Looking for new physics with high mass diboson searches

CLHCP workshop

26th November 2021

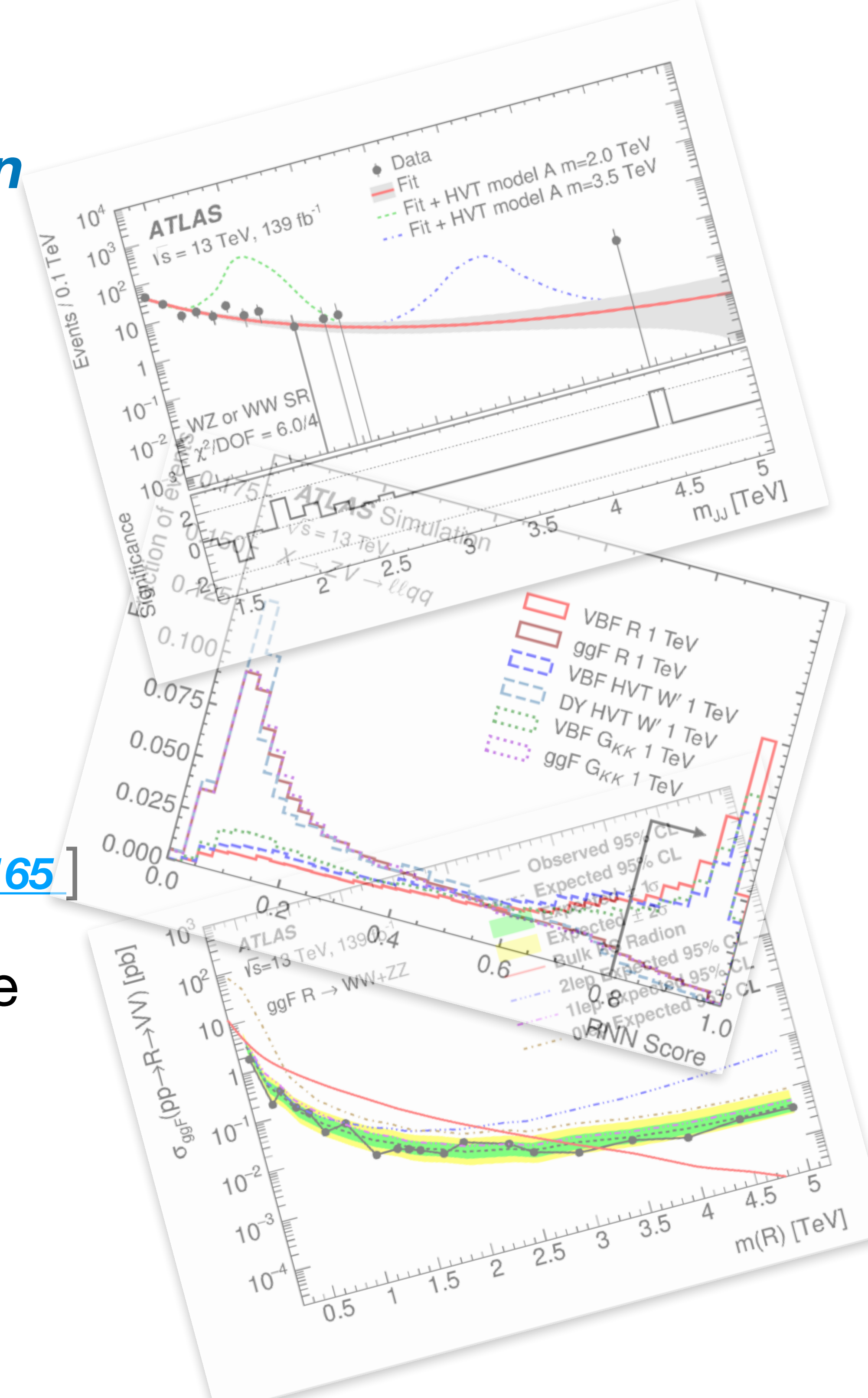
Antonio Giannini

University of Science and Technology of China

Outline

Looking for new physics with high mass diboson searches

- Introduction to VV searches
- Full Run-2 139 fb^{-1} data results
- Fully hadronic final states: [*JHEP09(2019)091*]
  - Boson-jet tagger
- Semi-leptonic final states: [*Eur. Phys. J. C 80 (2020) 1165*]
  - VBF event topology identification (Machine Learning)
- Conclusions and perspectives

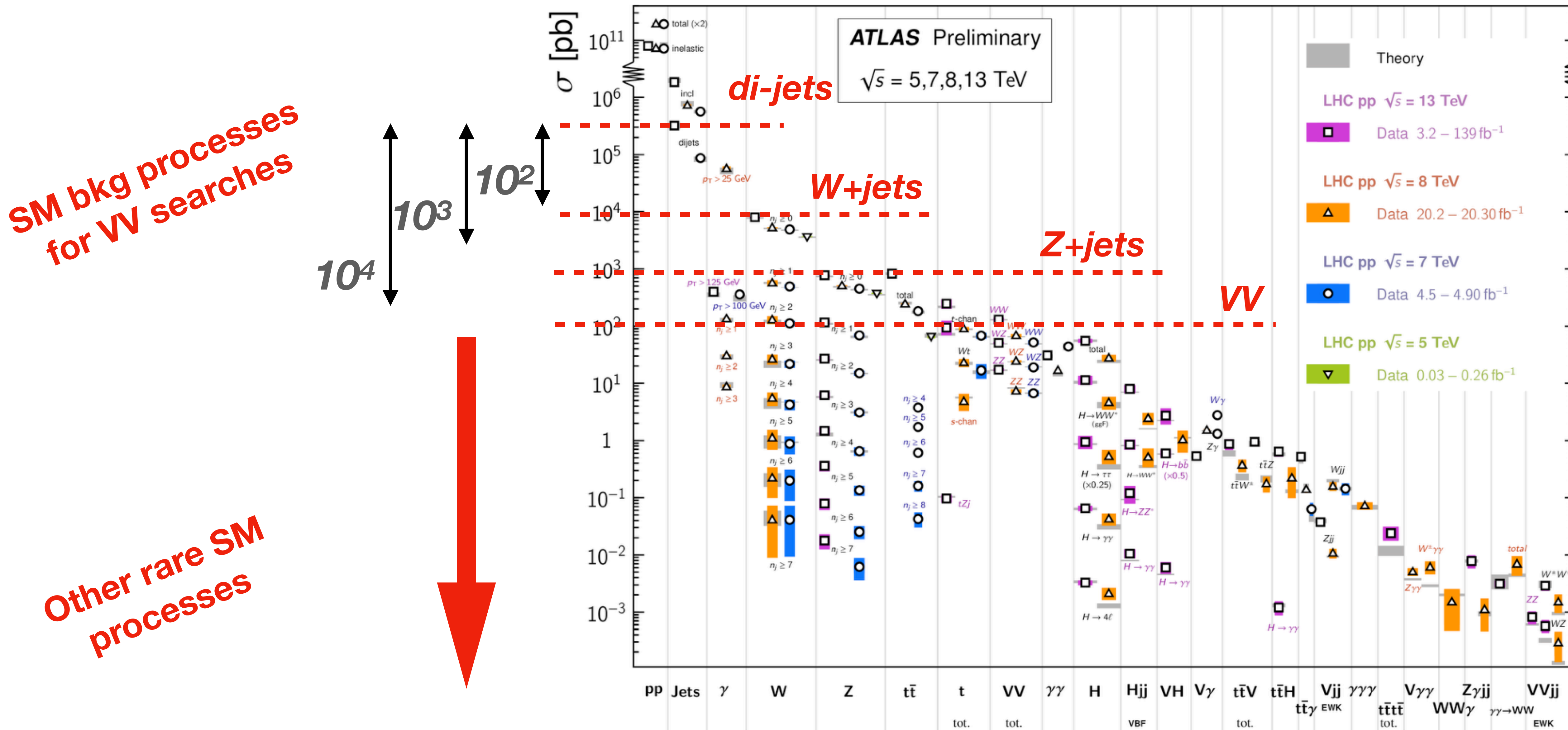


Introduction to diboson

- 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

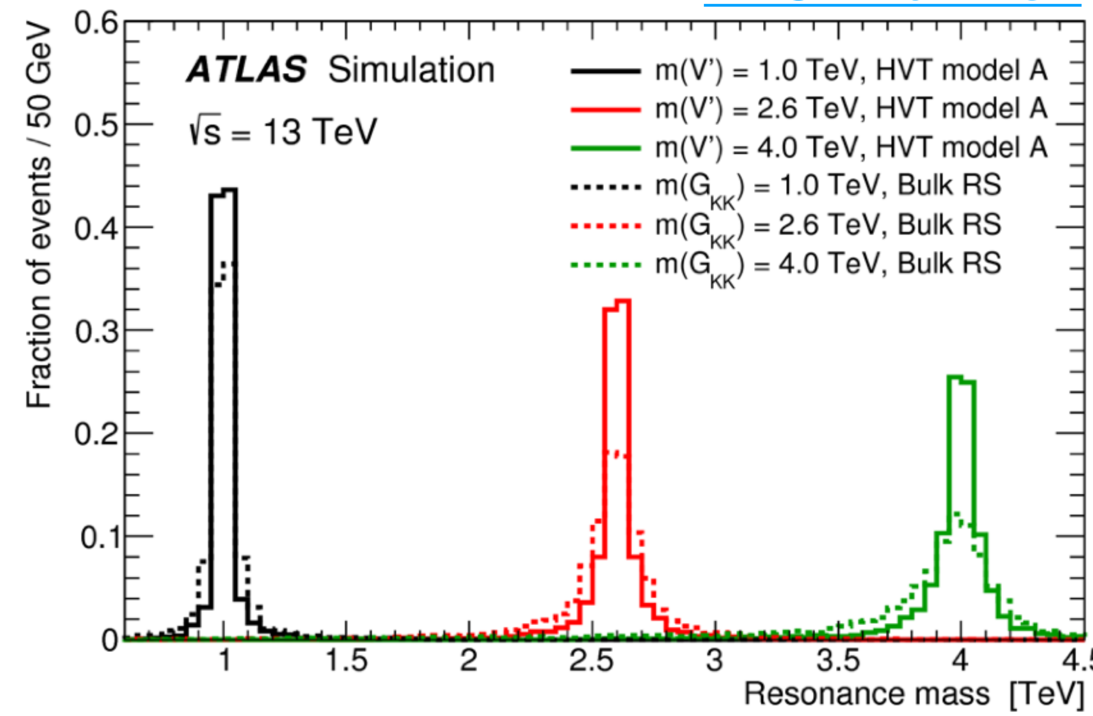
Standard Model Production Cross Section Measurements Status: March 2021



BSM models: heavy VV resonances

- Benchmark BSM models
  - Spin-0: RS Radions
  - Spin-1: Heavy Vector Triplet (HVT)
  - Spin-2: RS Gravitons
- Different production mechanisms are considered, ggF/DY/VBF
- Vector bosons from intermediate resonances are longitudinally state dominated

EXOT-2017-31

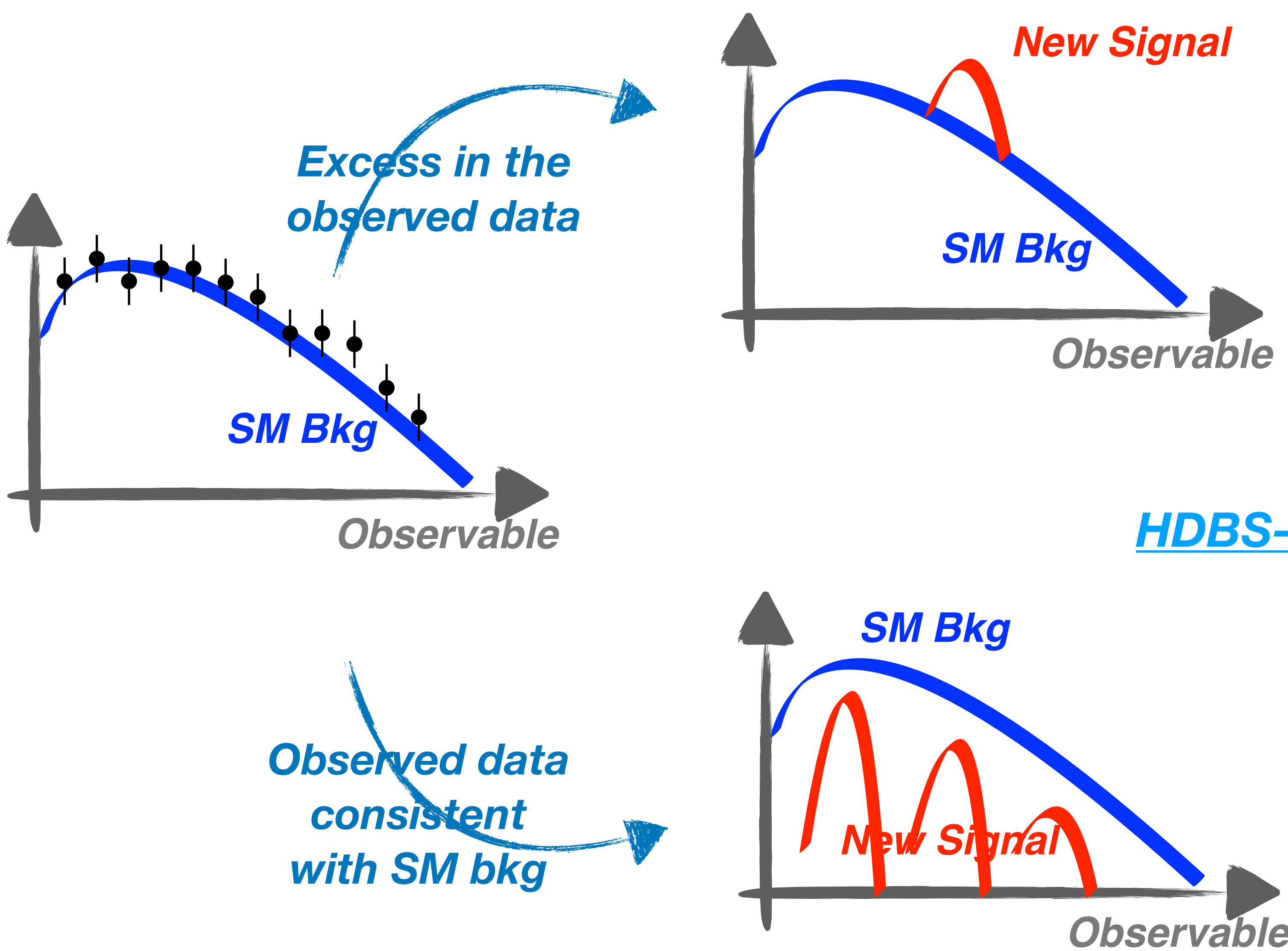


HDBS-2018-10

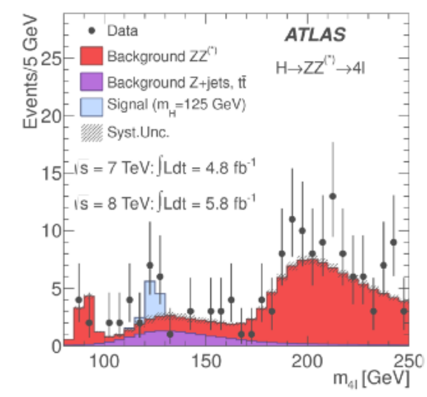
Model			Spin	m = 800 GeV			m = 3 TeV		
				σ [pb]	$\mathcal{B}$	Γ/m	σ [fb]	$\mathcal{B}$	Γ/m
RS radion ($k\pi r_c = 35$, $\Lambda_R = 3$ TeV)		$R \to WW$	0	0.54 (ggF)	0.43	2.6×10^{-3}	1.38 (ggF)	0.44	0.032
		$R \to ZZ$		1.1×10^{-3} (VBF)	0.21		5.5×10^{-3} (VBF)	0.22	
HVT	Model A	$W' \to WZ$	1	53	0.024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

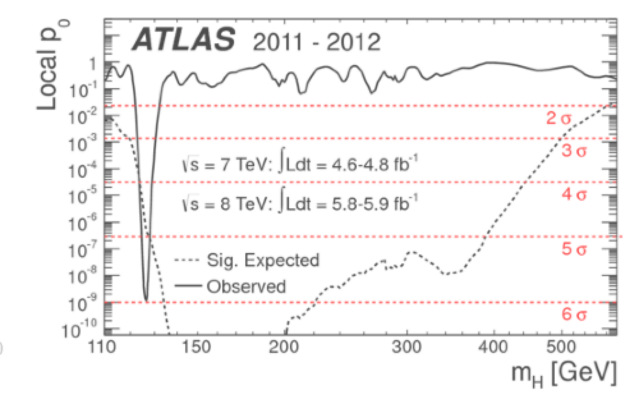
Analysis strategy

- Set of signal regions, SRs, to optimise the signal sensitivity
- Set of control regions, CRs, to constrain the SM bkg:
  - high purity regions to constraint the MC
  - data-driven techniques
- Unblind your SRs, what do you observe?

