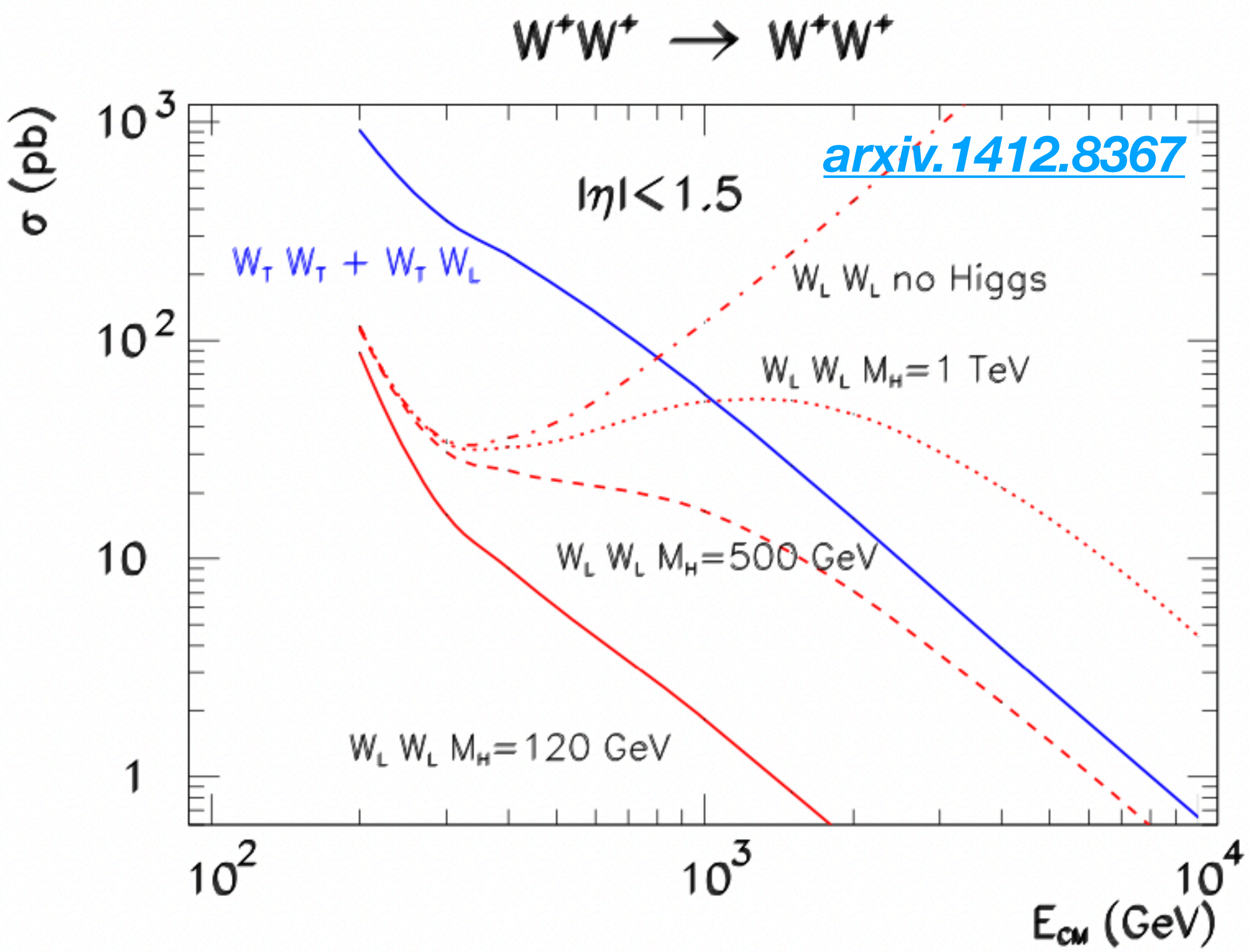
New physics searches with boosted objects tagging at the ATLAS experiment

- Why do we need boosted object tagging?
- Latest results about boosted jet tagging:
    - techniques and calibrations
    - double flavour tagging (bb)
    - anomaly detection
- Interplay of the novel techniques with the physics search results

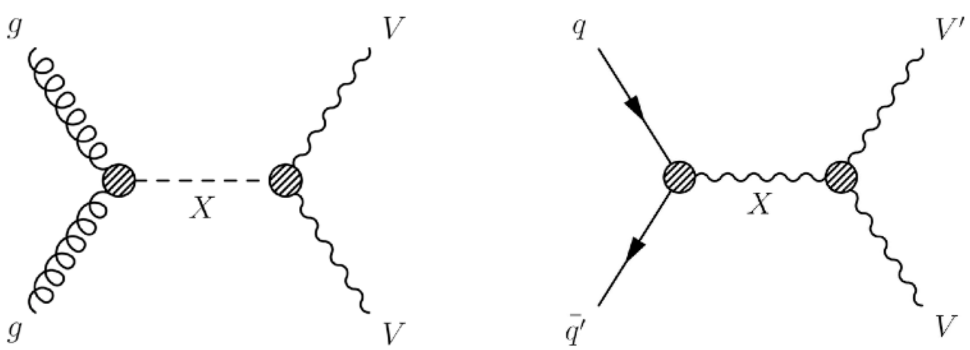


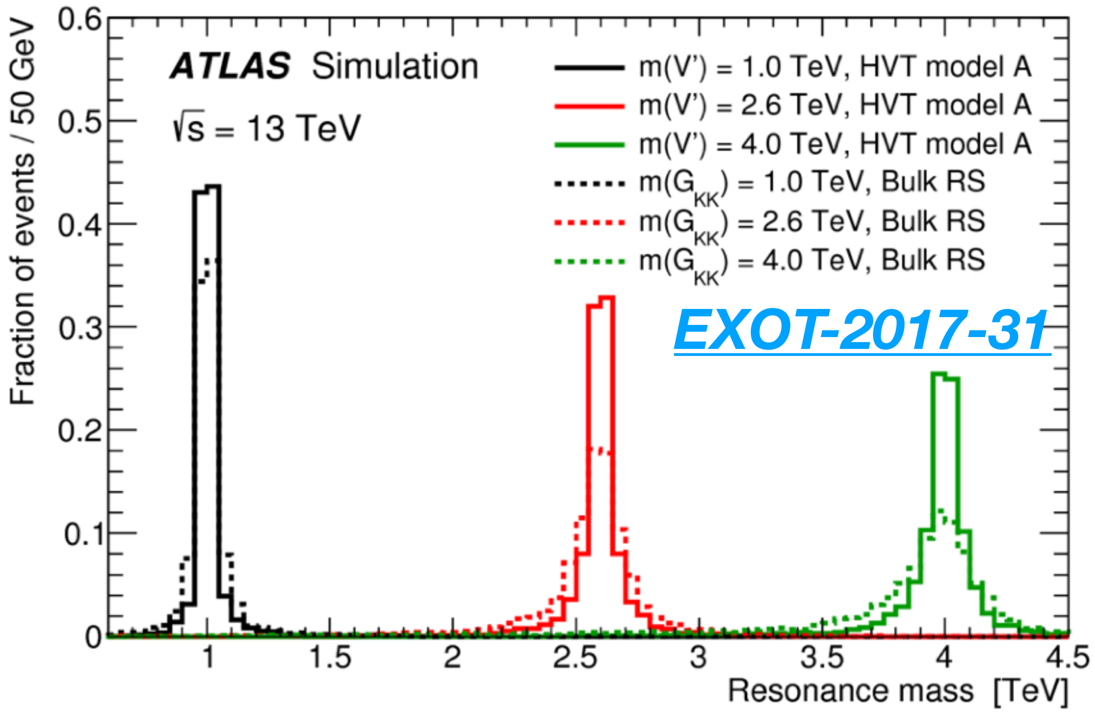
****** The material is a biased selection of all the approaches and works done in ATLAS on this topic!***

Why high-mass searches?



- Diboson interactions are a key process in the LHC program
  - Higgs discovery has a been a milestone probe of the SM
- However, ***diboson physics*** may still hide New Physics (NP)!!

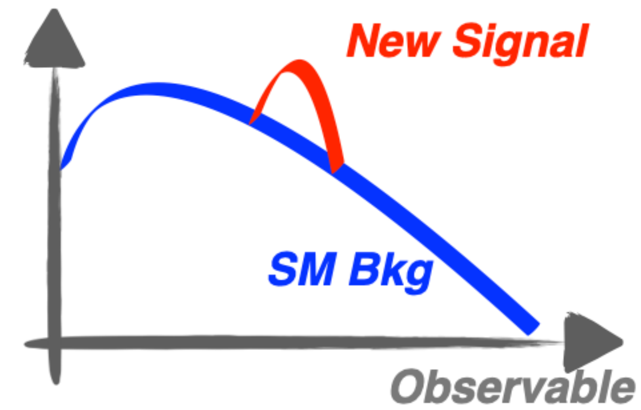




- Several theoretical models foresee the presence of NP beyond the SM
- Heavy new particles decaying into VV pairs:
  - ***Spin-0***, Randal Sundrum (RS) Radions
  - ***Spin-1***, Heavy Vector Triplet (HVT)
  - ***Spin-2***, RS Gravitons

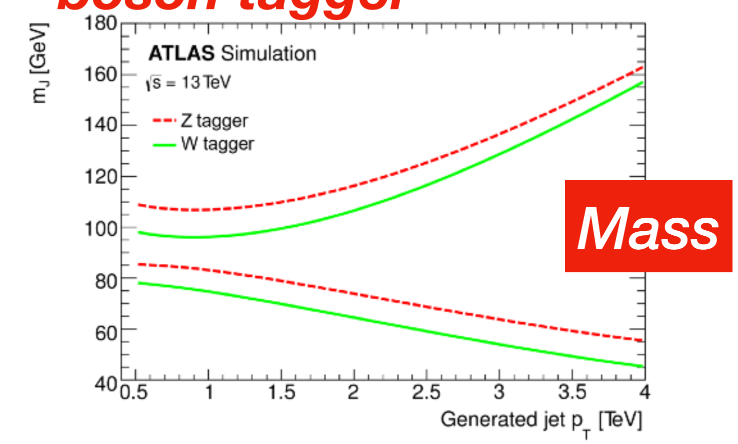
How high-mass searches?

- Many high mass diboson searches have been performed using run-2 data
  - [PUB-2021-018](), [ATLAS-CONF-2022-028]()
- The overall strategy is to look for bumps in the invariant mass spectrum in several final states:
  - ***VV*** ==>> fully leptonic, semi-leptonic, ***fully-hadronic***
  - VH ==>> semi-leptonic, fully hadronic
  - ***XH*** ==>> ***fully-hadronic***
- ***Why hadronic final states?***
  - higher BR of the bosons but more challenges to reduce the background processes!
- At high-mass, i.e. high-pT of the bosons, the decay products are boosted and merged reconstructed in ***one large-R jet (1.0)***

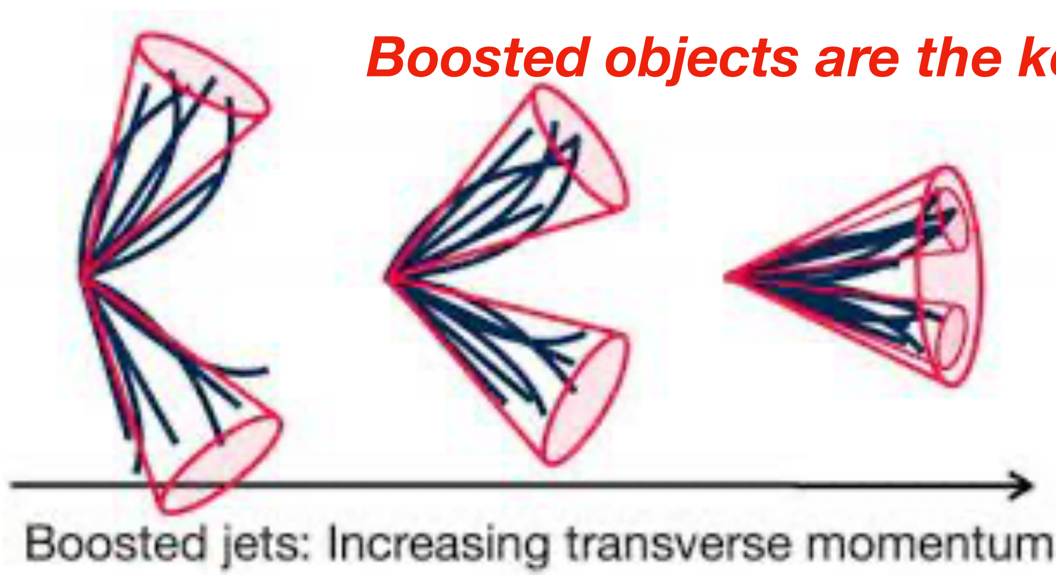


Historical W/Z boson tagger

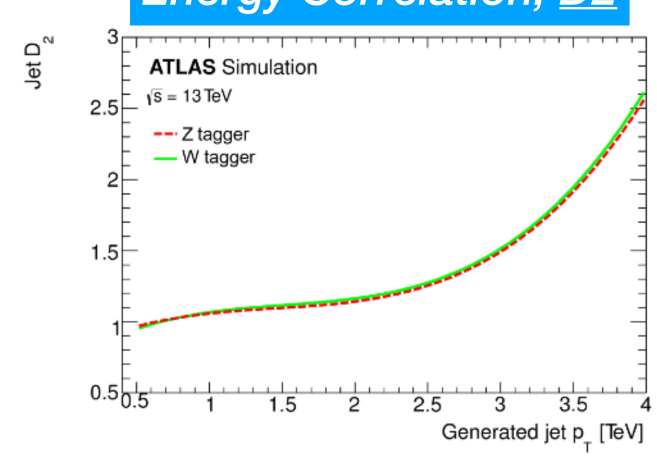
JHEP09(2019)091



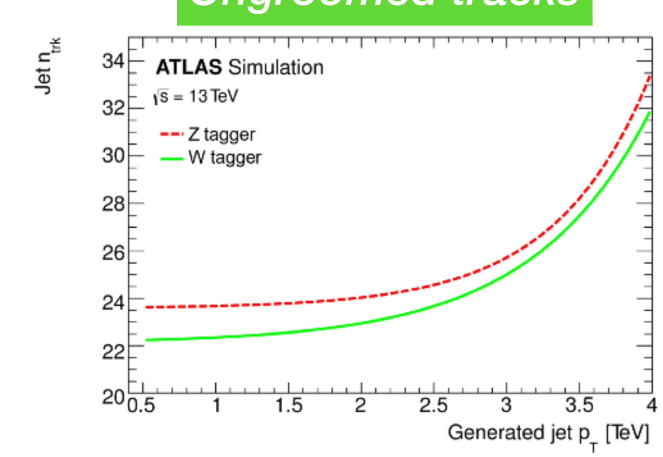
Boosted objects are the key!



Energy Correlation, D2



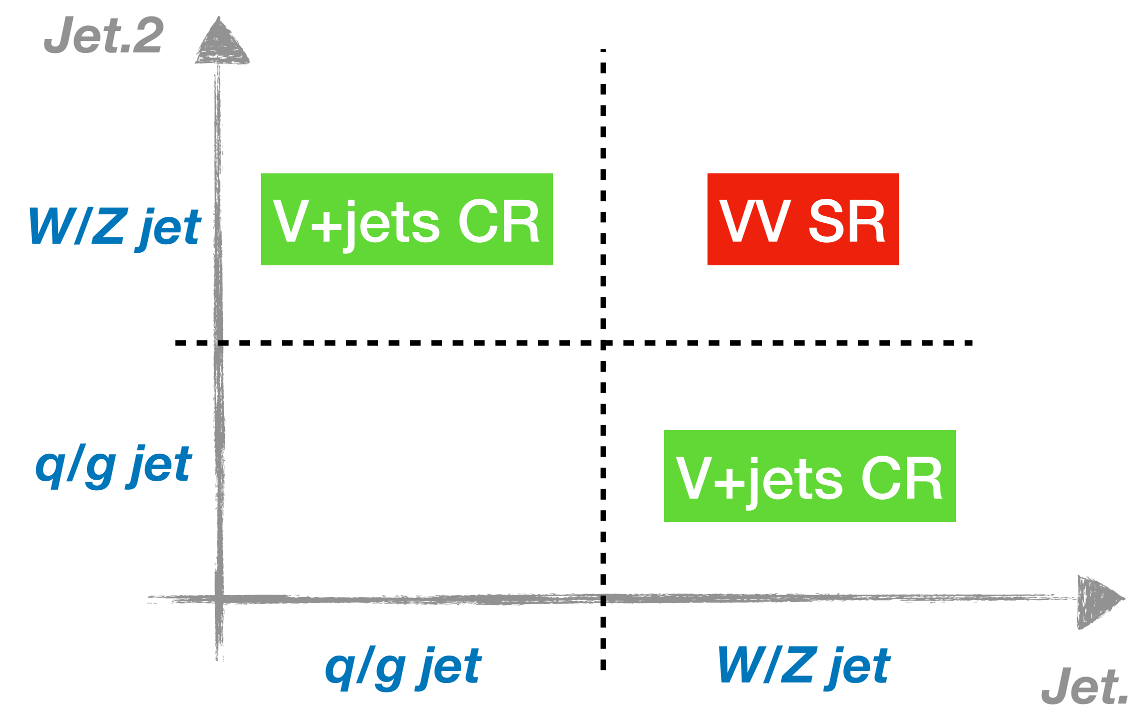
Ungroomed tracks

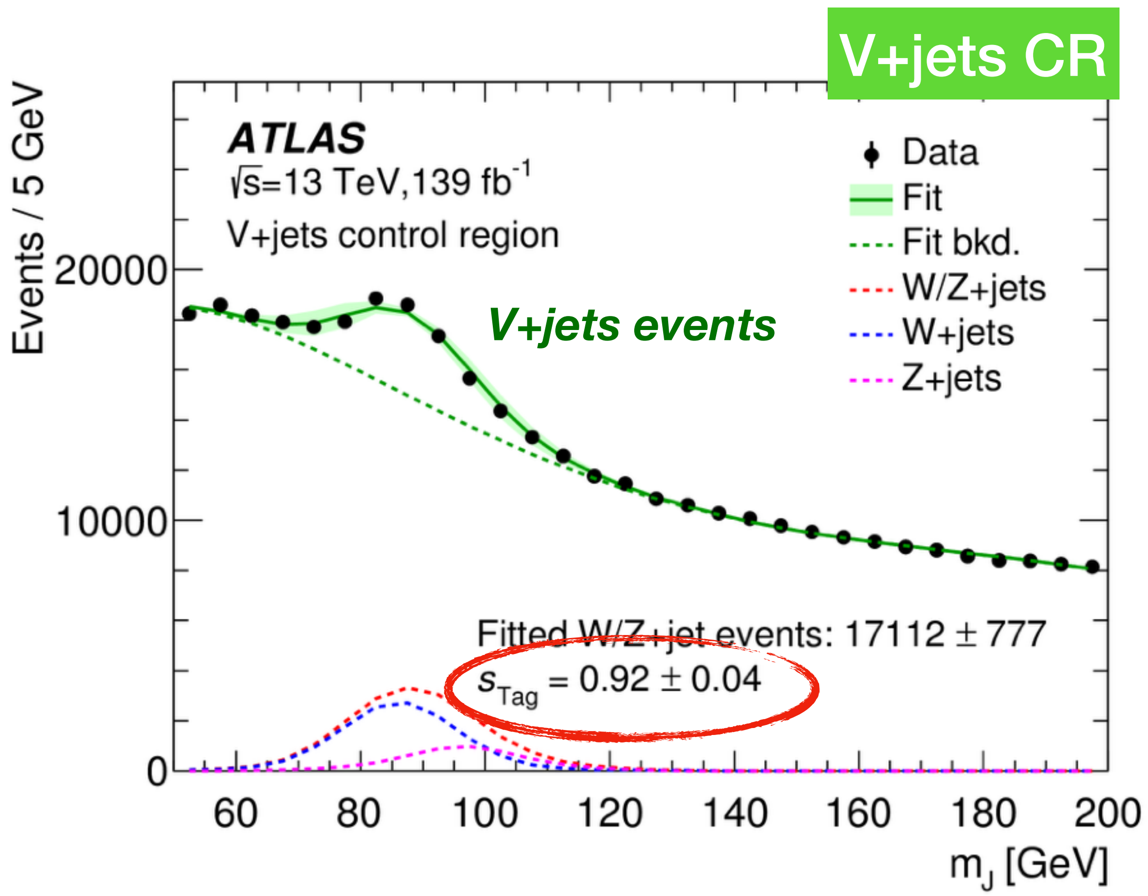


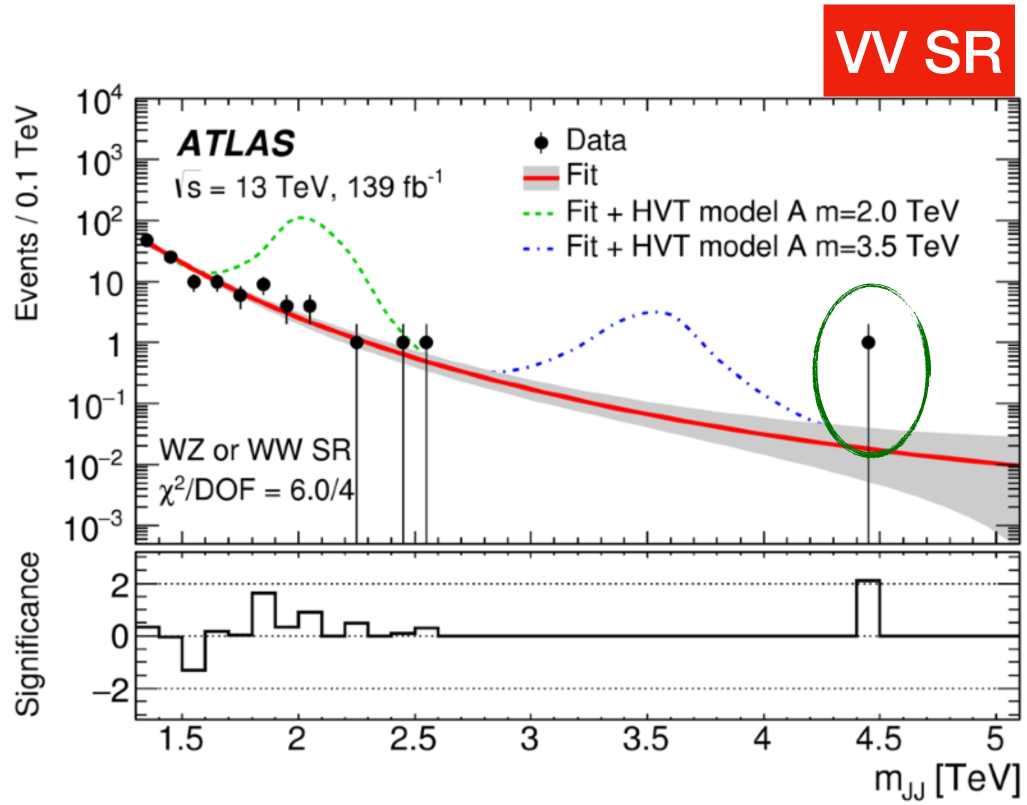
W/Z jet tagger calibration

[JHEP09(2019)091](#)

- Jet tagging tools are really improved using sophisticated reconstruction algorithms and ***JetSubStructure*** (JSS) information:
  - high bkg rejection
  - need to calibrate the tagging algorithm
- Other calibration approach available here, [***ATL-PHYS-PUB-2020-008***](#)