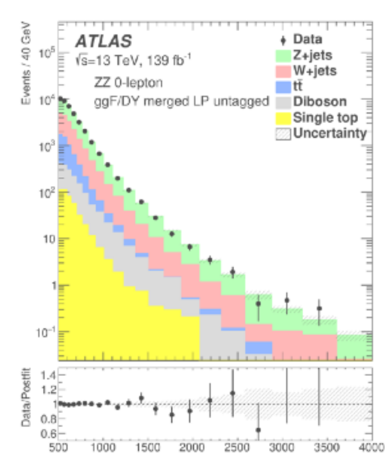


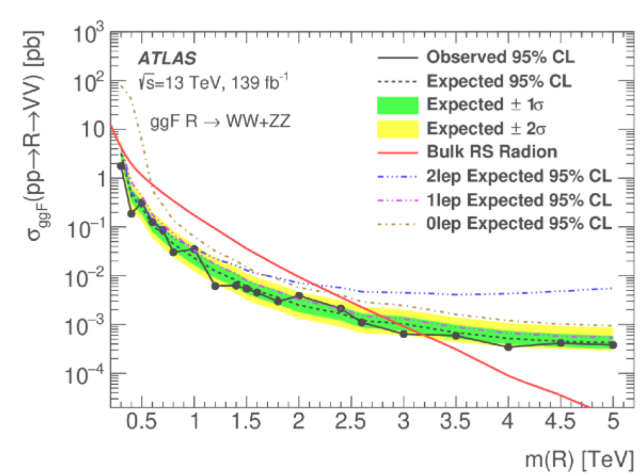
[link](#) **Claim your discovery!**



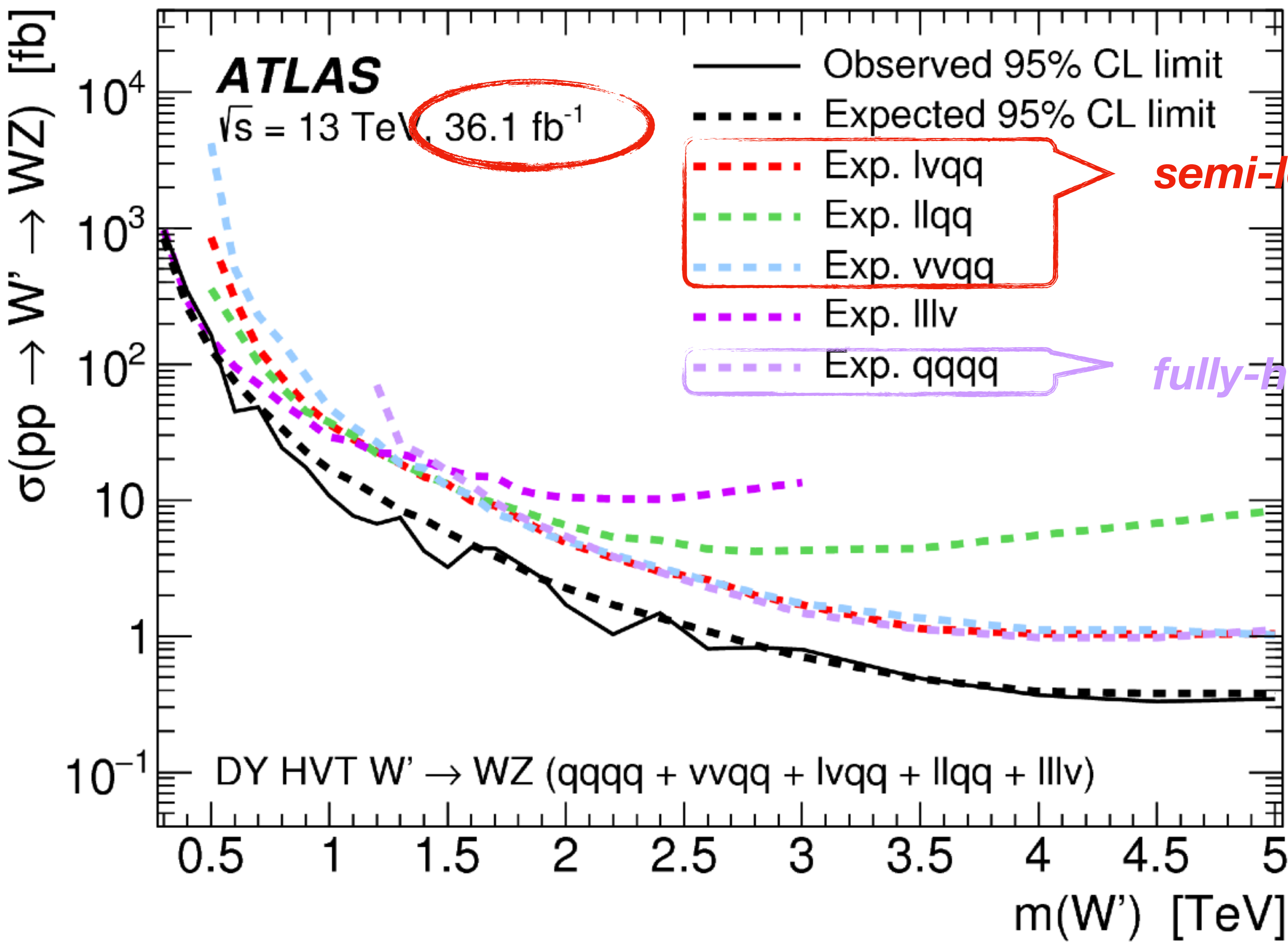


[HDBS-2018-10](#) **Put bounds on BSM models**





Diboson signatures sensitivity



Where we were with early run-2 data?

- Low mass: llqq/lllv
- High mass: vvqq/lvqq/qqqq

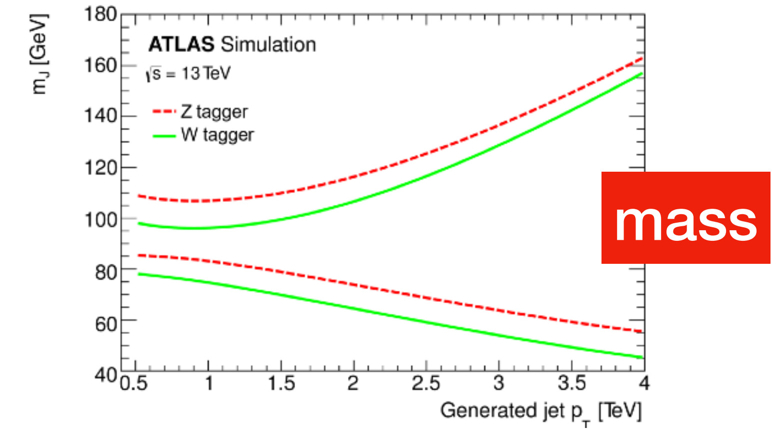
Channel	Lower limits on resonance mass [TeV]					
	HVT model A		HVT model B		Bulk RS	
	Obs	Exp	Obs	Exp	Obs	Exp
WW	2.9	3.1	3.6	3.5	1.7	1.9
WZ	3.6	3.6	3.9	3.9	-	-
ZZ	-	-	-	-	1.5	1.7
VV	3.7	3.7	4.0	3.9	2.3	2.2
WH	2.6	2.8	2.8	3.1	-	-
ZH	2.7	2.5	2.8	2.8	-	-
VH	2.8	3.1	3.0	3.4	-	-
$\ell\nu$	4.6	4.6	-	-	-	-
$\ell\ell$	4.5	4.4	-	-	-	-
$\ell\nu/\ell\ell$	5.0	5.0	-	-	-	-
VV/VH	4.3	4.3	4.5	4.4	-	-
$VV/VH/\ell\nu/\ell\ell$	5.5	5.3	-	-	-	-

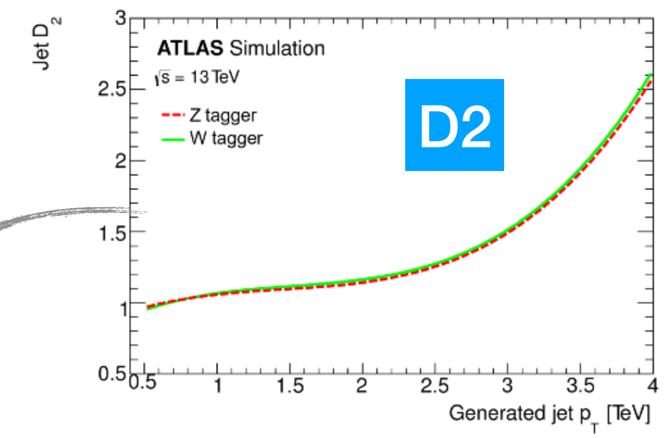
EXOT-2017-31

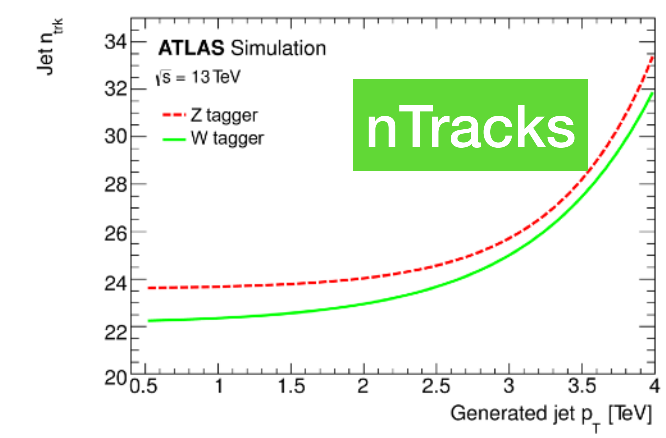
Boson jets and taggers

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



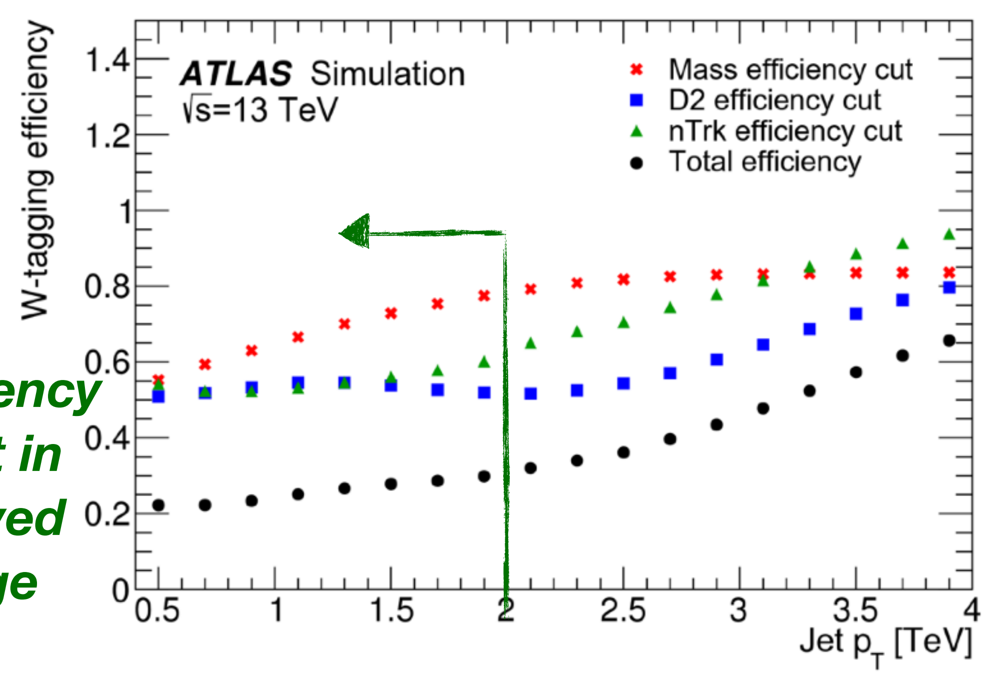


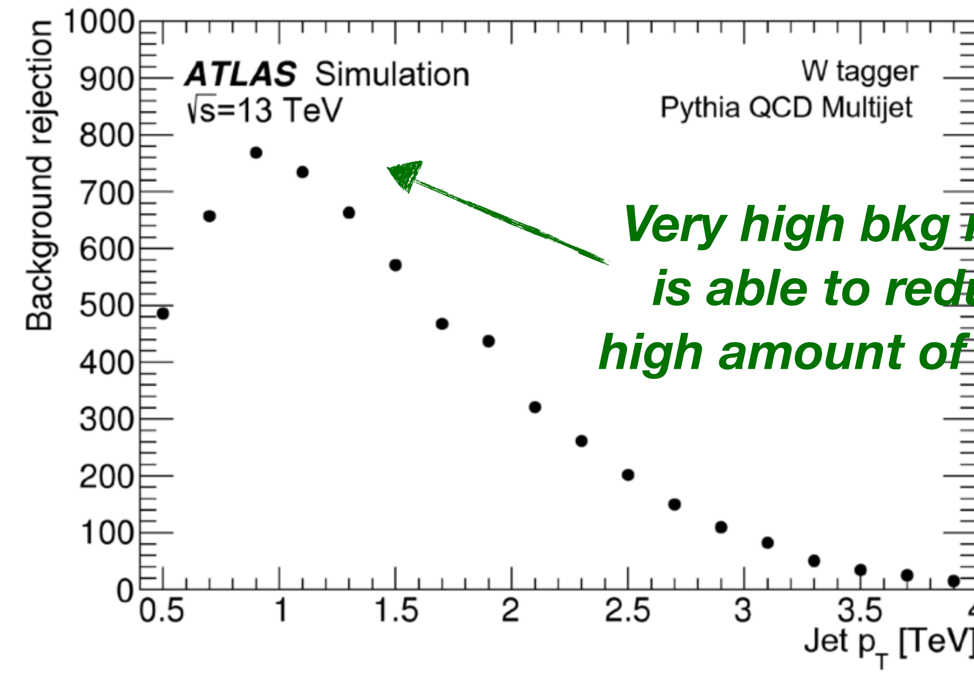


signal efficiency

bkg rejection

signal efficiency mostly flat in the observed data range

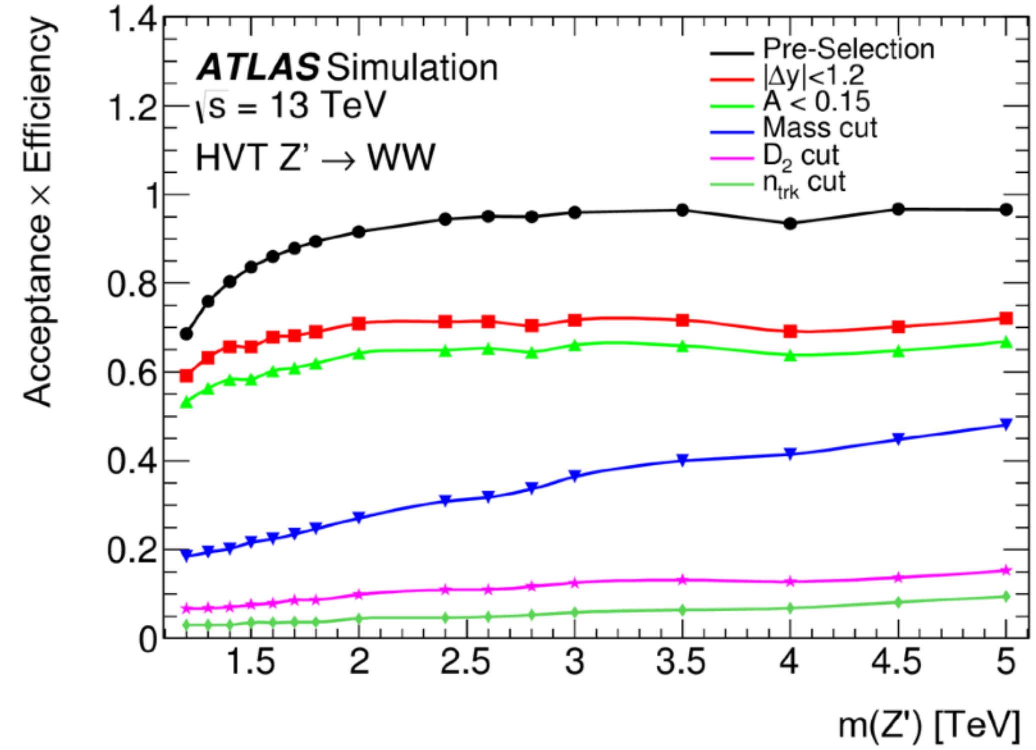




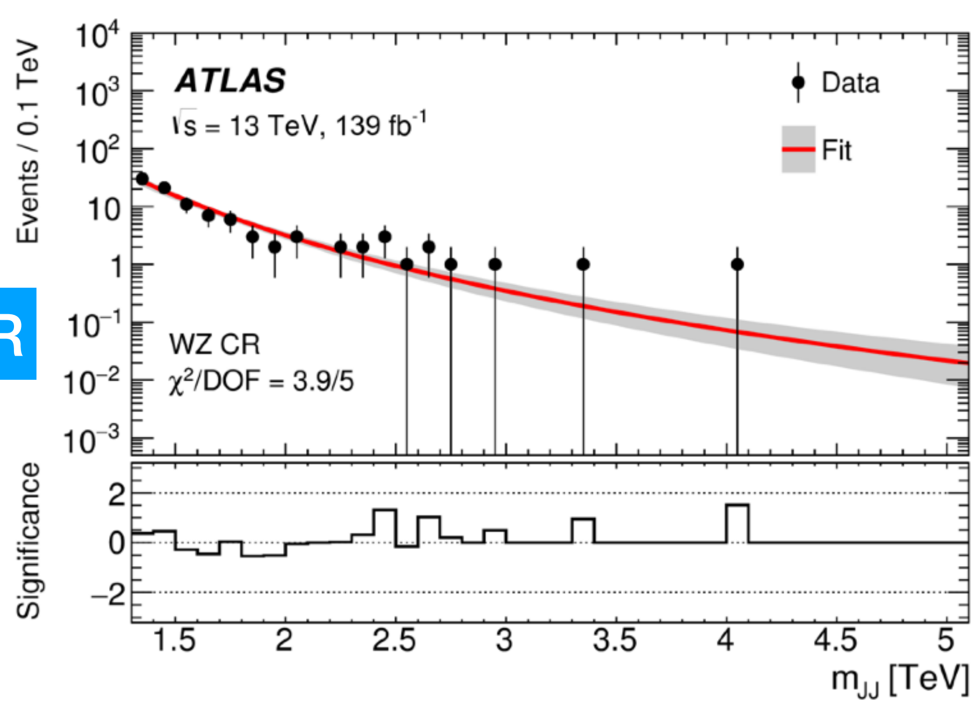
Very high bkg rejection is able to reduce the high amount of QCD bkg

Fully-hadronic strategy

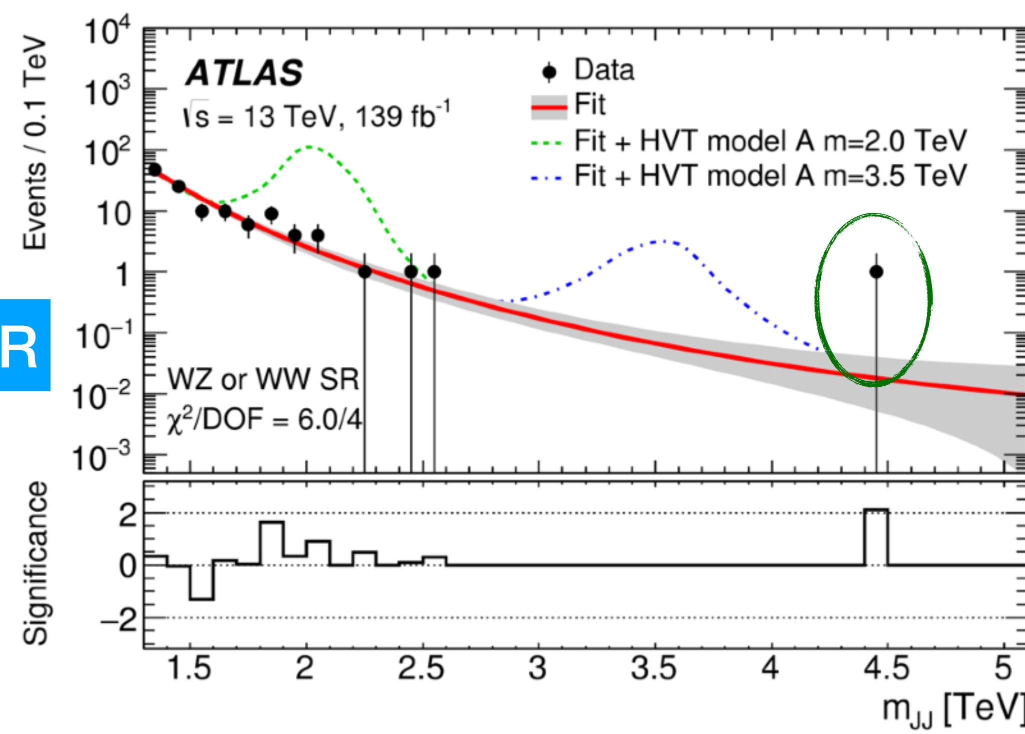
- Very simple analysis:
  - di-jets requirements: ***ΔY<1.2*** (reduce t-channel QCD) ***p_T-asymmetry<0.15*** (to reduce mis-reconstructed events)
  - di-tagging: ***boson tagger x 2***
- Fully data-driven bkg estimation:
  - modelling of the fit shape tested in CRs
  - bkg uncertainty driven by the ***fit uncertainty***
  - ***tagger uncertainty*** in dedicated V+jets CR



CR



SR

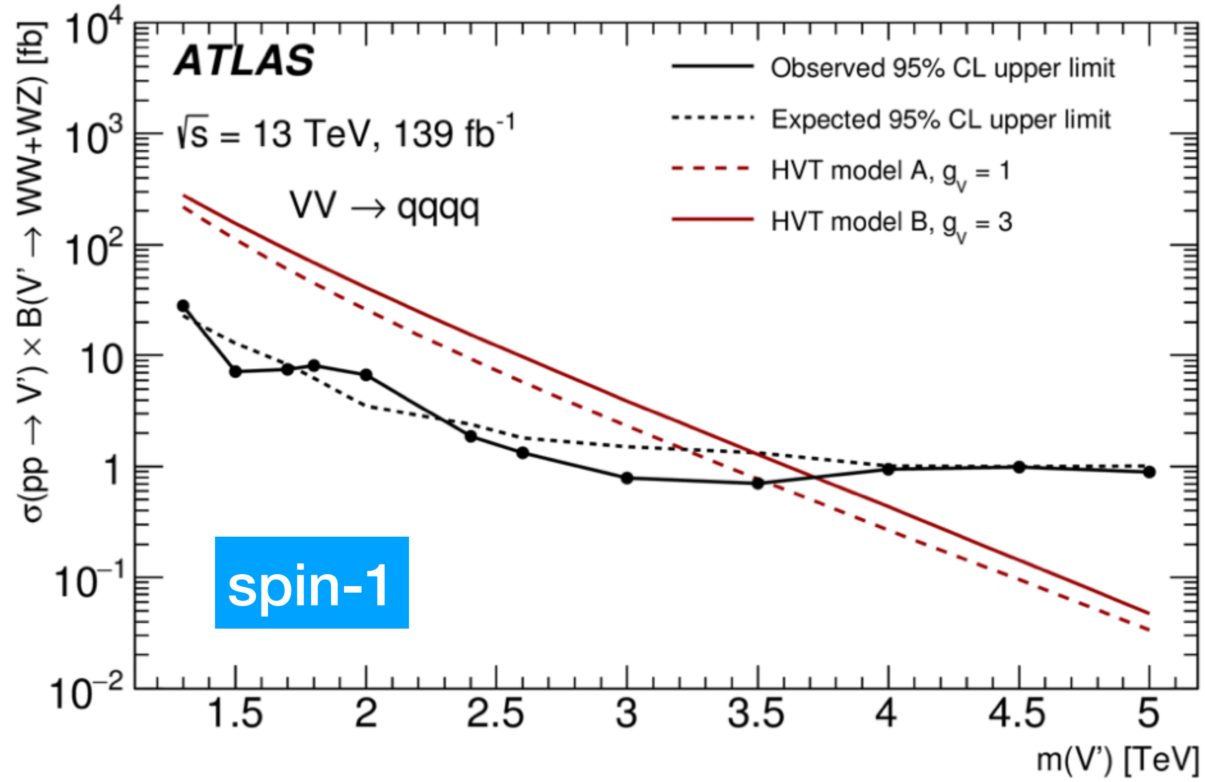


We collect data at very high m_{JJ} range

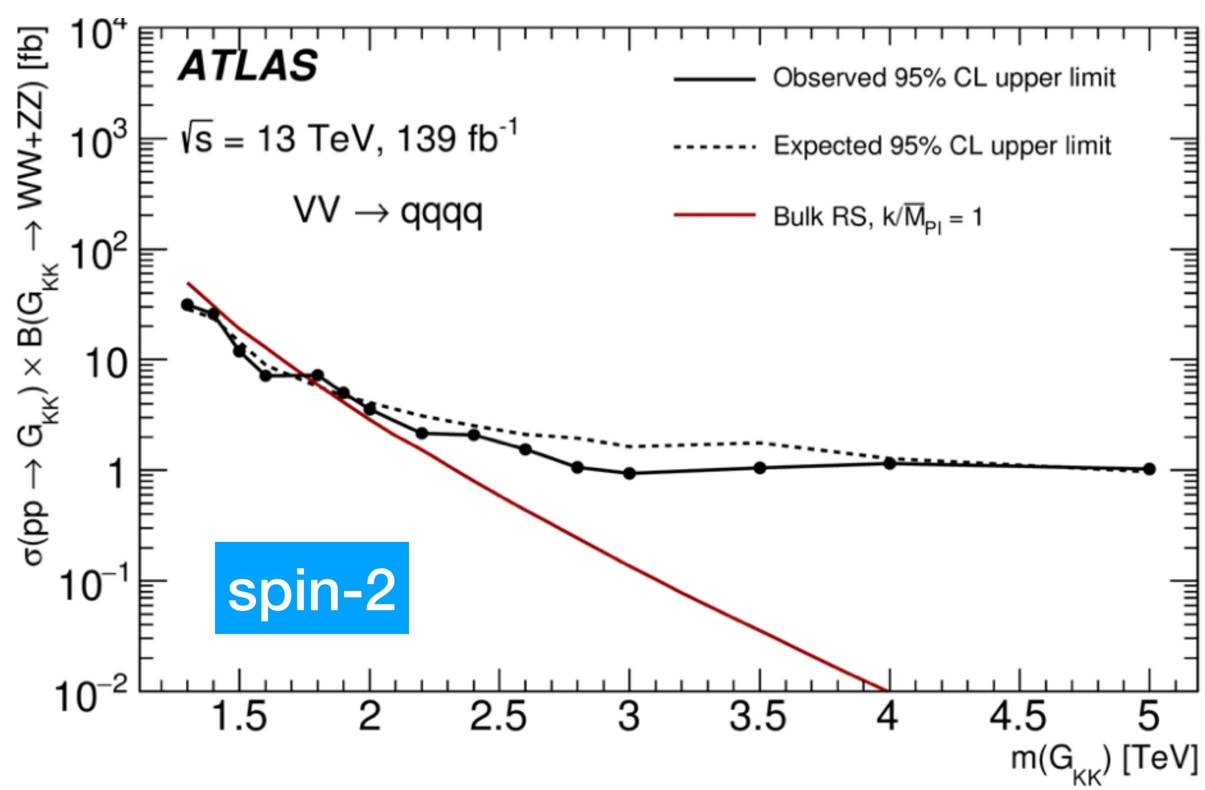
VV fully-hadronic results

HDBS-2018-31

- No significant deviation observed
  - interpretation of the observed data in terms of BSM models
- Very competitive results in the high mass regime
  - we can not say anything below 1.3 TeV