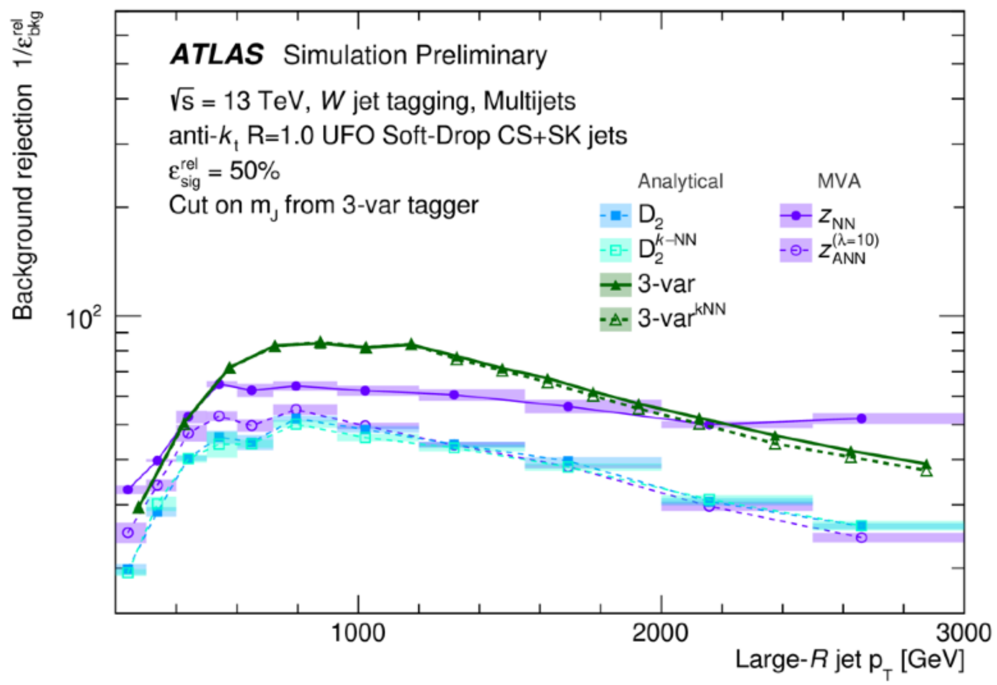


We collect data at very high m_{JJ} range

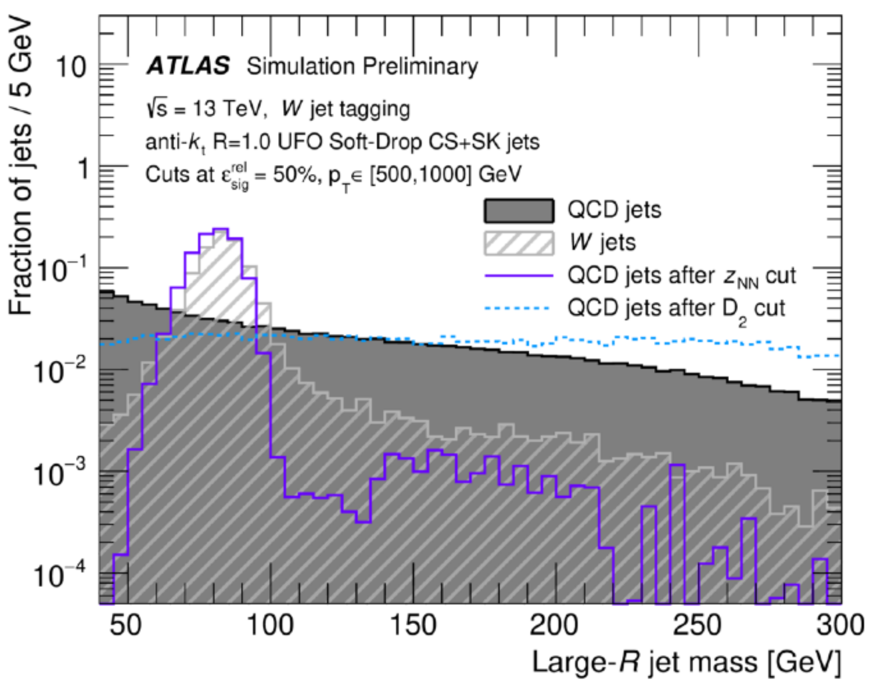
Moving forward: Machine Learning

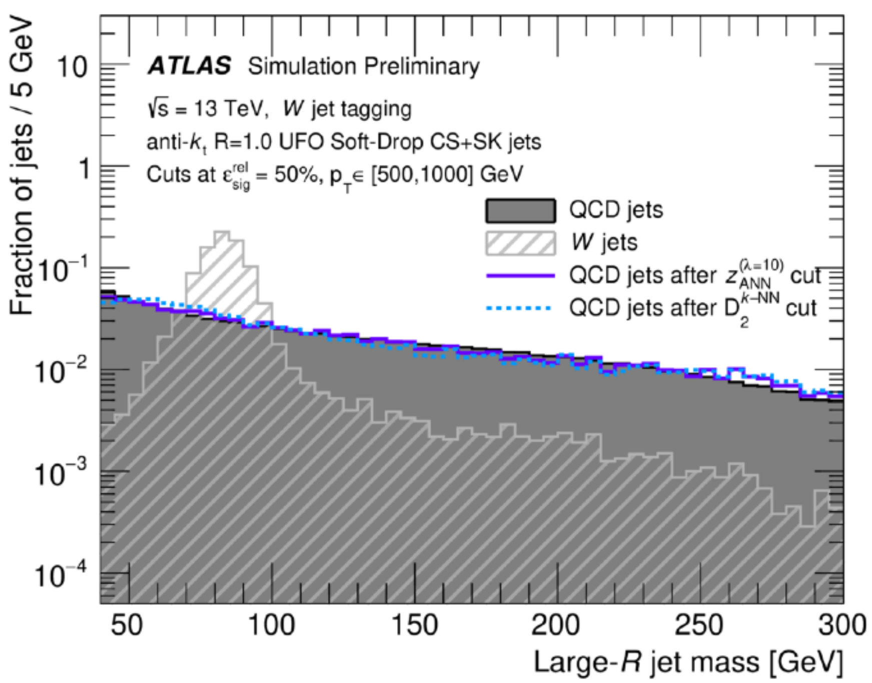
ATL-PHYS-PUB-2021-029

- The ATLAS community has been hardly working on improving large-R jets and tagging performances
  - new ***UnifiedFlowObjects (UFO)*** reconstruction, ready for late run-2 analyses, [*Eur. Phys. J. C 81 (2021) 334*]
  - ***new boson tagger approaches***, [*ATL-PHYS-PUB-2021-029*]



Adversarial-NN to decorrelated w.r.t. JetMass





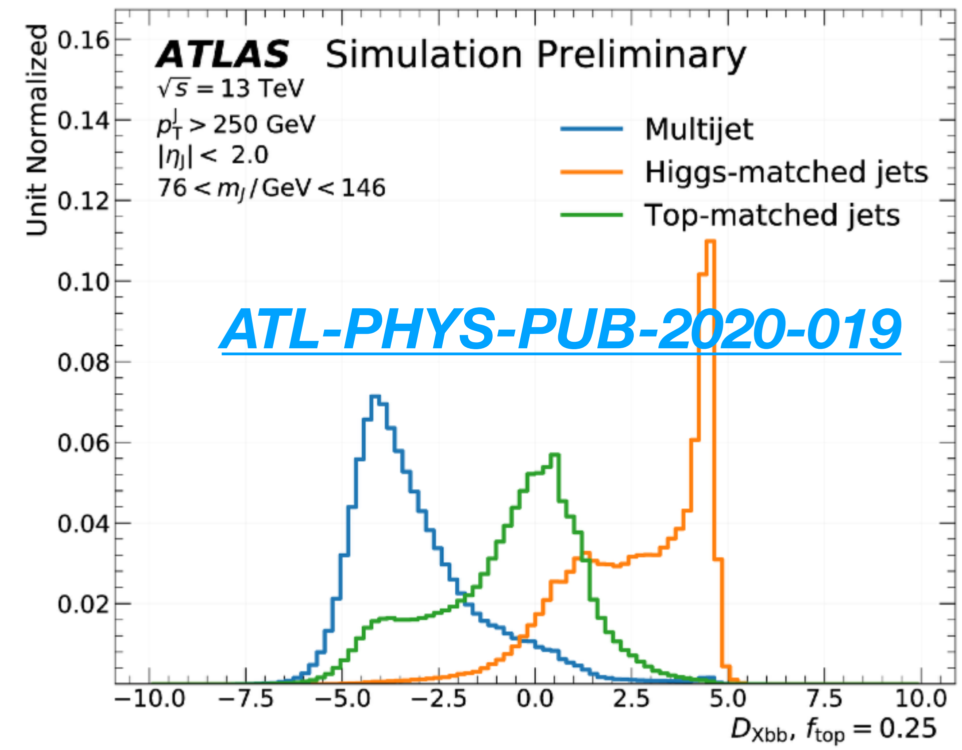
complexity of the tagger →

← ***jet mass de-correlation***

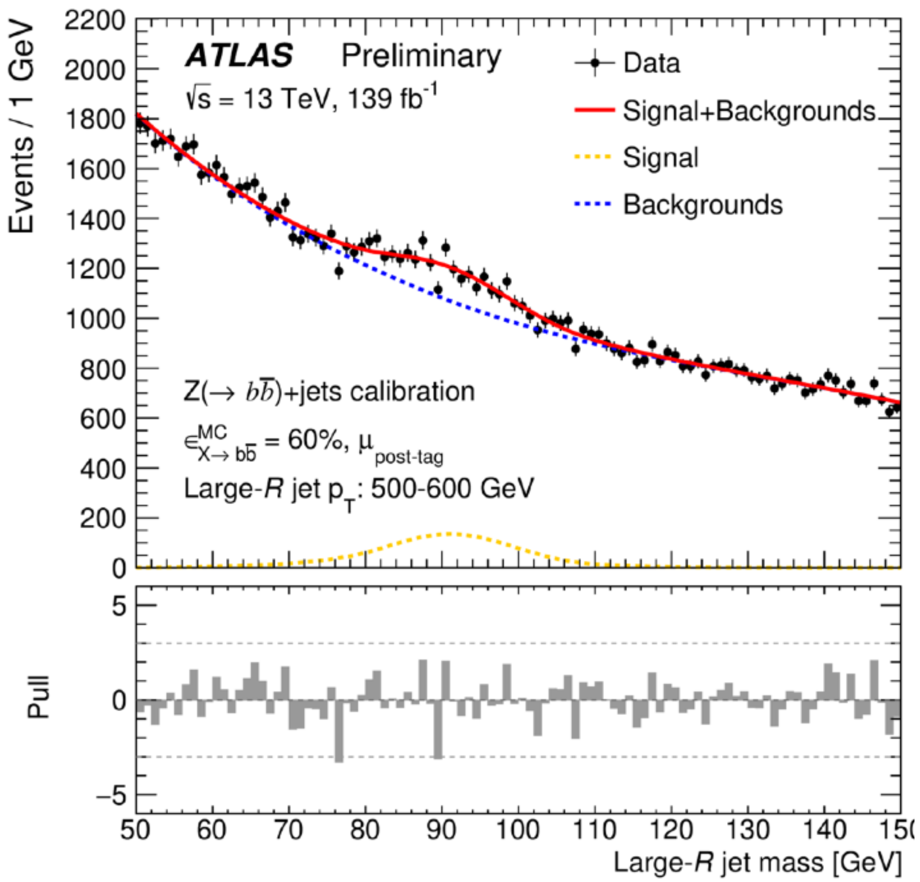
The mass de-correlation is a key features of this task: higher performances might not allow to calibrate the tagger

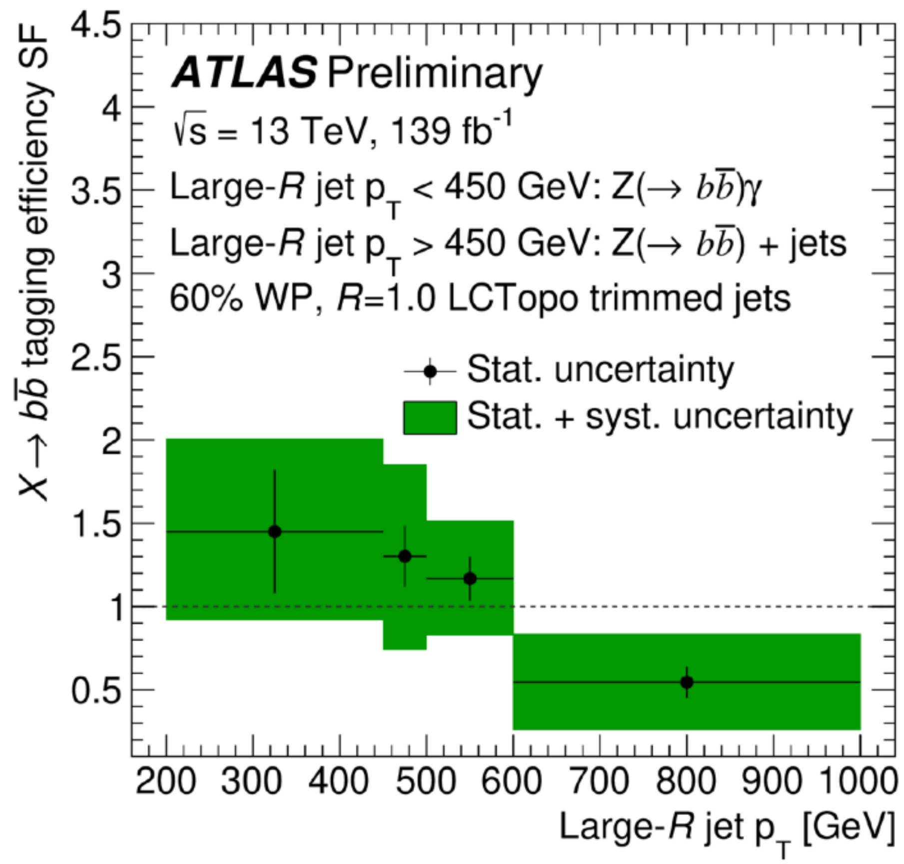
What about the Higgs? Boosted Xbb

- We have very powerful flavour tagging algorithm to identify b-jets (DL1r):
  - *FTAG-2018-01*, *ATL-PHYS-PUB-2020-014*
- Flavor tagging info are plugged into a DNN:
  - up to 3 VariableRadiusTrackJets reconstructed within the large-R jet
  - DL1r information only, to avoid any mass sculpting