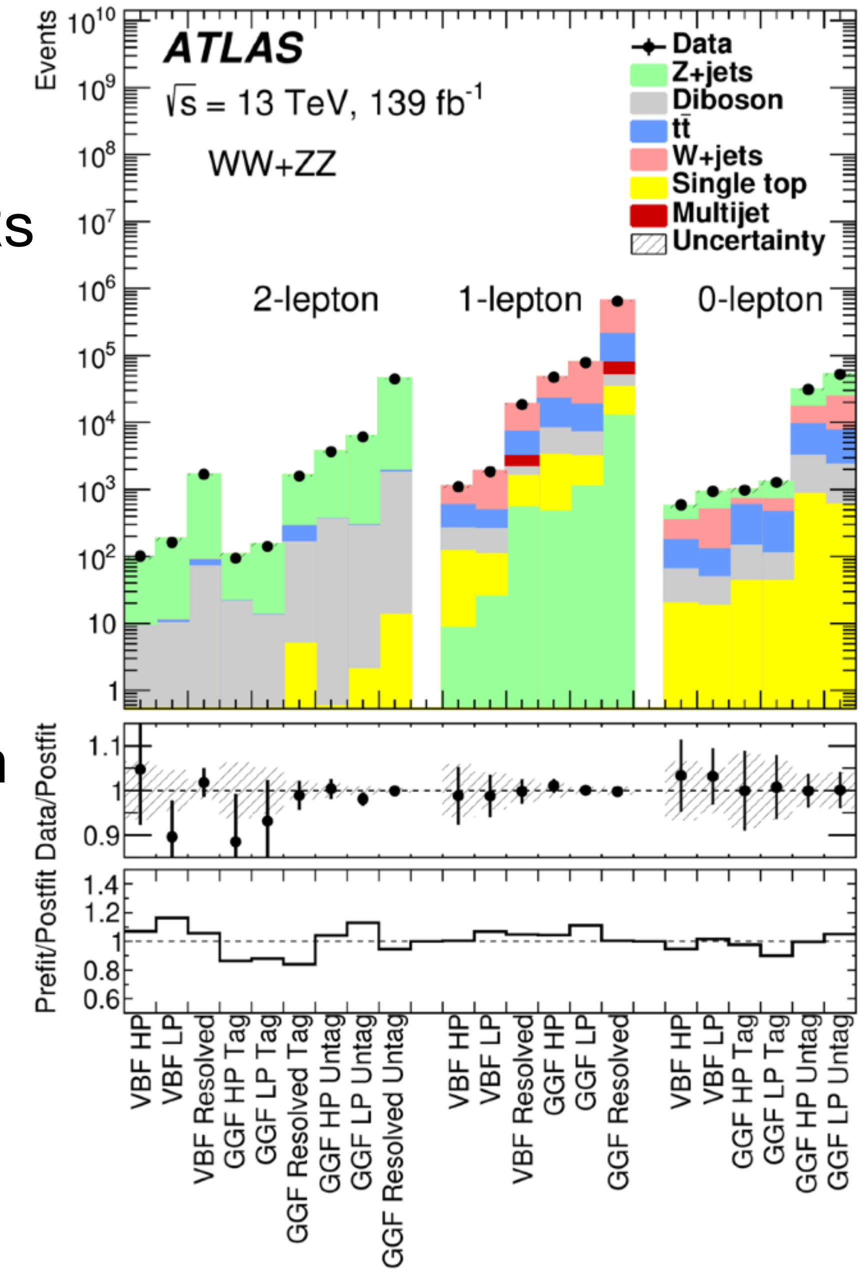
Model	Signal Region	Excluded mass range [TeV]
Radion	WW	none
	ZZ	none
	$WW+ZZ$	none
HVT model A, $g_V = 1$	WW	1.3–2.9
	WZ	1.3–3.4
	$WW+WZ$	1.3–3.5
HVT model B, $g_V = 3$	WW	1.3–3.1
	WZ	1.3–3.6
	$WW+WZ$	1.3–3.8
Bulk RS, $k/\overline{M}_{\rm Pl} = 1$	WW	1.3–1.6
	ZZ	none
	$WW+ZZ$	1.3–1.8

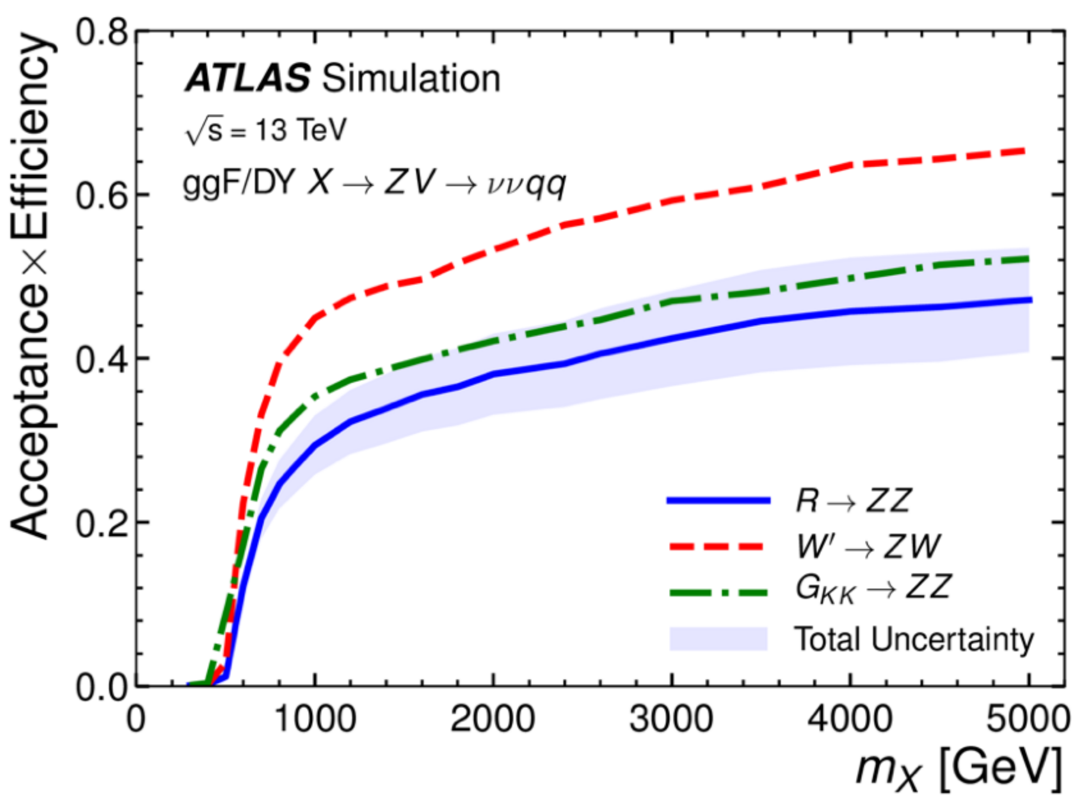


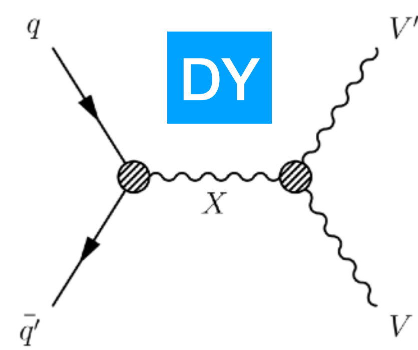
Semi-leptonic signatures

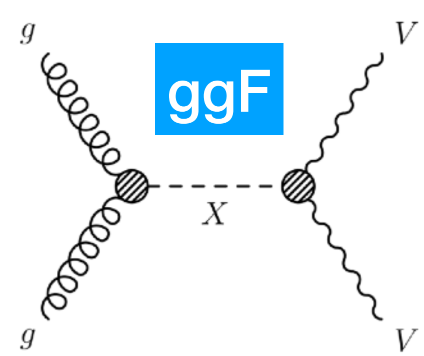
[HDBS-2018-10]

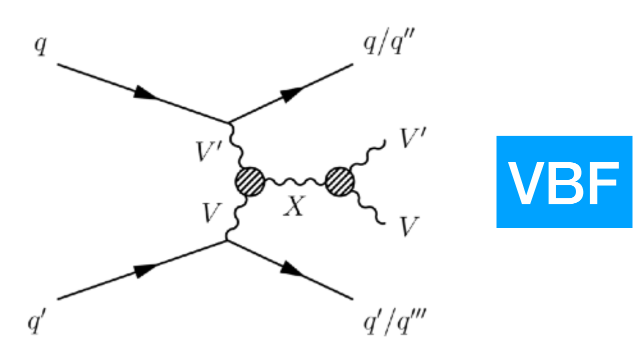
- Semi-leptonic channels has ***lower BR but cleaner environment*** w.r.t. fully-hadronic channel
  - MC-based bkg estimation with quite complex SRs/CRs scheme definition
- Different channels to be combined:
  - ***3 leptons channels***: ll, lv, vv
  - 2 V-hadronic reconstruction techniques: ***resolved/boosted***
  - ***VBF vs ggF/DY*** production mechanism categorisation









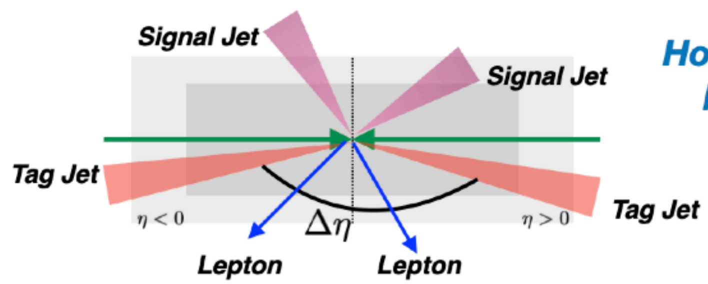


VBF categorisation using ML

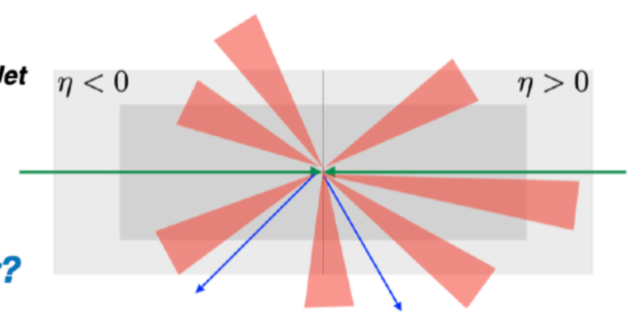
HDBS-2018-10

PhDThesis.cds.2743418

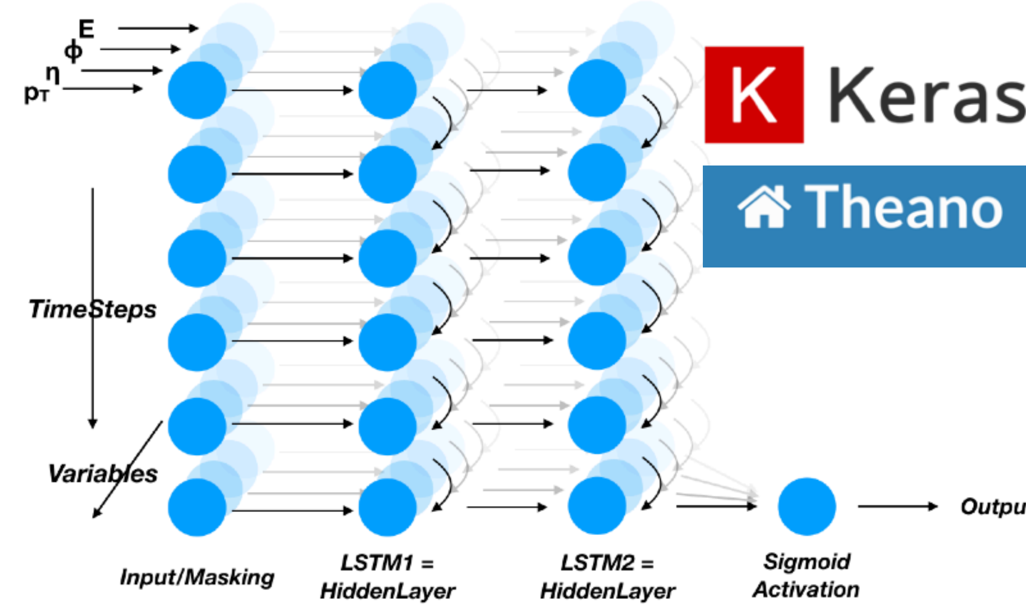
- Novel ML technique introduced to categorise VBF vs ggF/DY events
  - RNN dealing with ***low-level jets*** input variables
- improved performances w.r.t. cut based approaches
  - ***1-jet VBF*** events naturally recovered

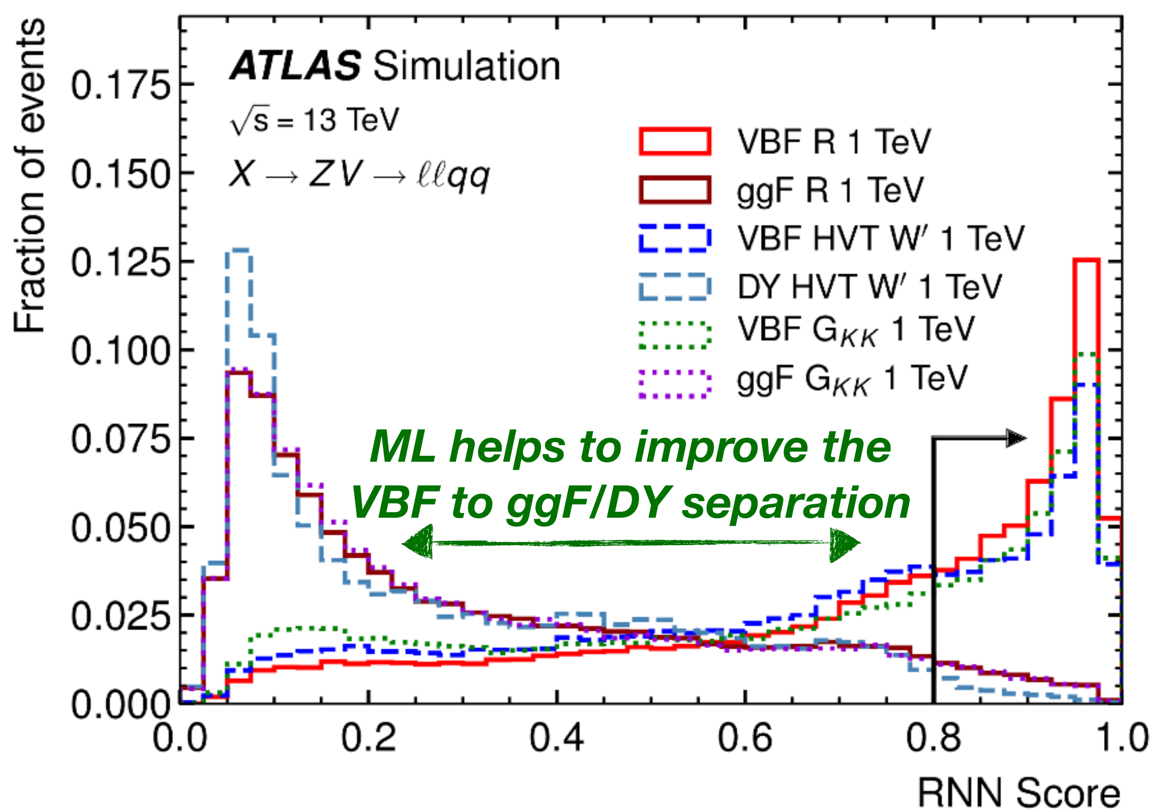


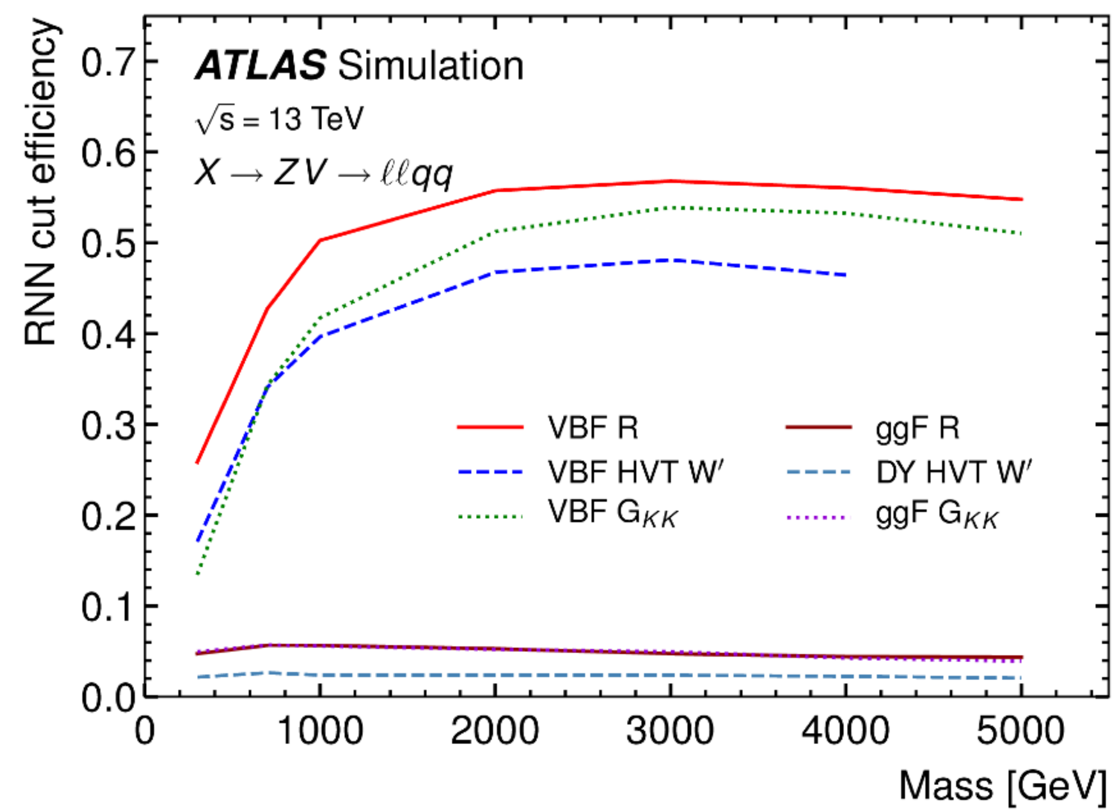
How do we expect to be a VBF event?