ATL-PHYS-PUB-2020-019





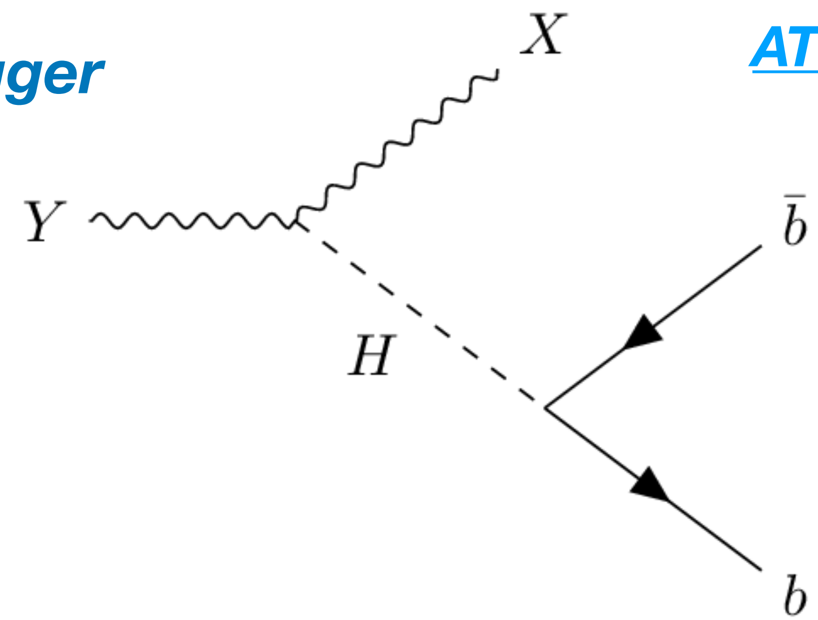
The last bins are affecting more high mass searches ===>> techniques statistics dominated

ATL-PHYS-PUB-2021-035

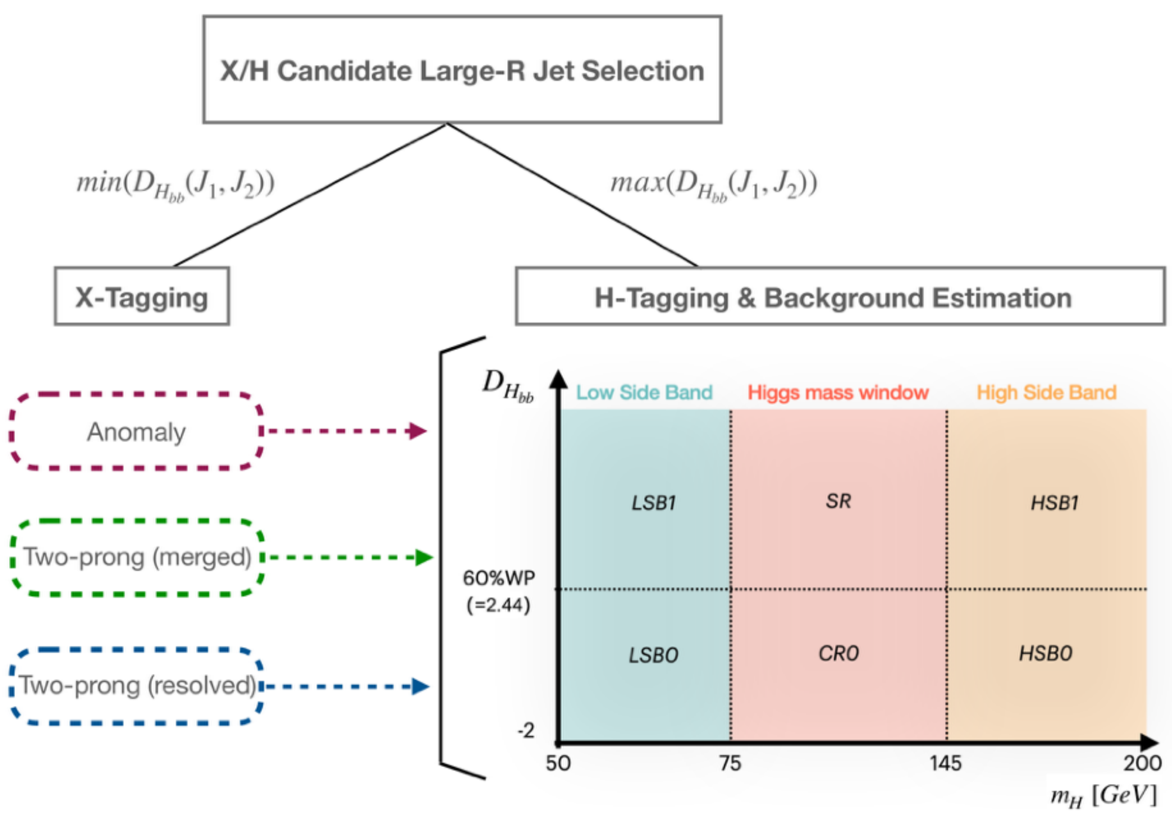
Y —> XH search use case

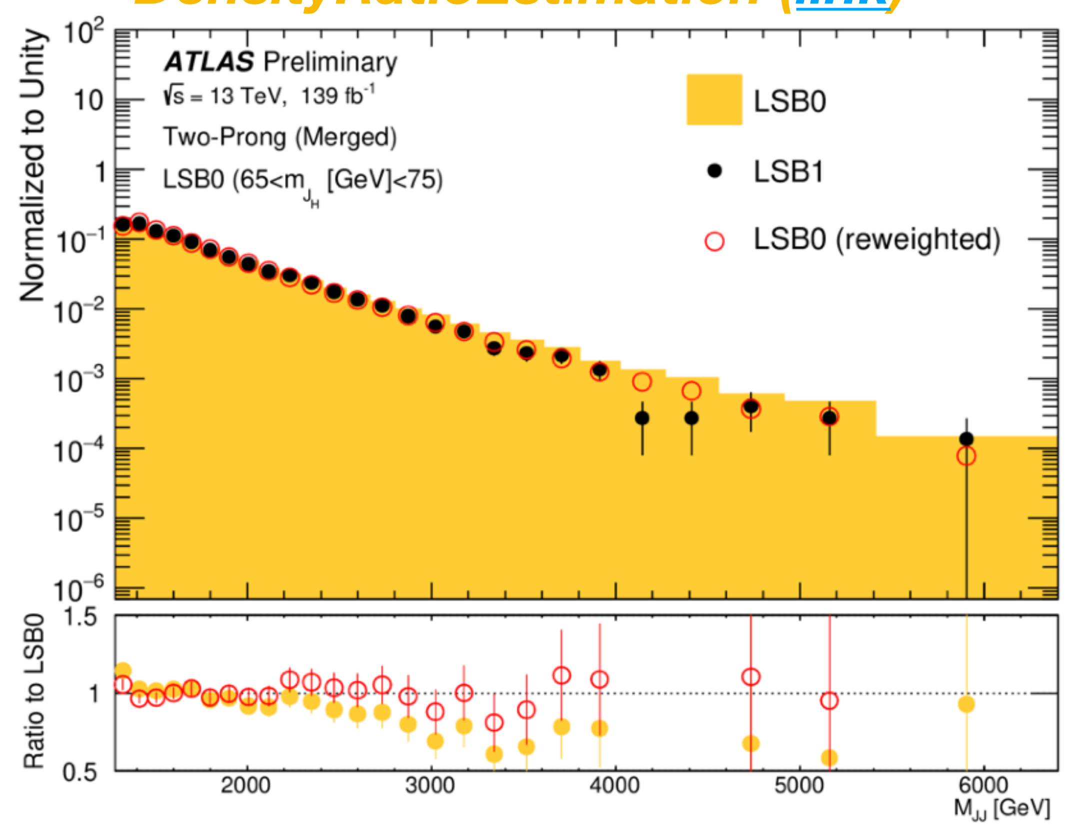
ATLAS-CONF-2022-045

- ***First analysis using the boosted Xbb tagger***
  - Y—> XH high mass search
- ***Interesting final state:***
  - SM Higgs + BSM X boson jet
  - few assumption on the signal model
- ***Challenges:***
  - QCD multijets bkg dominated
  - boosted jet tagging



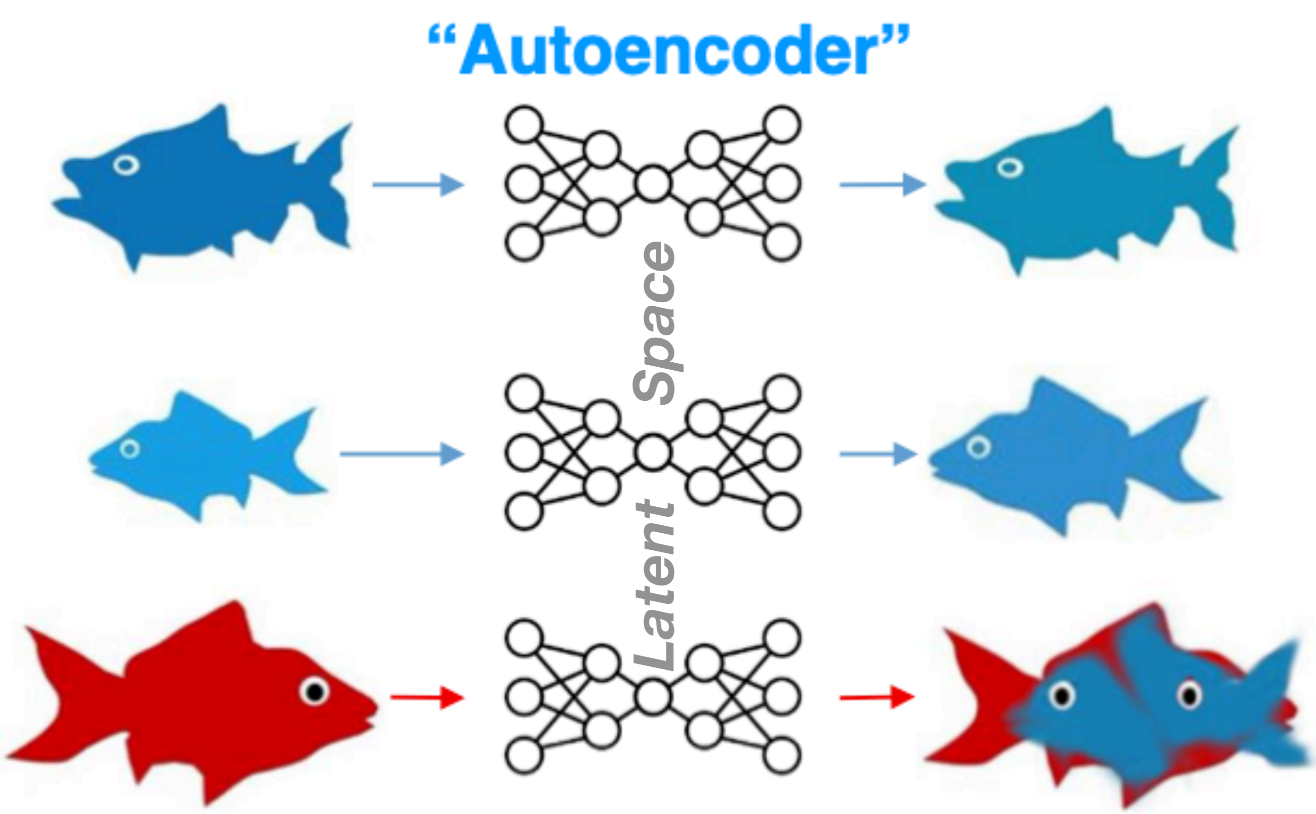
The bkg is estimated via a ML-based data-driven technique using the DensityRatioEstimation (link)