How do we see an event in the detector?

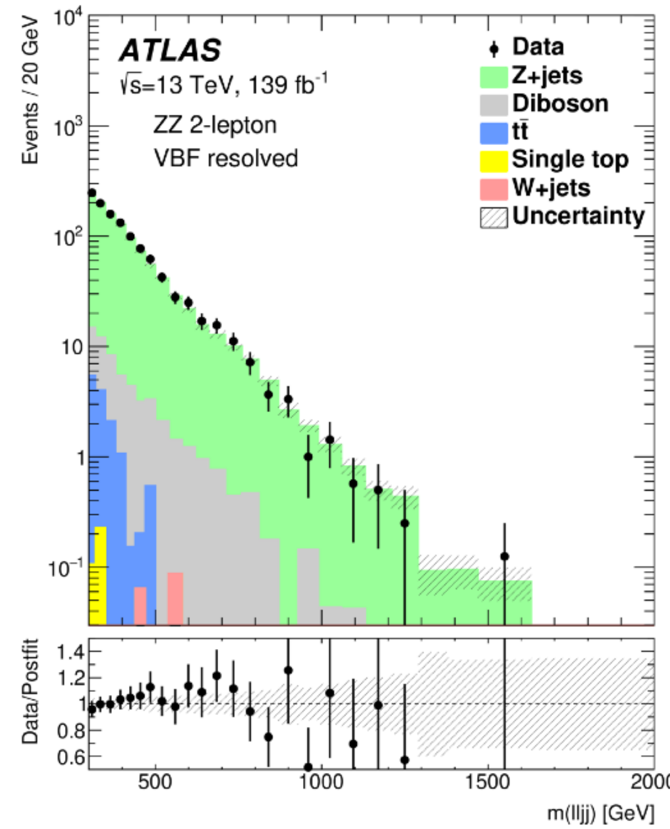


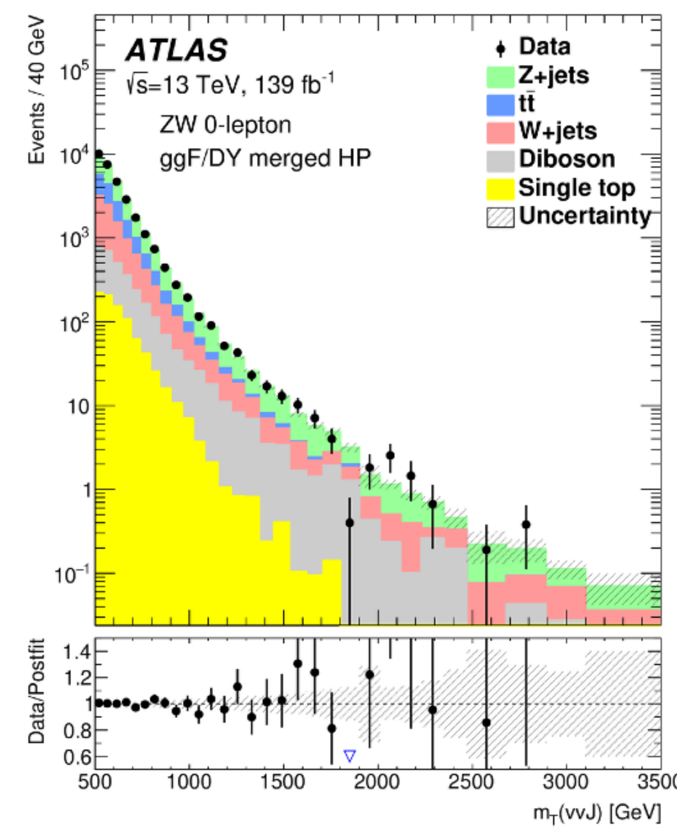


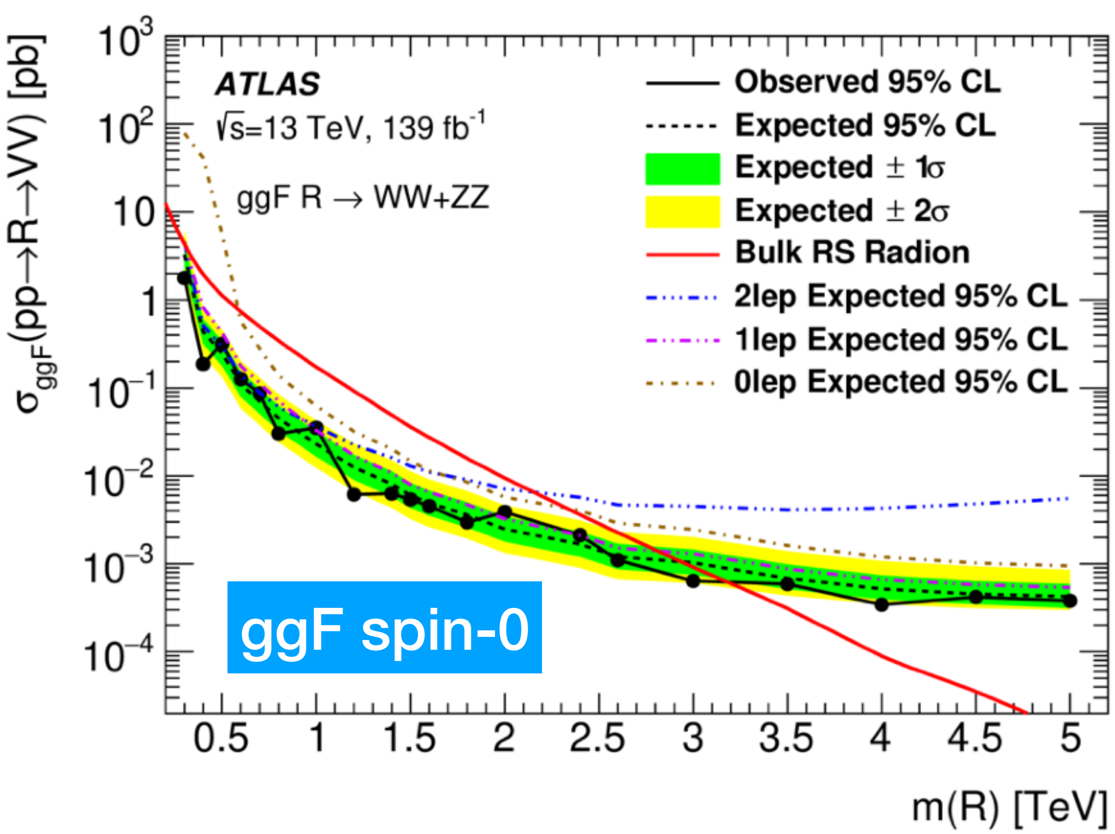


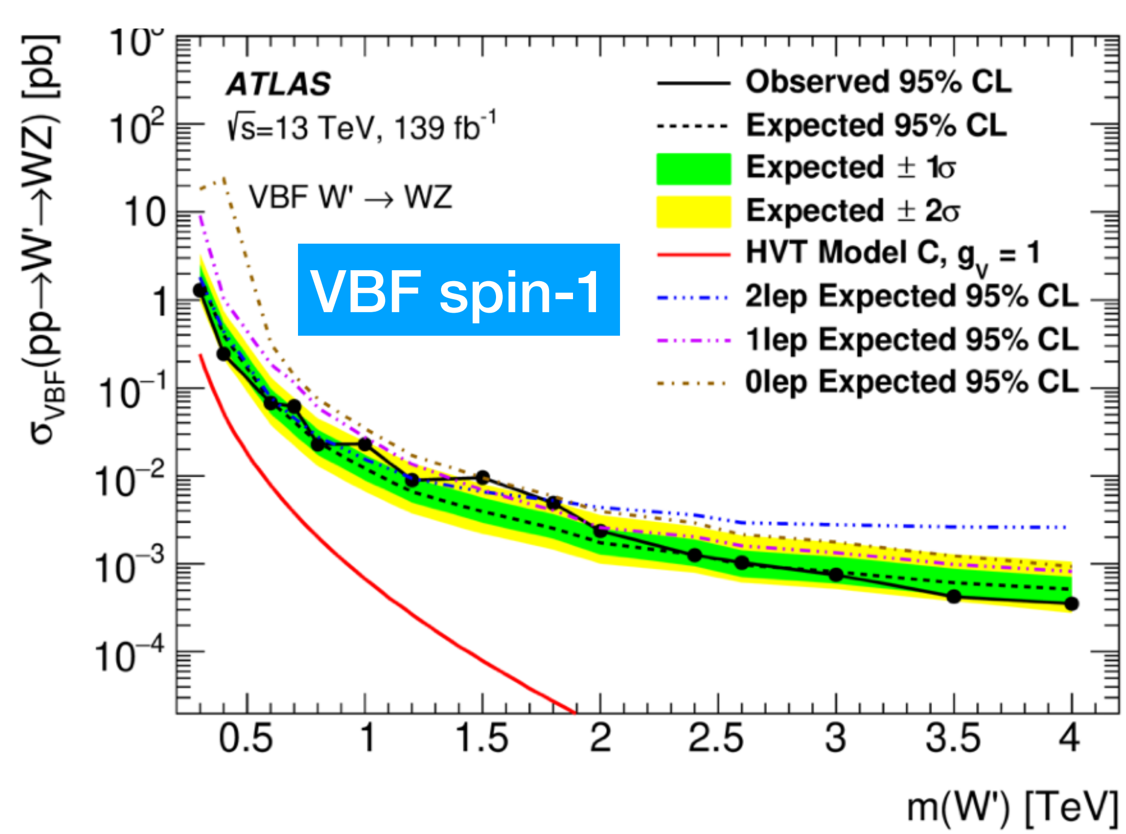
VV semi-leptonic results

- No significant deviation observed
  - new data are interpreted as upper bounds on the cross section of new models
- Different channels let to cover large p.s.:
  - 0/1 lepton channels contributing at high-mass, 2 lepton at low-mass
  - VBF regions covered, still need to be improved to get exclusions



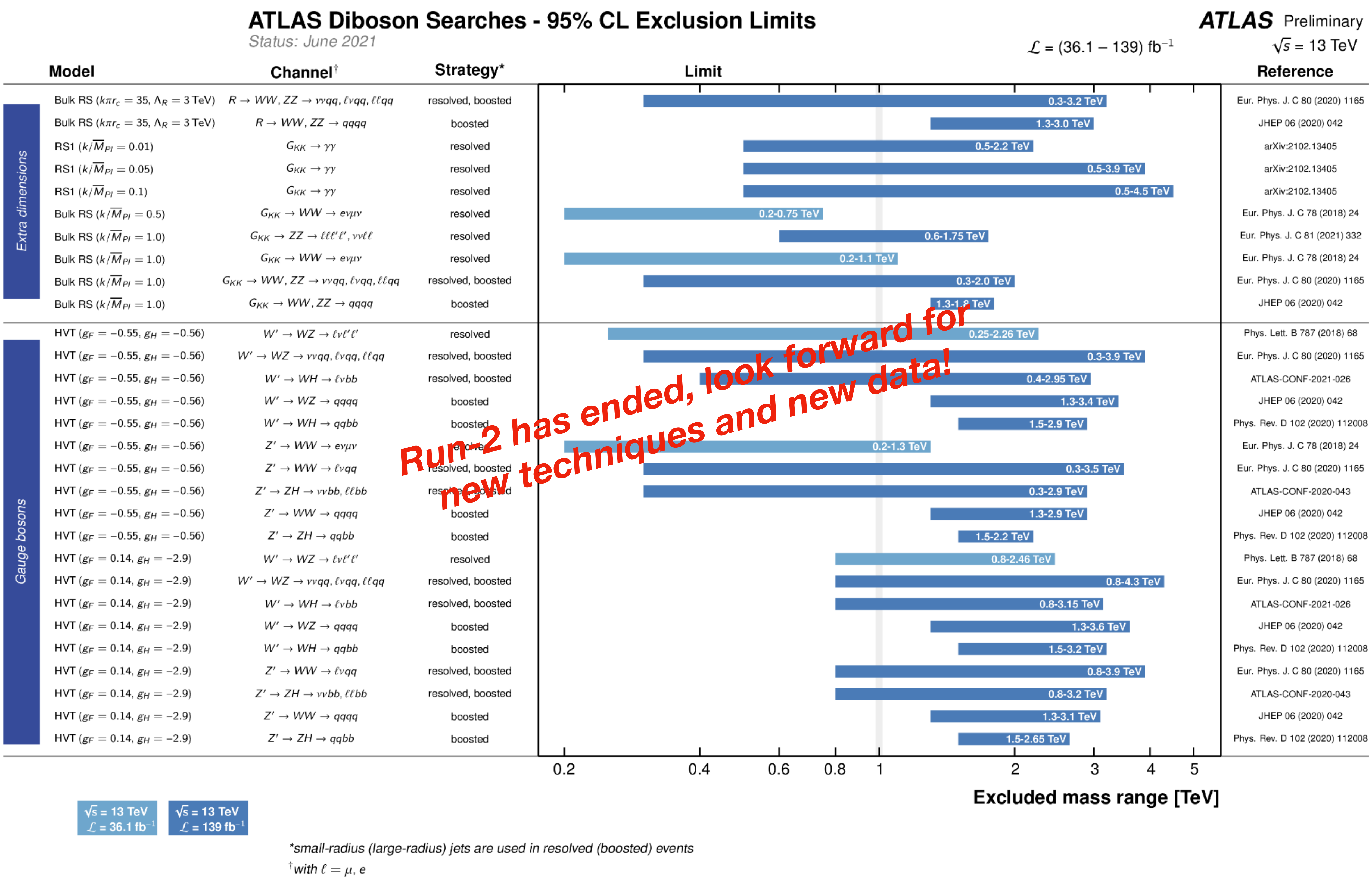






Conclusions & perspectives

PUB-2021-018



Run-2 has ended, look forward for new techniques and new data!