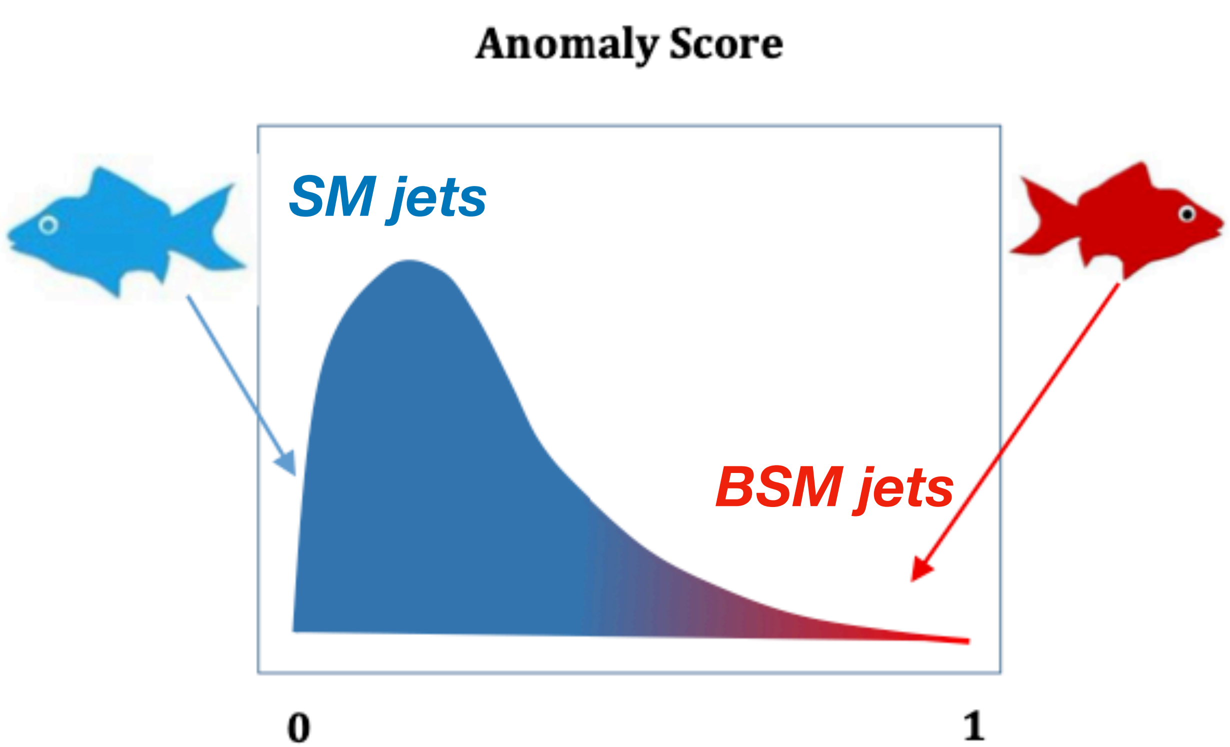




Looking for anomalies in jets

- Many many many BSM searches performed in the last years
  - are we looking for in the right places?
  - are we using the best signal models to optimise our analyses?
- ***Revert the strategy:***
  - model independent searches: minimal assumptions, loose phase spaces
  - train our ML on data to learn directly on real data and try to spot any anomaly

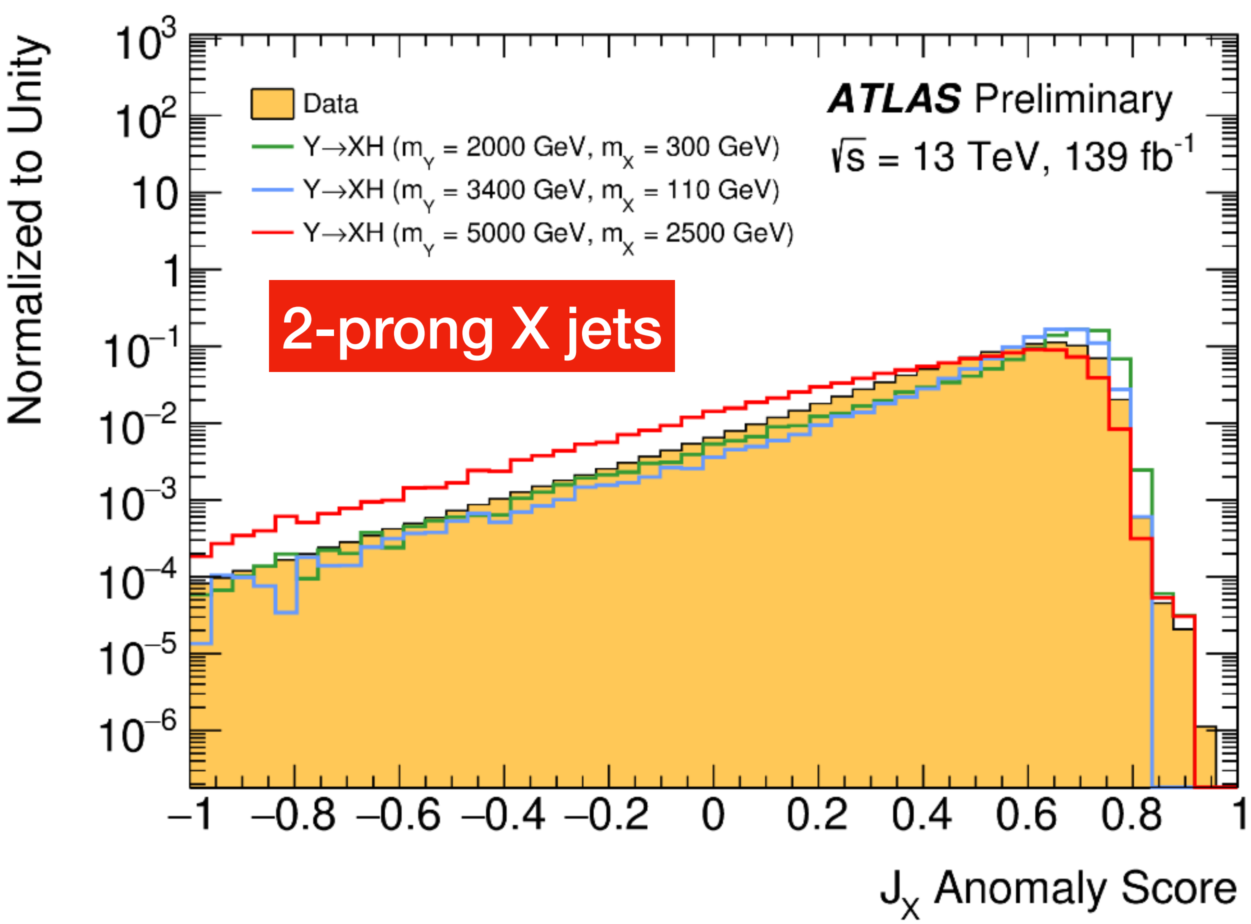


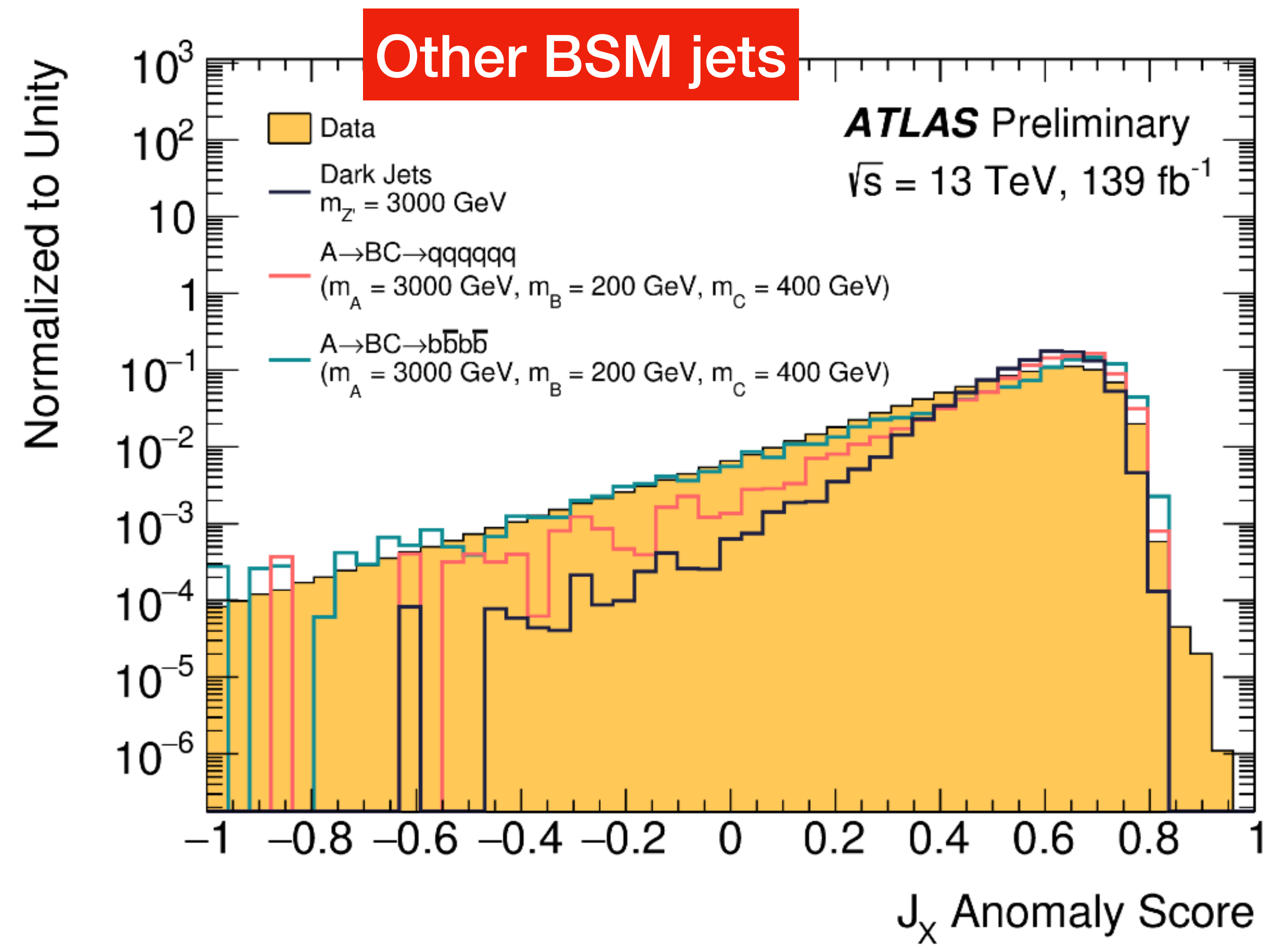


Anomaly detection via jets

ATLAS-CONF-2022-045

- A novel ML technique is used to train on jets:
  - VAE-RNN: Variational Auto-Encoder (VAE) to exploit the latent space and RNN to deal with jets constituents (paper link)
- ***First published results on ATLAS data:***
  - again, YXH has been a milestone effort
  - super interesting challenges to develop new algorithms,
    - LHC Olimpics, Dark Machines

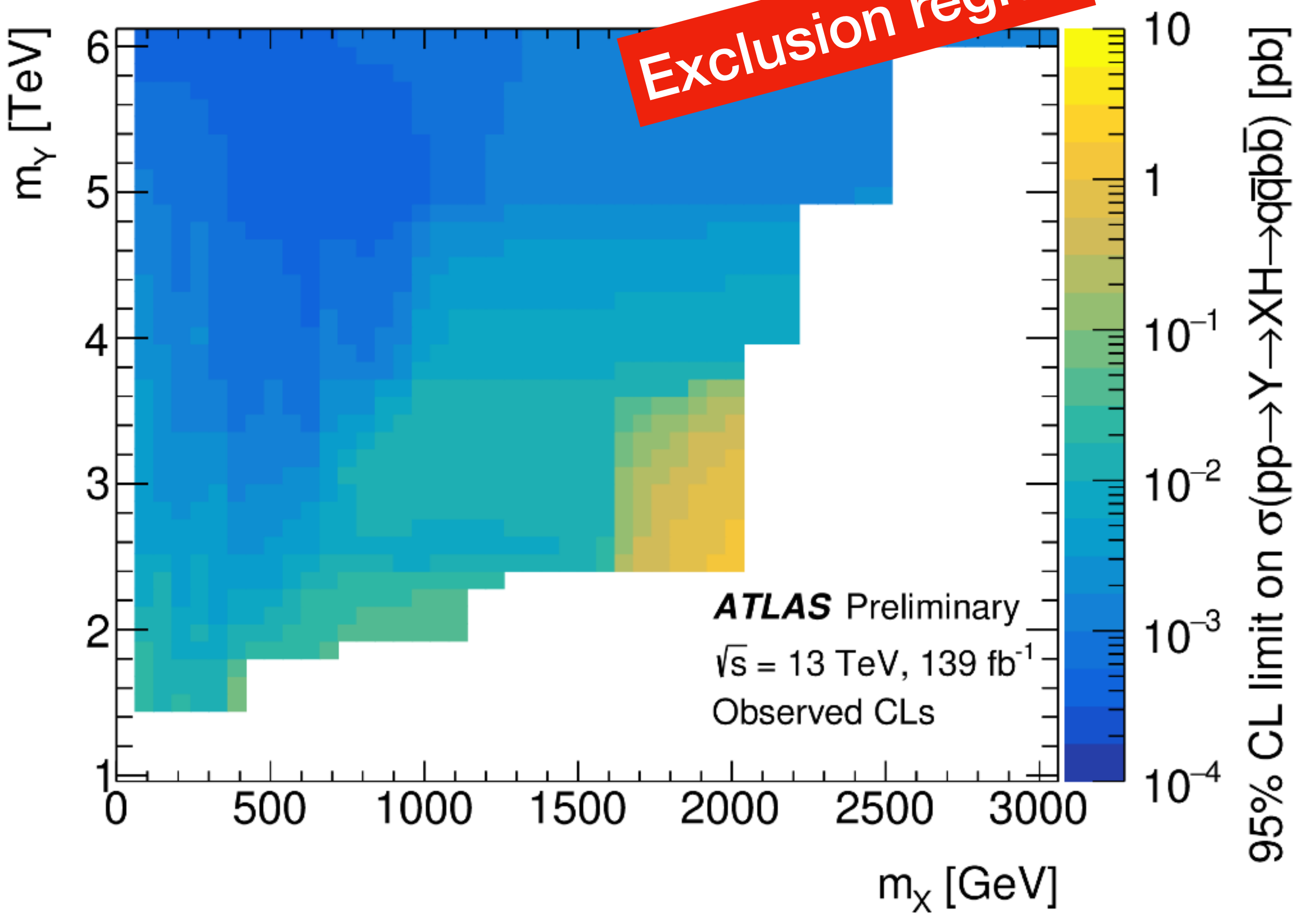


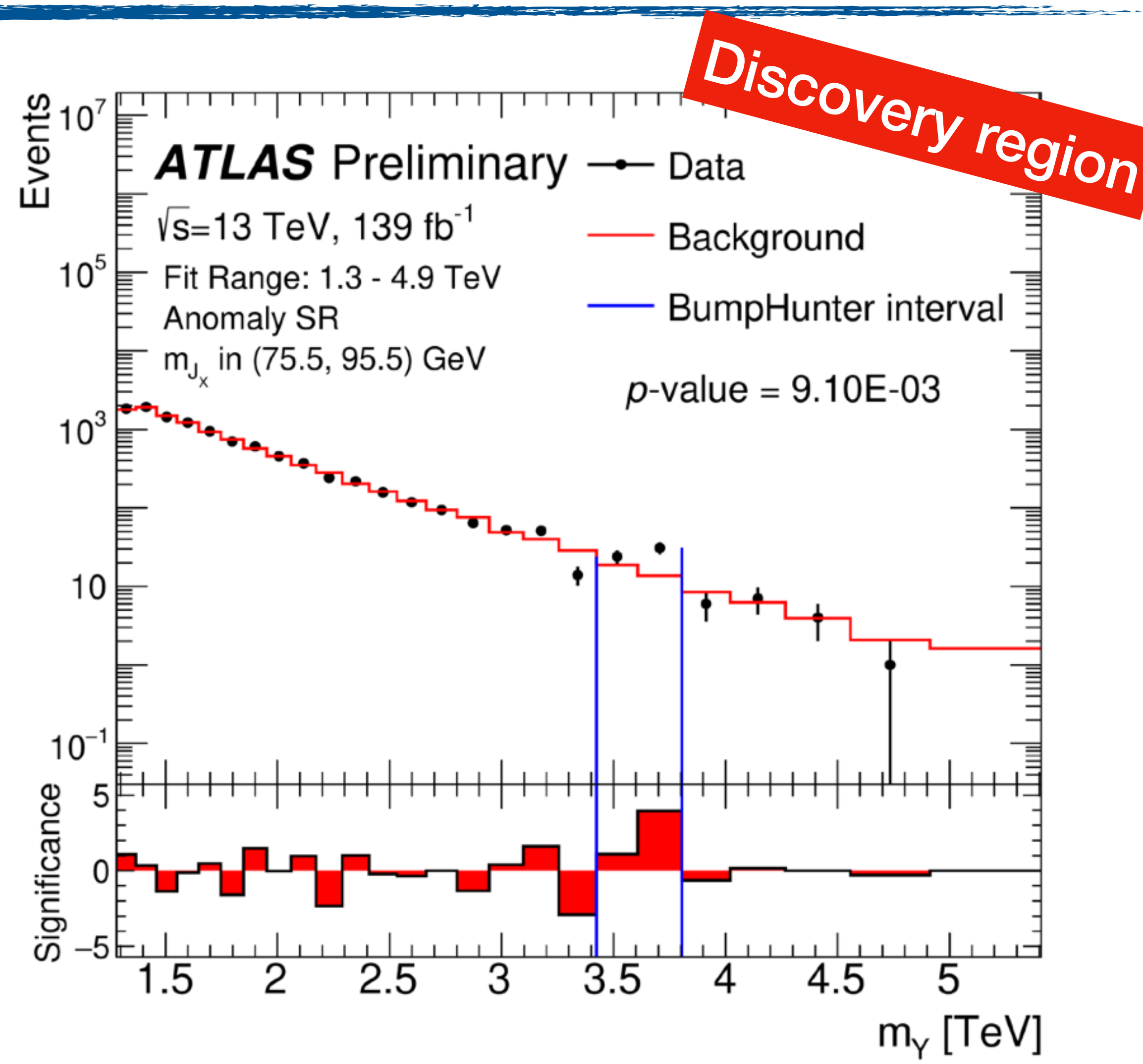


Y —> XH: results

[ATLAS-CONF-2022-045](#)

- 3 SRs are defined according to tagging of the X boson
  - ***resolved regime***, X —> jj (2 prongs)
  - ***boosted regime***, X —> J (2 prongs)
  - ***anomaly regime***, X —> whatever with high Anomaly-Score
- A Bump Hunter search is performed on the smoothly falling invariant mass spectrum of the reconstructed final state





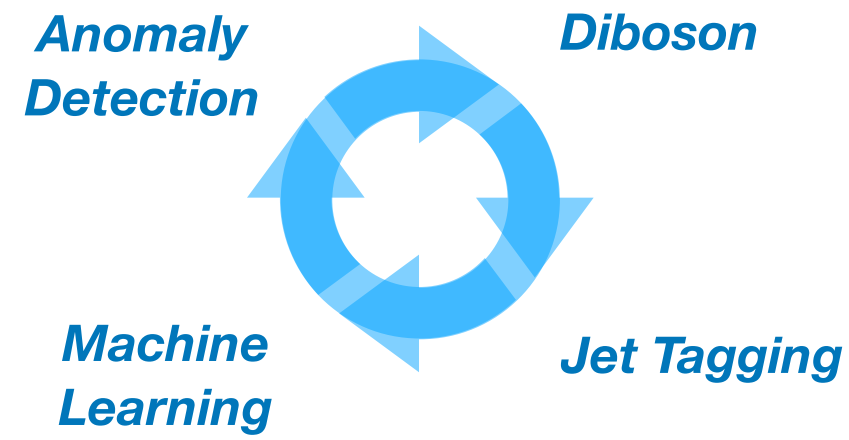
Significant improvement in the exclusion limits in this final states w.r.t. early run-2 search ([Phys. Lett. B 779 (2018) 24](#)), given the new Xbb tagger and other analysis optimisations

The largest excess is found in the anomaly SR with a global significance of 1.47 σ considering all mX and mY bins

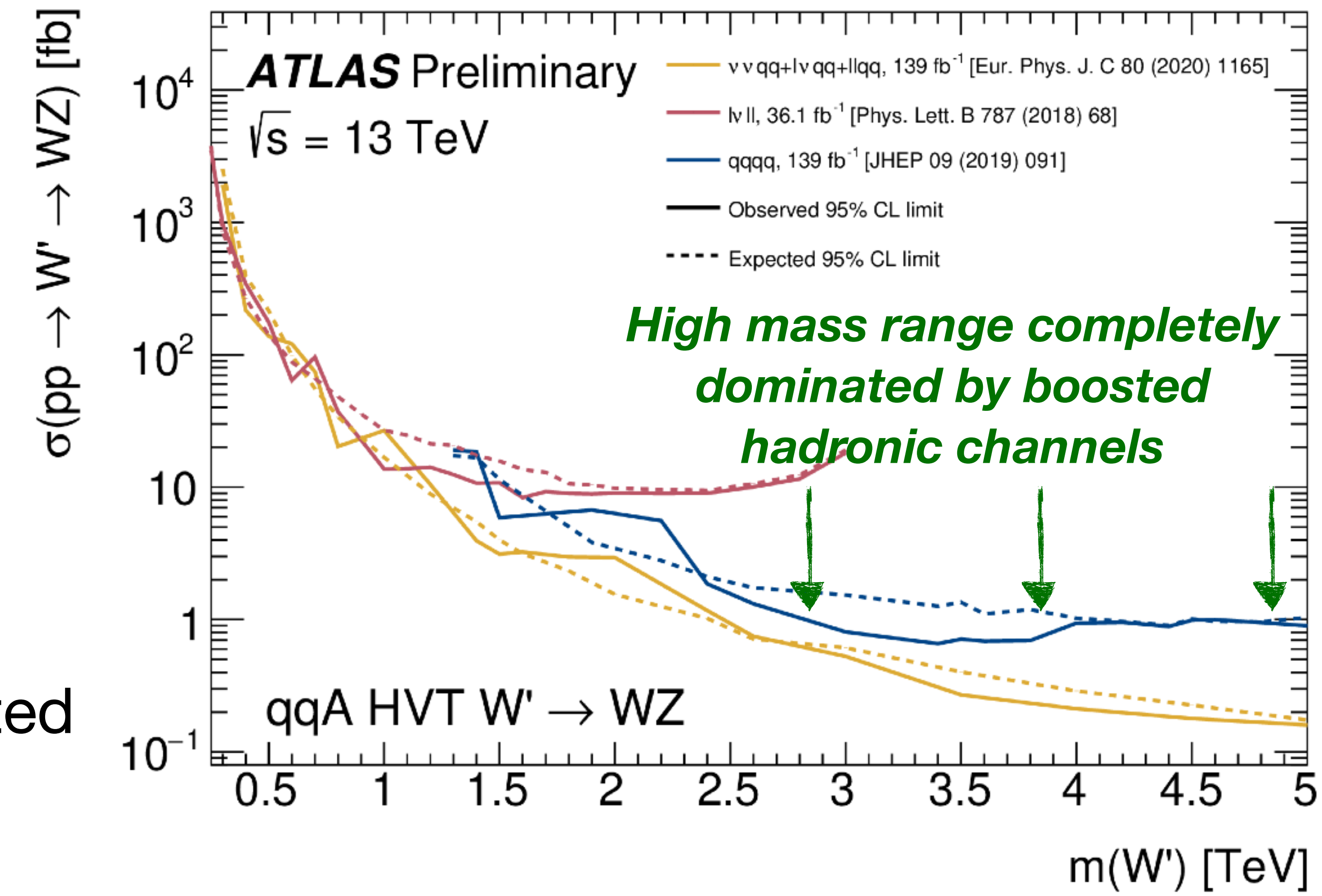
Conclusions and perspectives

New physics searches with boosted objects tagging at the ATLAS experiment

- ***Boosted objects tagging*** are largely exploited in the ATLAS community!
- High performances allow to reach very sensitive analysis and to put new constraints on NP models
  - ***anomaly detection*** approach might be the way for long term research programs
- New techniques and searches expected during ***run-3***, stay tuned!!!