Conclusions & perspectives

High-mass VV resonance searches during Run-2

- Different channels very competitive

  Fully-hadronic channel:

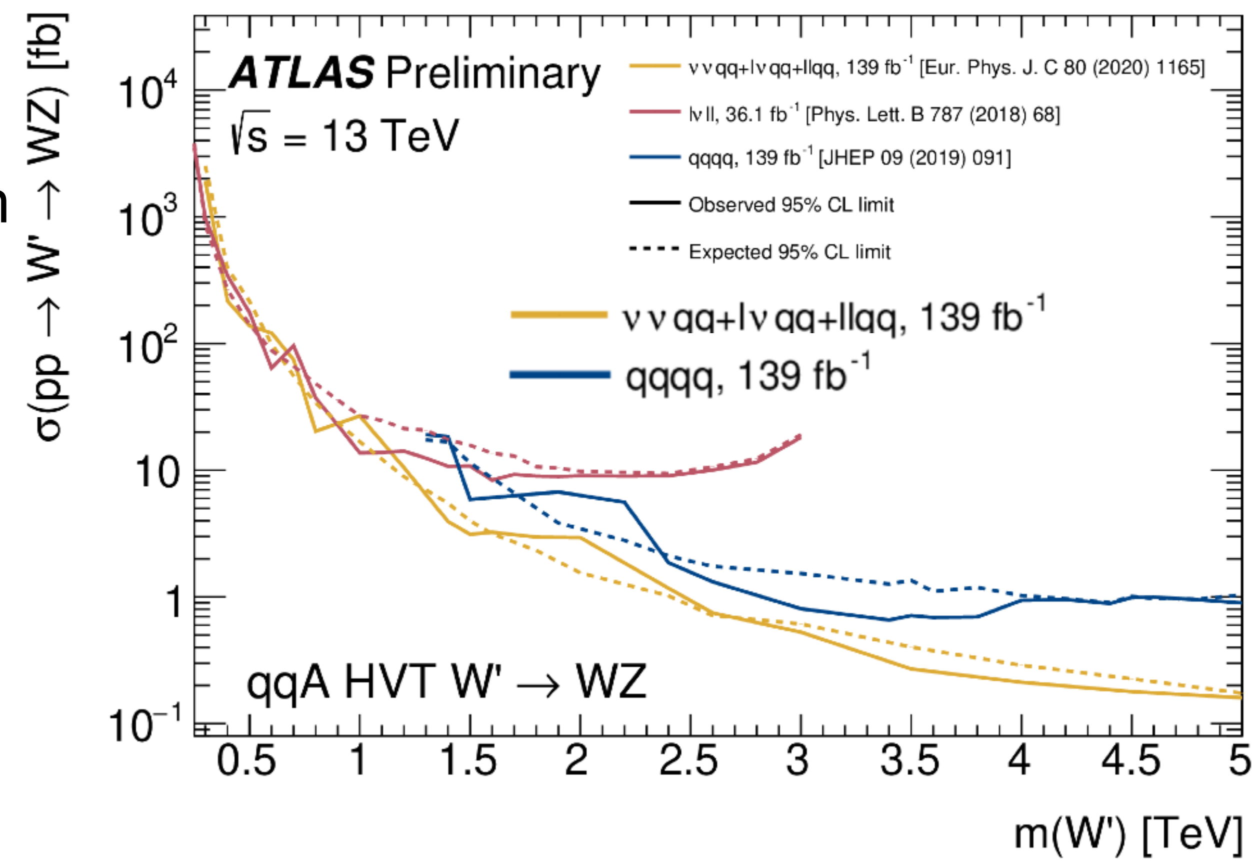
  - large-R(=1.) jets and very powerful boson tagger technique are the keys to exploit high mass range

  Semi-leptonic channel:

  - new ML techniques already exploited for VBF categorisation events (VBF-RNN)

- Several models and mass ranges excluded

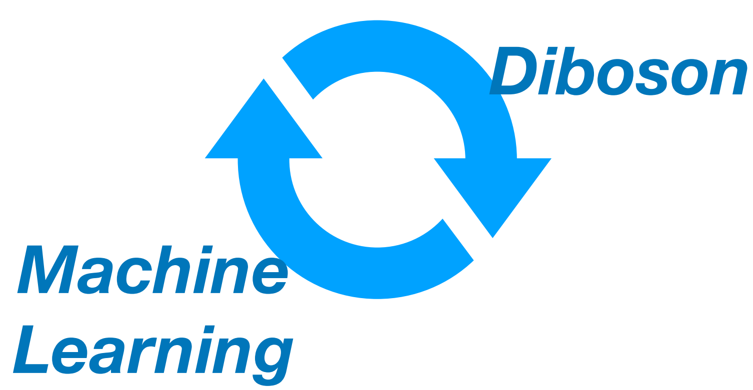
[PUB-2021-018]



What is next?

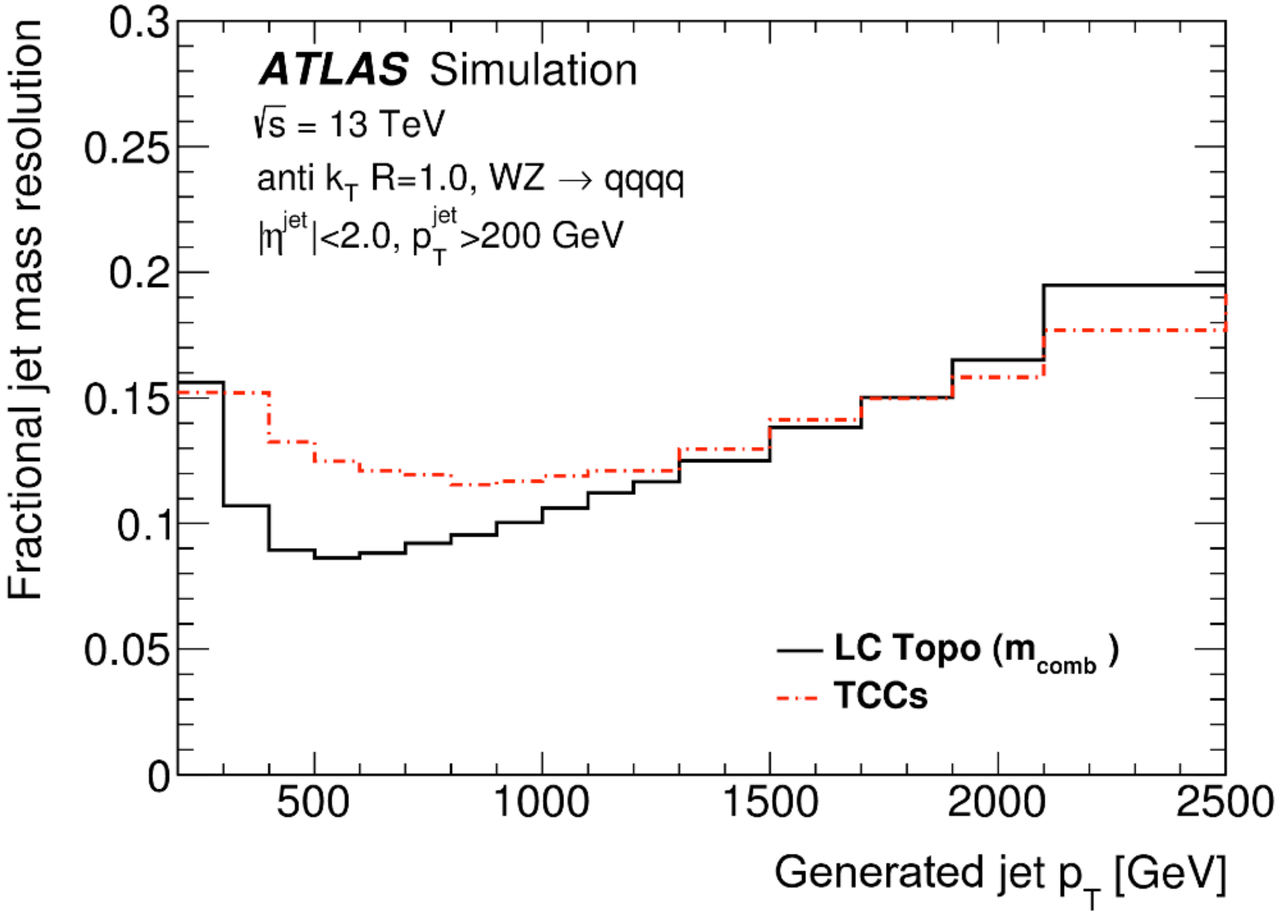
Both channels are re-analysing Run-2 for improved results and preparation for Run-3:

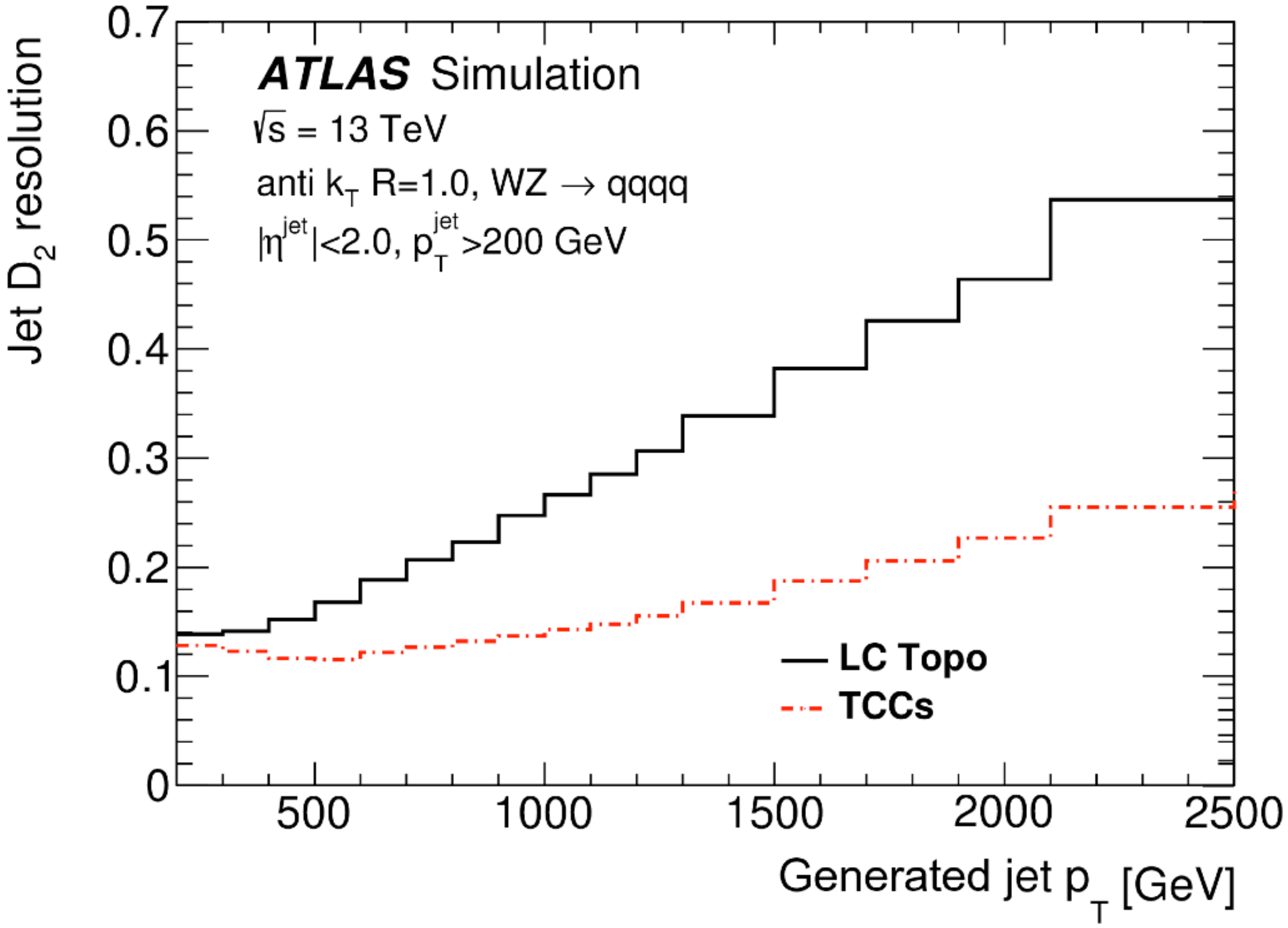
- fully-hadronic: new jets reconstruction and ML tagging techniques
- semi-leptonic: extension of ML approaches