PUB-2021-018



backup

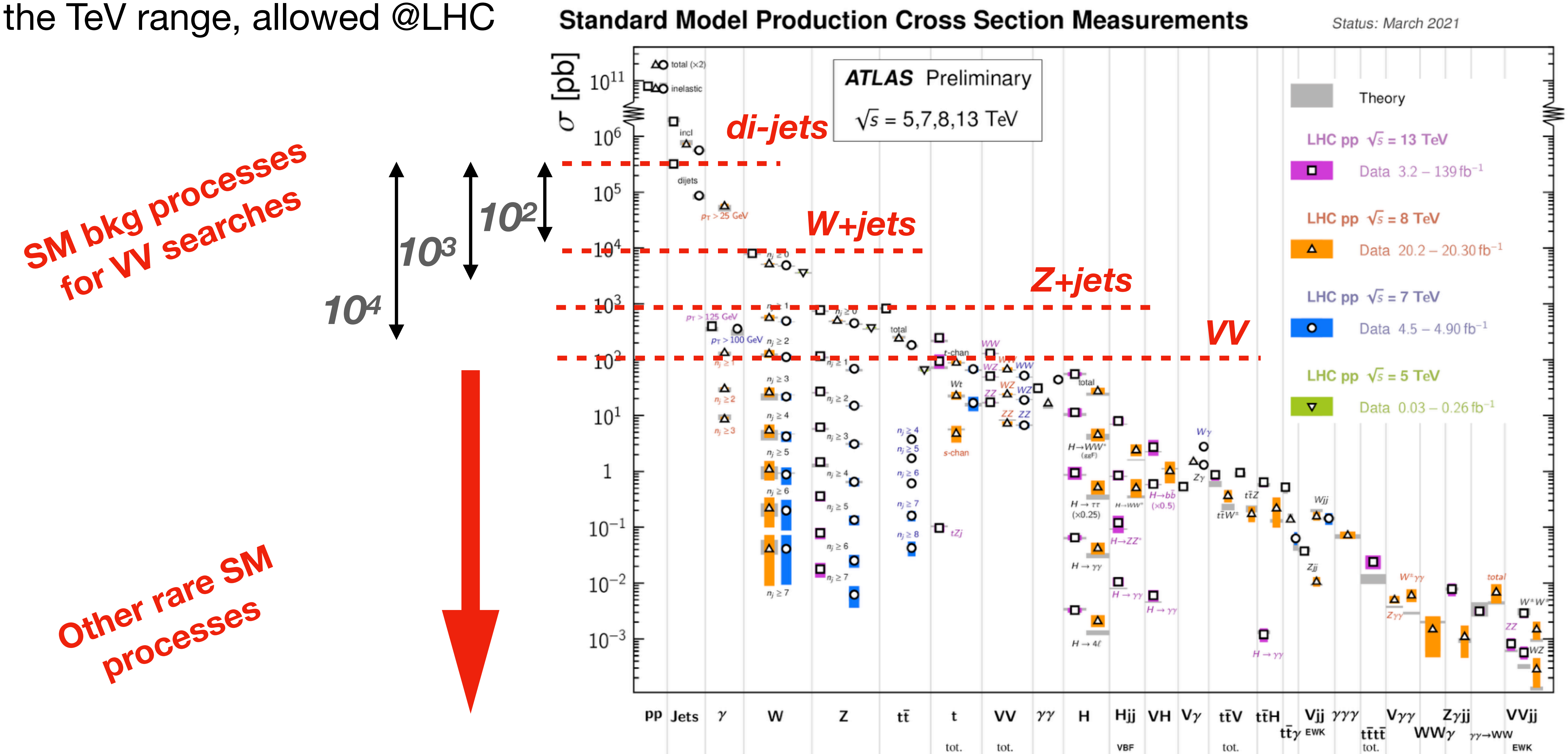
Model			Spin	m = 800 GeV			m = 3 TeV		
				σ [pb]	$\mathcal{B}$	Γ/m	σ [fb]	$\mathcal{B}$	Γ/m
RS radion ($k\pi r_c$ = 35, Λ_R = 3 TeV)		$R \to WW$ $R \to ZZ$	0	0.54 (ggF) 1.1×10^{-3} (VBF)	0.43 0.21	2.6×10^{-3}	1.38 (ggF) 5.5×10^{-3} (VBF)	0.44 0.22	0.032
HVT	Model A	$W' \to WZ$ $Z' \to WW$	1	53 26	0.024 0.023	0.026	79 36	0.020	0.025
	Model B	$W' \to WZ$ $Z' \to WW$		1.6 0.86	0.43 0.41	0.040	5.5 2.5	0.47	0.031
	Model C (VBF)	$W' \to WZ$ $Z' \to WW$		4.0×10^{-3} 2.7×10^{-3}	0.50 0.49	3.5×10^{-3}	1.6×10^{-3} 1.0×10^{-3}	0.50	3.3×10^{-3}
Bulk RS G_{KK} ($k/\overline{M}_{Pl}$ = 1.0)		$G_{KK} \to WW$ $G_{KK} \to ZZ$	2	1.9 (ggF) 0.050 (VBF)	0.28 0.14	0.037	0.47 (ggF) 1.6×10^{-2} (VBF)	0.20 0.10	0.062

HDBS-2018-10

Introduction to diboson

- Diboson signatures open testing scenario of new physics decaying into SM sector
- Several BSM scenarios foresee new massive particles in the TeV range, allowed @LHC

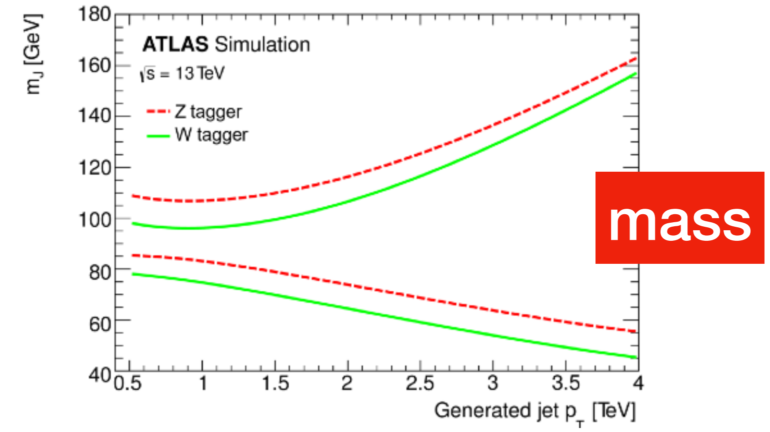
ATL-PHYS-PUB-2021-005

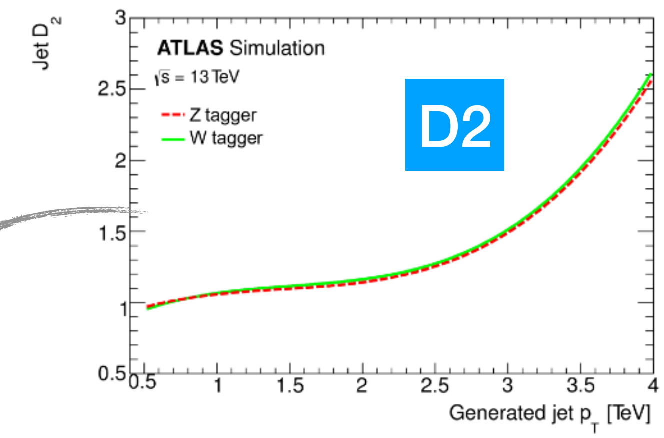


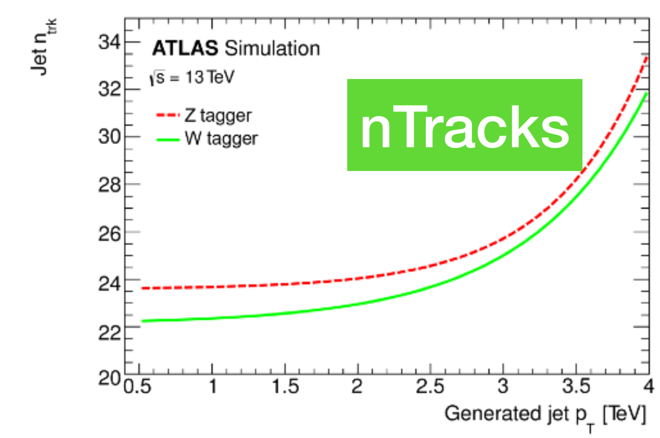
Boson jets and taggers

HDBS-2018-31

- We look for BSM hints in the diboson events
  - the SM VV events are hidden in the ***large amount of the QCD bkg***
- We need powerful ***boson-jet tagger*** to select our signal topology
  - di-jets topology, TCC large-R=1. jets
- 3-variables based tagger
  - WP optimised using significance

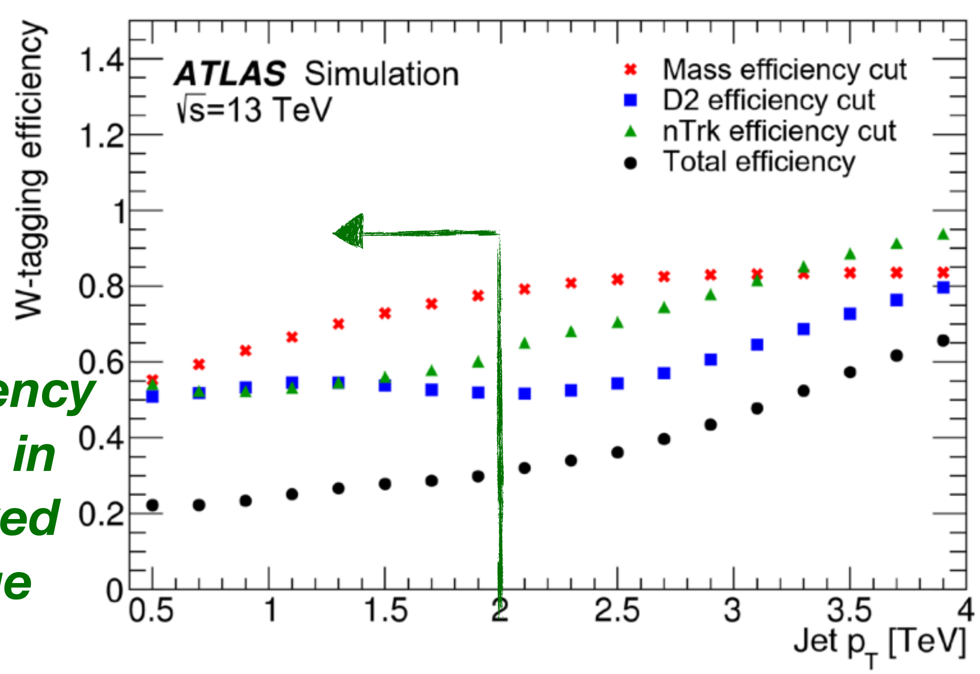




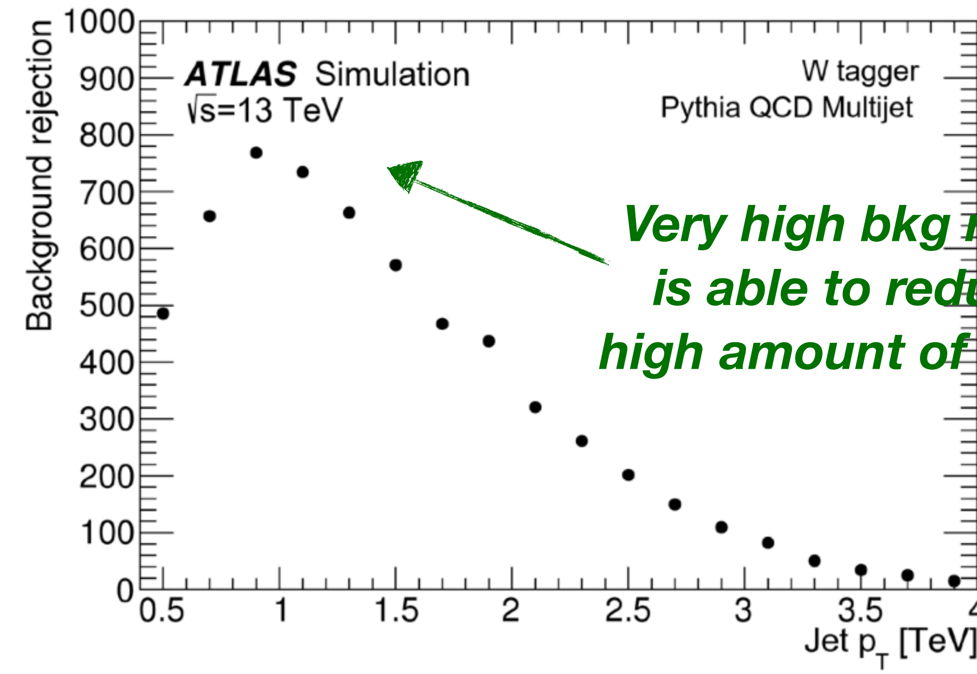


signal efficiency

bkg rejection



signal efficiency mostly flat in the observed data range



Very high bkg rejection is able to reduce the high amount of QCD bkg

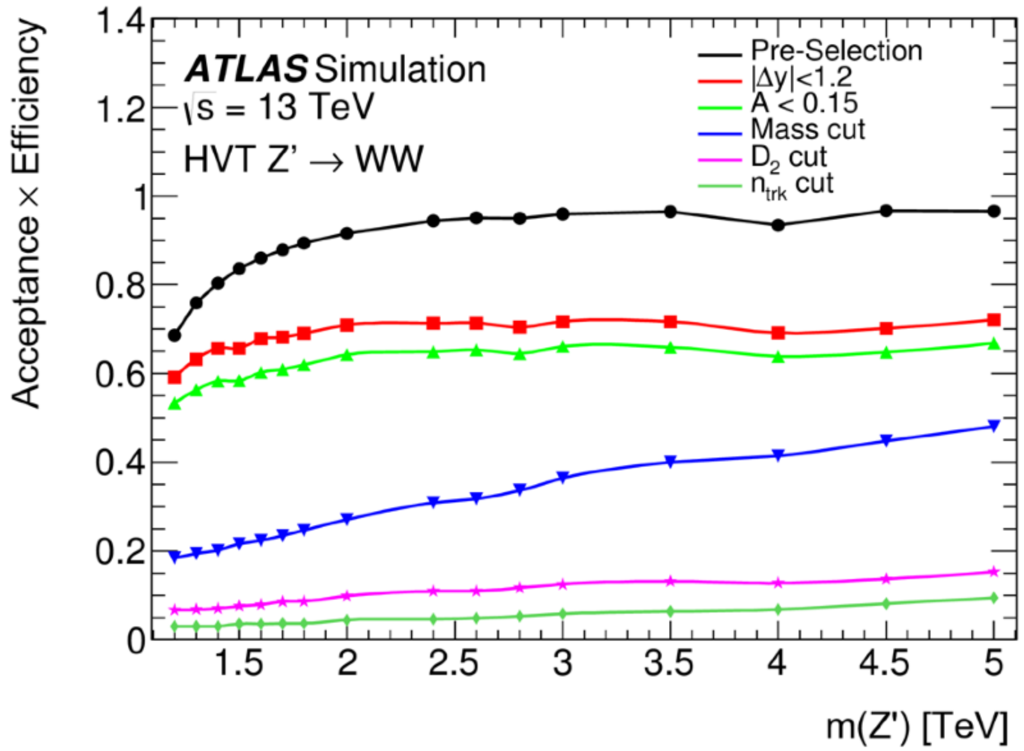
Parameter	Preselection requirements
m_{JJ} [GeV]	> 1300
$p_{\mathrm{T}}(J_1)$ [GeV]	> 500
m_J [GeV]	$m_{J_1} > 50 \parallel m_{J_2} > 50$
$D_{H_{bb}}$	> -2