backup

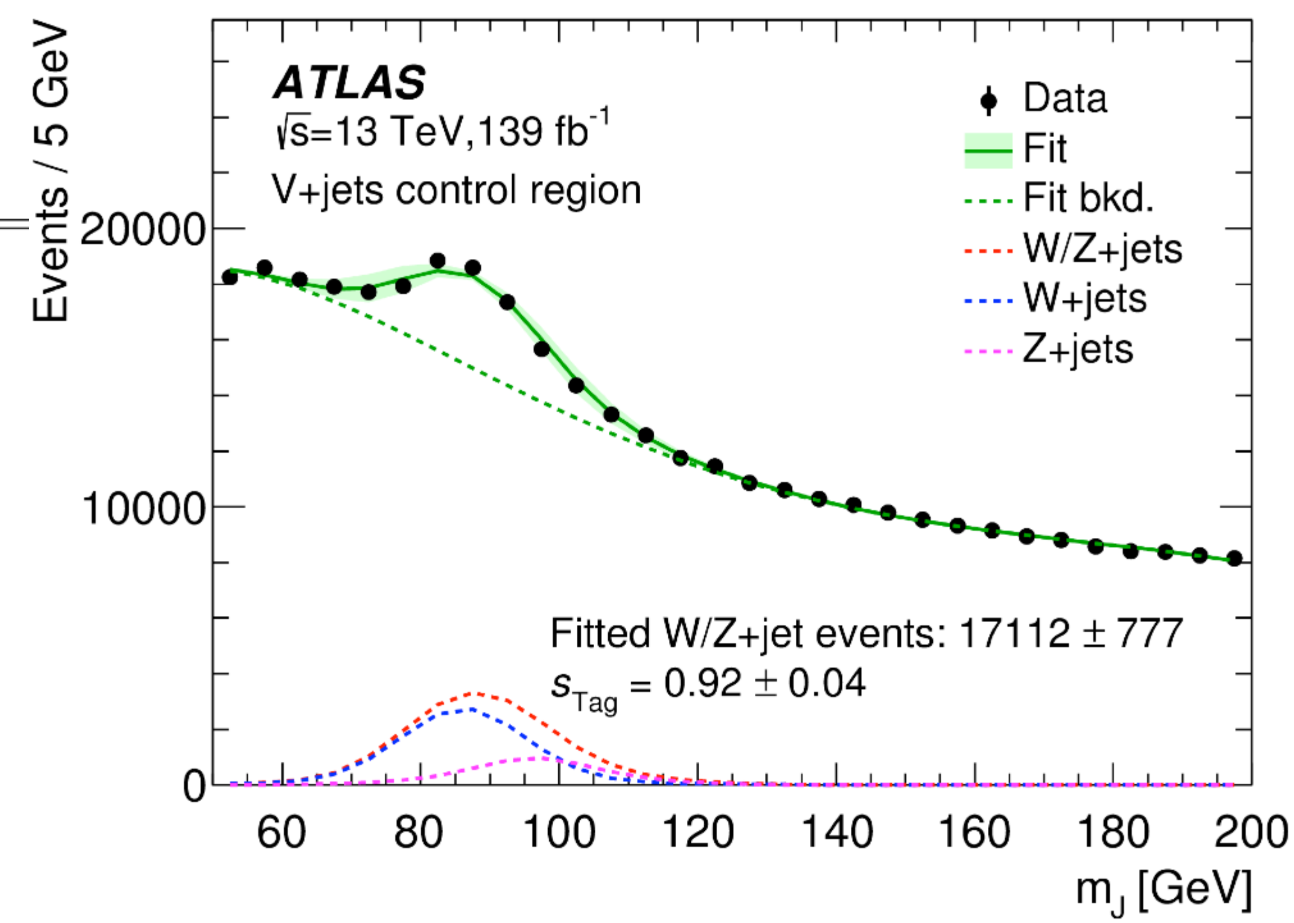
LCTopo VS TCC



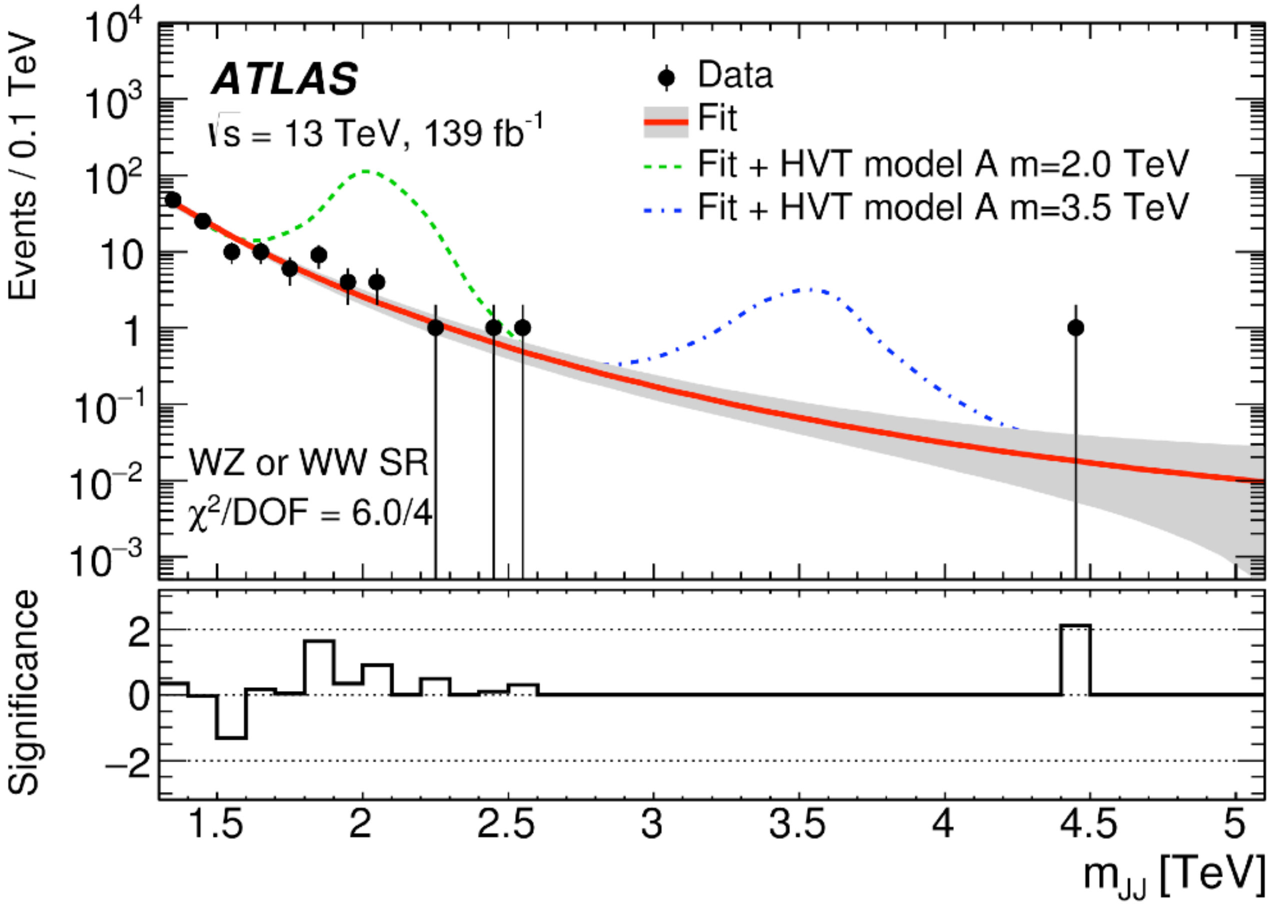


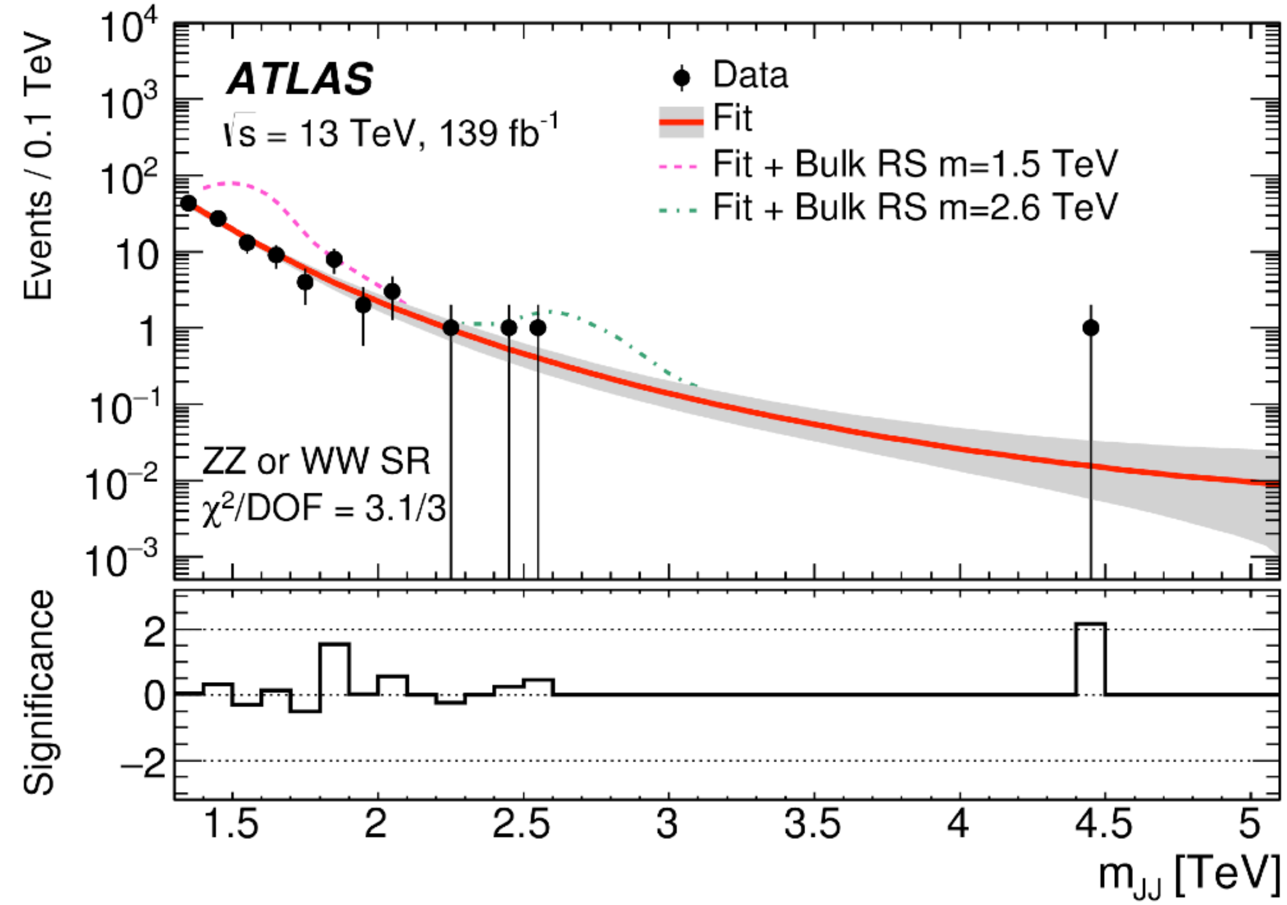
Event Selection

Signal region	Veto events with leptons:
	No e or μ with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$
	Event pre-selection:
	≥ 2 large-R jets with $|\eta| < 2.0$ and mass > 50 GeV
	$p_{\mathrm{T1}} > 500$ GeV and $p_{\mathrm{T2}} > 200$ GeV
	$m_{\mathrm{JJ}} > 1.3$ TeV
	Topology and boson tag:
	$|\Delta y| = |y_1 - y_2| < 1.2$
	$A = (p_{\mathrm{T1}} - p_{\mathrm{T2}}) / (p_{\mathrm{T1}} + p_{\mathrm{T2}}) < 0.15$
	Boson tag with D_2 variable, n_{trk} variable, and W or Z mass window
V+jets control region	Veto events with leptons:
	No e or μ with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$
	V+jets selection:
	≥ 2 large-R jets with $|\eta| < 2.0$
	$p_{\mathrm{T1}} > 600$ GeV and $p_{\mathrm{T2}} > 200$ GeV
	Boson tag with D_2 and n_{trk} variables on either jet
	Anti-boson tag with D_2 variable on other jet



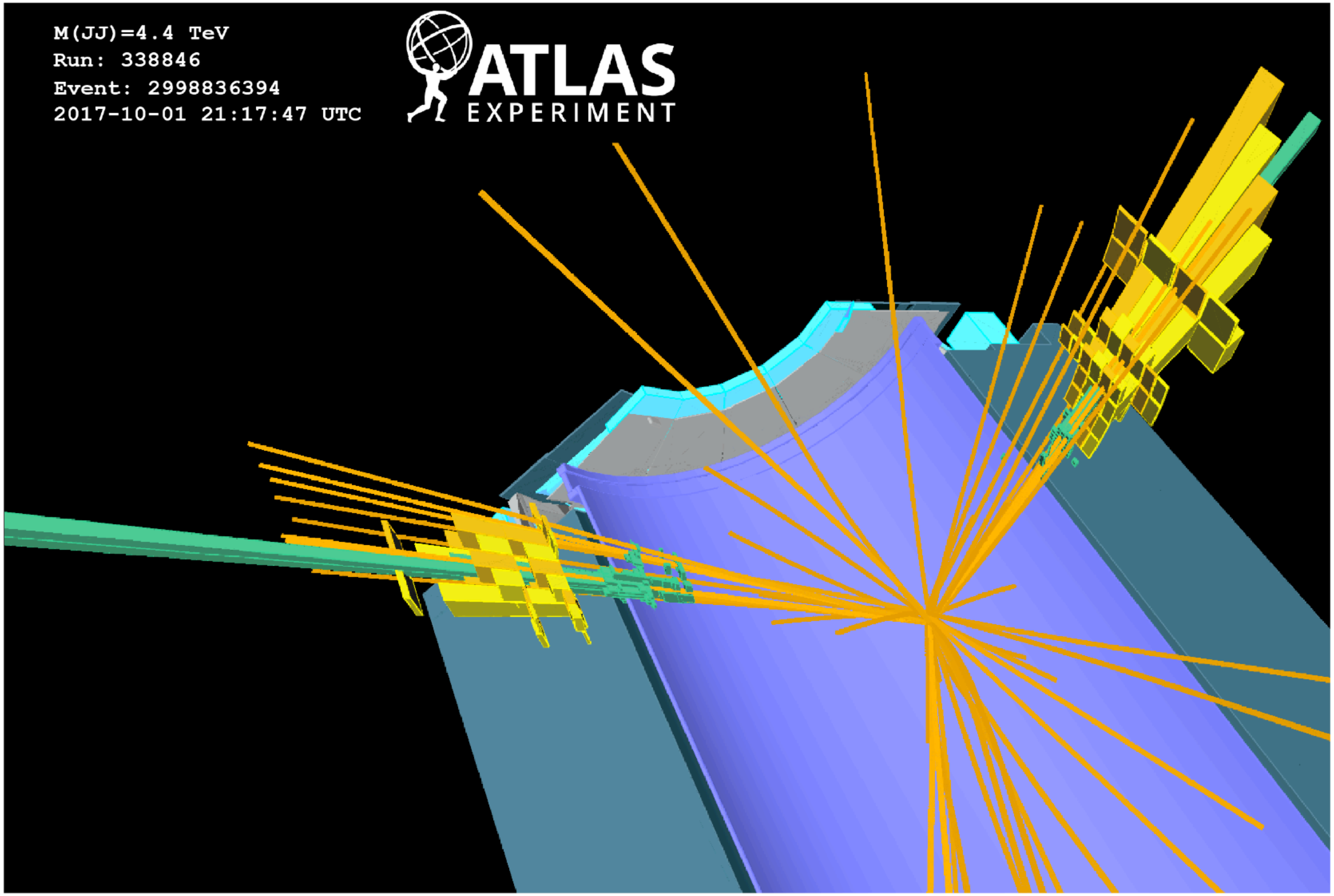
SRs, fit, uncertainties



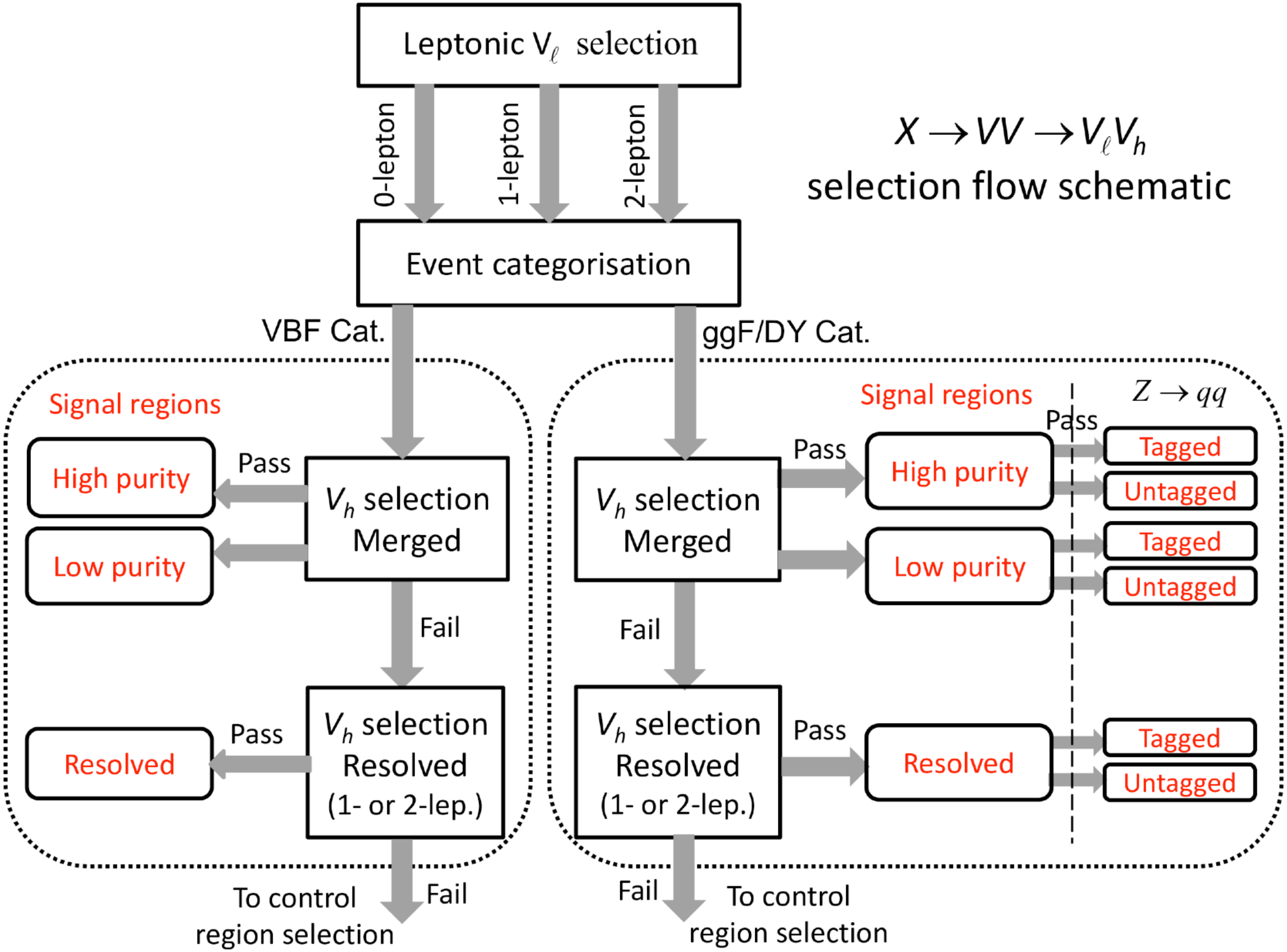


Source	$m(V') = 1.4$ TeV	$m(G_{KK}) = 2.8$ TeV
Fit	32%	2.4%
Jet p_{T} scale	29%	1.5%
Total systematic uncertainty	79%	17%
Statistical uncertainty	61%	99%