	Signal regions		
	Merged	Resolved	Anomaly
m_H [GeV]	(75, 145)		
$D_{H_{bb}}$	> 2.44		
D_2^{trk}	< 1.2	> 1.2	-
$\lvert\Delta y_{j_1,j_2}\rvert$	-	< 2.5	-
p_{T}^{bal}	-	< 0.8	-
Anomaly Score	-	-	> 0.5

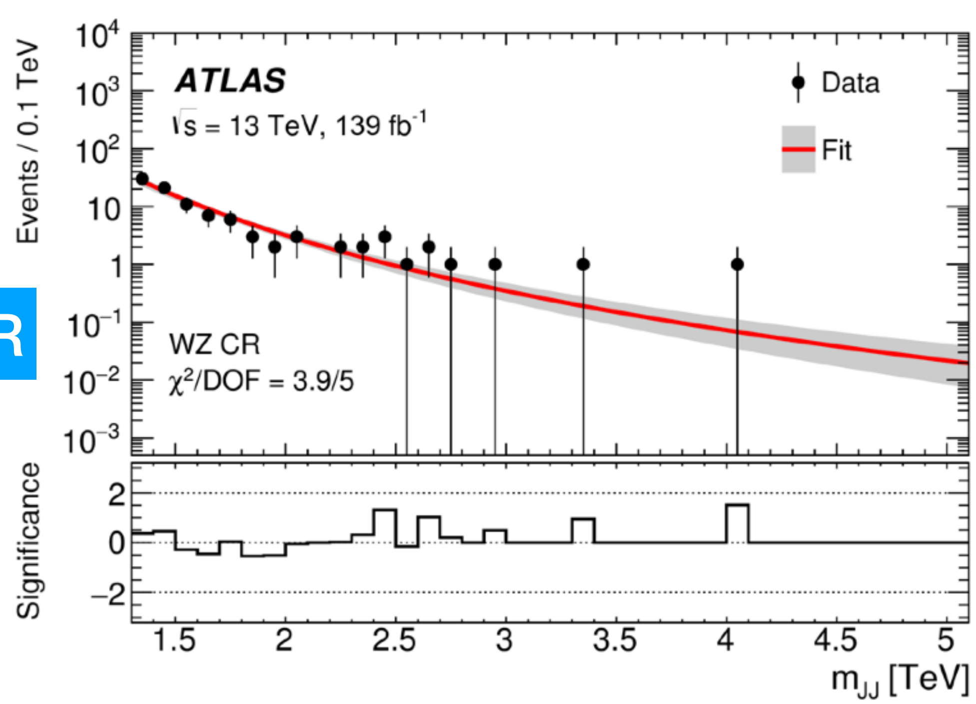
	Background estimation regions				
	CR0	HSB0	HSB1	LSB0	LSB1
m_H [GeV]	(75, 145)	(145, 200)		(65, 75)	
$D_{H_{bb}}$	< 2.44	< 2.44	> 2.44	< 2.44	> 2.44

VV fully hadronic use case

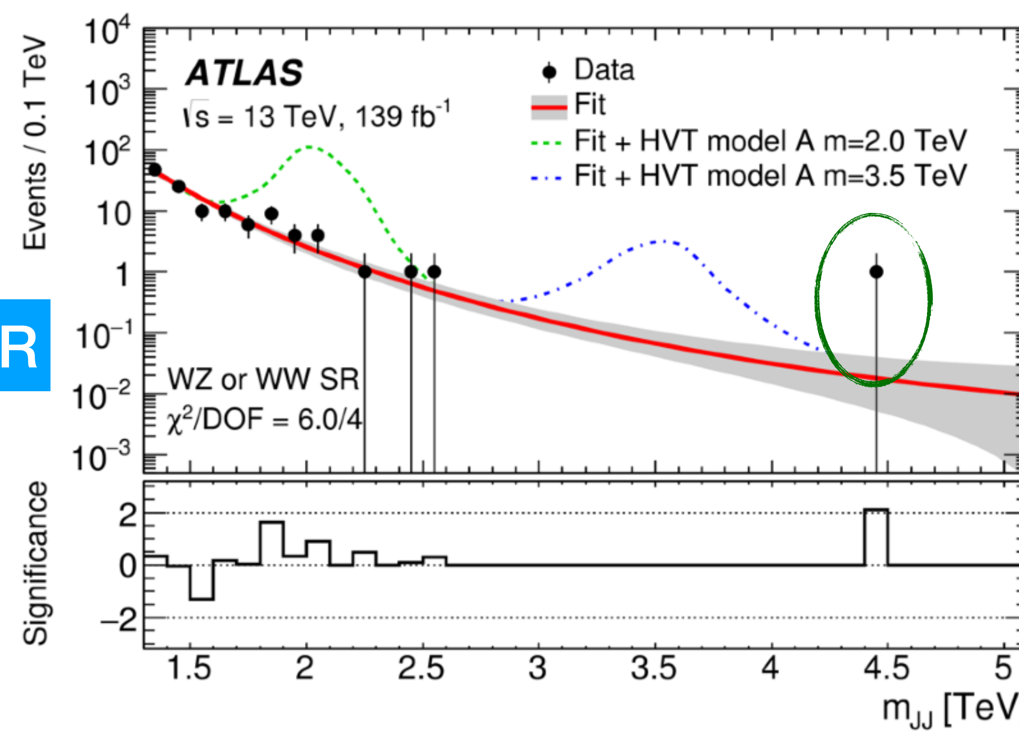
HDBS-2018-31



CR



SR

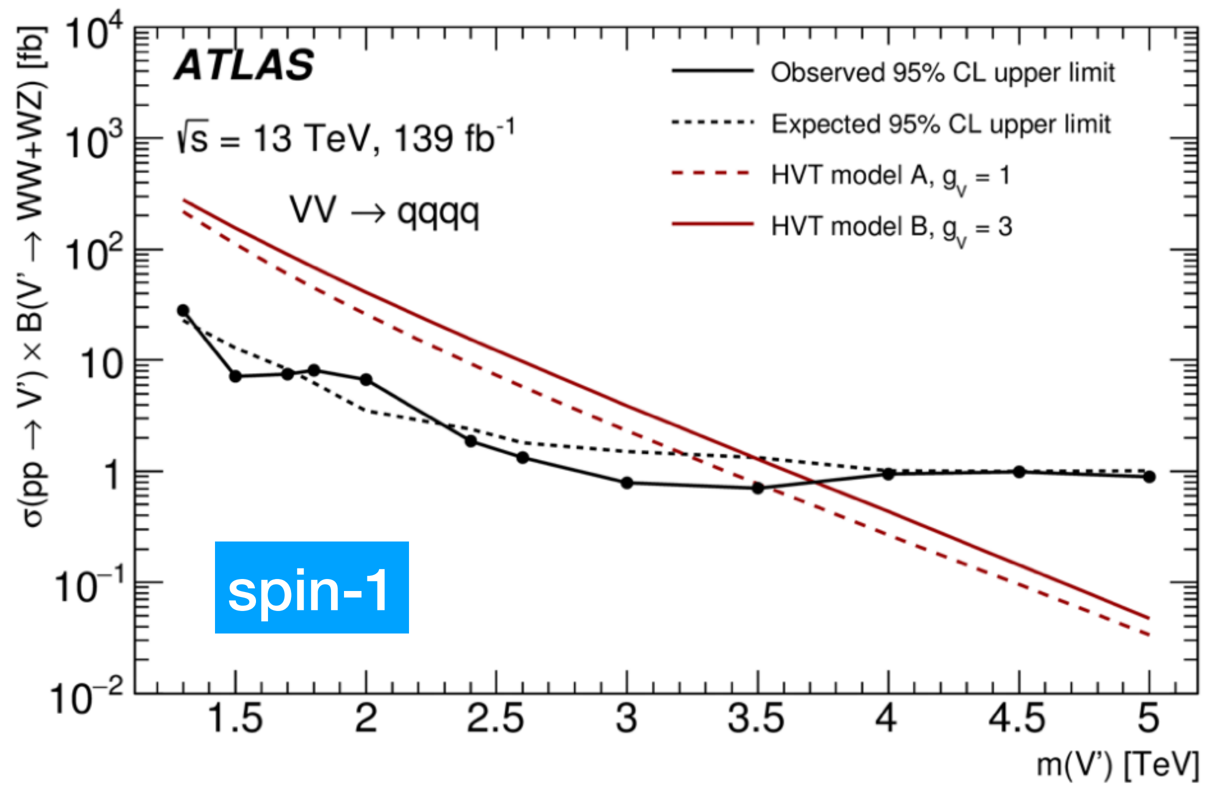


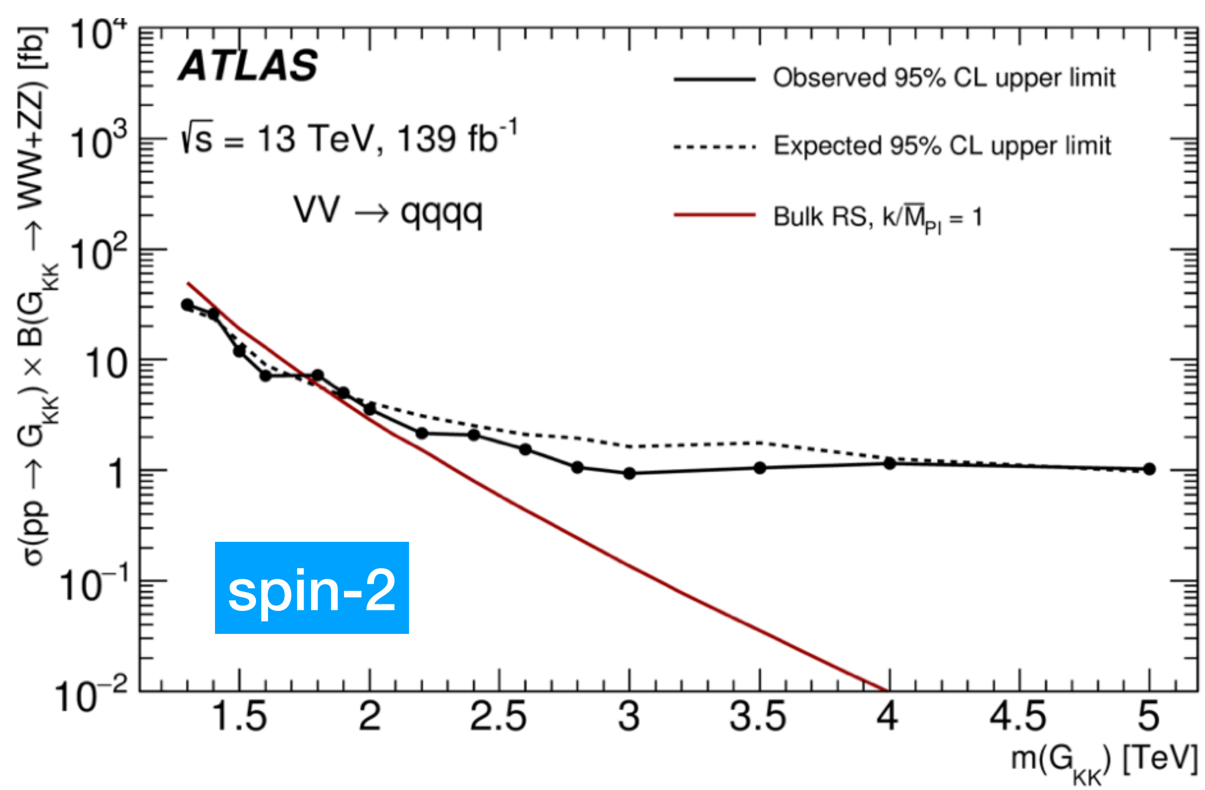
We collect data at very high m_{JJ} range

VVJJ: results

HDBS-2018-31

Model	Signal Region	Excluded mass range [TeV]
Radion	WW	none
	ZZ	none
	$WW+ZZ$	none
HVT model A, $g_V = 1$	WW	1.3–2.9
	WZ	1.3–3.4
	$WW+WZ$	1.3–3.5
HVT model B, $g_V = 3$	WW	1.3–3.1
	WZ	1.3–3.6
	$WW+WZ$	1.3–3.8
Bulk RS, $k/\overline{M}_{\rm Pl} = 1$	WW	1.3–1.6
	ZZ	none
	$WW+ZZ$	1.3–1.8





W/Z boosted tagging

ATL-PHYS-PUB-2021-029