Event display mJJ = 4.4 TeV



Analysis flow



Event selection

Event selection	0-lepton ($ZV \to \nu\nu V_h$)	1-lepton ($WV \to \ell\nu V_h$)	2-lepton ($ZV \to \ell\ell V_h$)
V_ℓ selection	No *Loose* lepton $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV	1 *Tight* electron or 1 *Medium* muon with $p_{\mathrm{T}}^{\ell} > 30$ GeV $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV $p_{\mathrm{T}}^{V_\ell} > 75$ GeV	2 *Loose* leptons with $p_{\mathrm{T}}^{\ell} > 30$ GeV from the $Z \to \ell\ell$ candidate
Event veto	No additional *Loose* leptons Veto events with b-jets not associated with the $Z \to qq$ candidate		
Event categorisation	≥ 1 large-R jets or ≥ 2 small-R jets VBF and ggF/DY classification according to RNN score		
V_h selection (Merged)		$E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV $p_{\mathrm{T}}^{V_\ell} > 200$ GeV	
	≥ 1 large-R jets The leading jet passing p_{T}-dependent m_J requirement		
		$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)	$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)
V_h selection (Resolved)	Not Performed	Failed merged selection ≥ 2 small-R jets with $\lvert\eta\rvert < 2.5$ $62 < m_{jj} < 97$ GeV for $W \to jj$ $70 < m_{jj} < 105$ GeV for $Z \to jj$	
		$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.25$ (VBF)	$\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (ggF/DY) $\mathcal{R}_{p_{\mathrm{T}}/m} > 0.35$ (VBF)

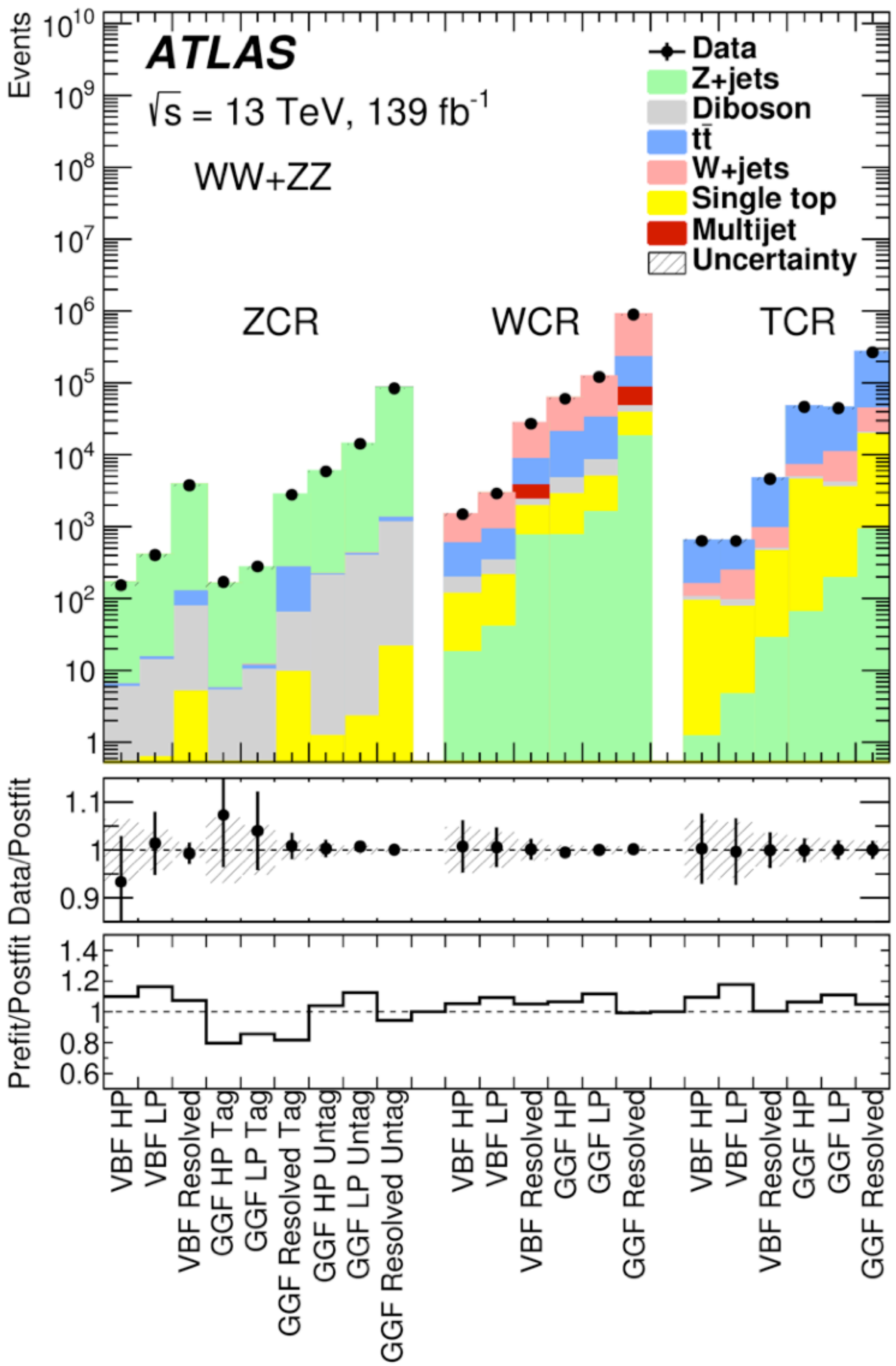
Uncertainties

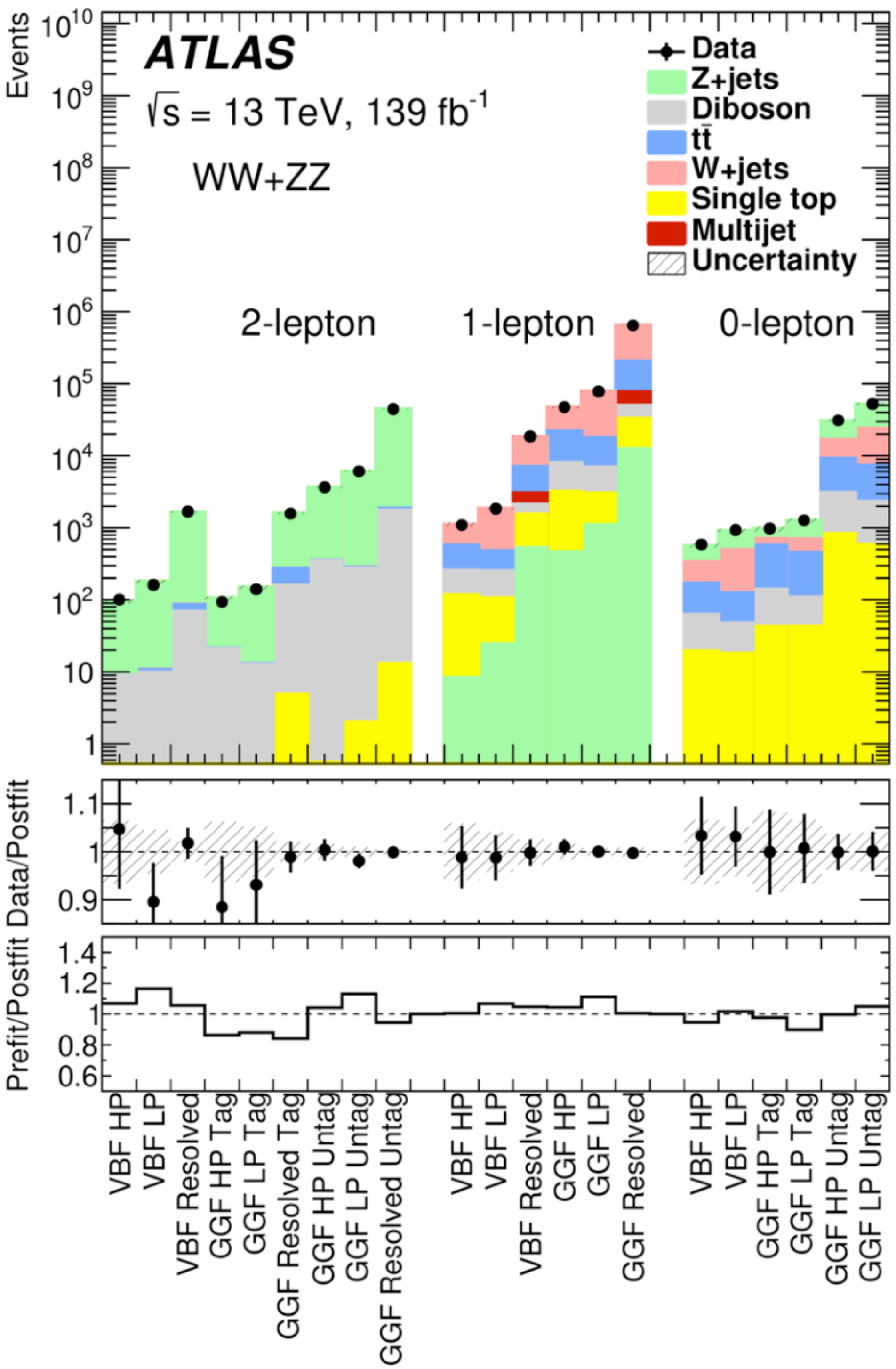
$m(G_{KK}) = 600$ GeV		$m(G_{KK}) = 2$ TeV	
Uncertainty source	$\Delta\mu/\mu$ [%]	Uncertainty source	$\Delta\mu/\mu$ [%]
Total	50	Total	59
Statistical	29	Statistical	48
Systematic	41	Systematic	34
Large-R jet	18	Large-R jet	24
MC statistics	16	MC statistics	17
Background normalisations	15	W/Z+jets modelling	15
Diboson modelling	12	Flavour tagging	5.5
W/Z+jets modelling	11	$t\bar{t}$ modelling	4.2
Small-R jet	9.7	Diboson modelling	3.9
$t\bar{t}$ modelling	8.1	Single-t modelling	3.3

Some yields…

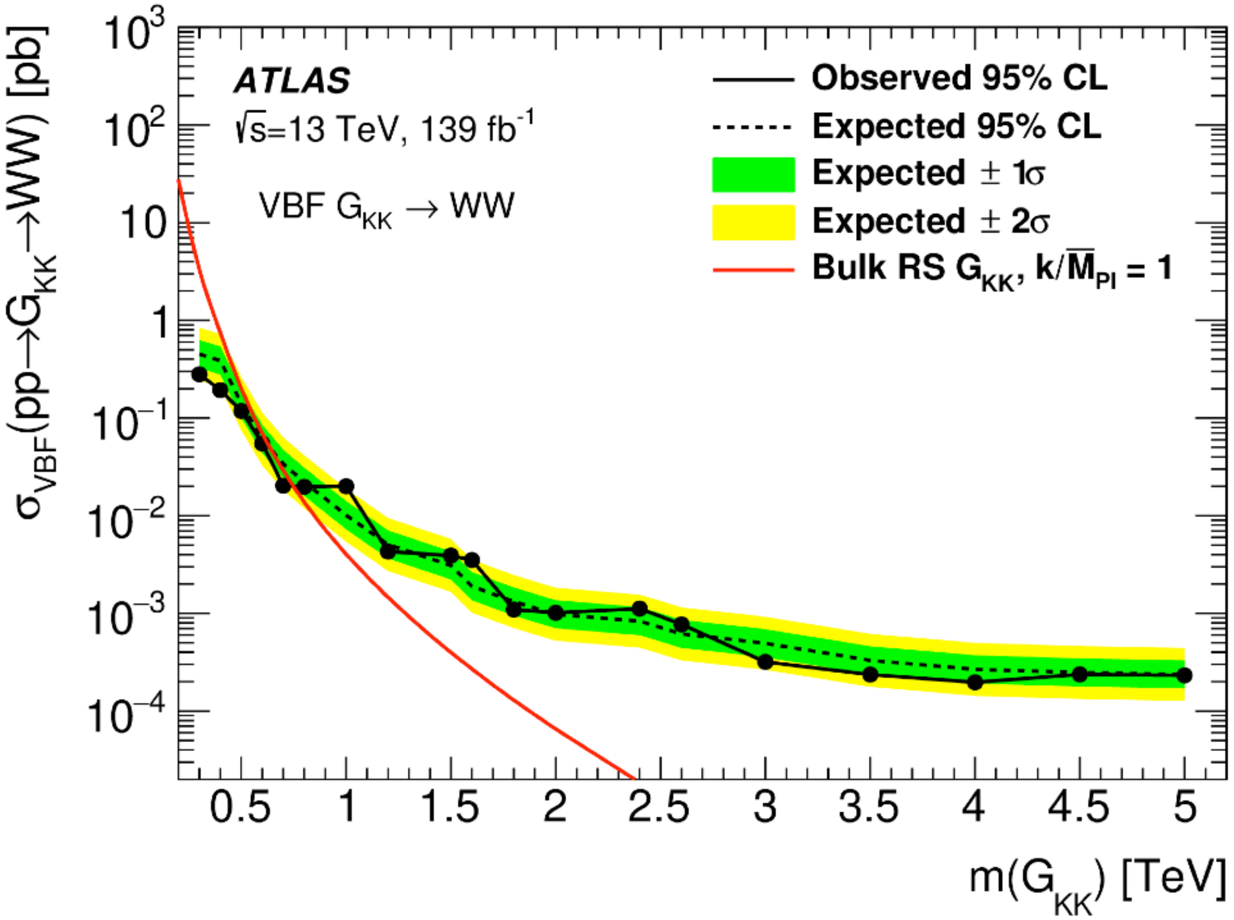
Channel	$V \to qq$ recon.	Signal regions	Background estimates: W+jets	Z+jets	$t\bar{t}$	Diboson	Single-t	Multijet	Total
			VBF category						
0-lepton (ZZ)	Merged	HP	169 ± 12	228 ± 16	102 ± 10	51 ± 10	24 ± 4	-	574 ± 25
		LP	370 ± 23	411 ± 20	75 ± 8	30 ± 4	21 ± 4	-	906 ± 33
			ggF/DY category						
	Merged	HP Tag	133 ± 14	270 ± 40	437 ± 31	100 ± 10	45 ± 7	-	982 ± 60
		HP Untag	7600 ± 400	14 300 ± 600	6030 ± 270	2300 ± 180	840 ± 110	-	31 100 ± 800
		LP Tag	259 ± 28	560 ± 50	342 ± 24	67 ± 7	43 ± 7	-	1270 ± 70
		LP Untag	16 300 ± 900	28 600 ± 1100	5040 ± 220	1760 ± 150	600 ± 80	-	52 400 ± 1500
			VBF category						
1-lepton (WW)	Merged	HP	530 ± 28	8.3 ± 0.5	321 ± 22	141 ± 27	113 ± 21	-	1110 ± 50
		LP	1380 ± 40	24.5 ± 1.1	228 ± 17	150 ± 33	83 ± 16	-	1870 ± 60
	Resolved		11 360 ± 190	530 ± 10	4060 ± 130	590 ± 80	1070 ± 210	960 ± 110	18 570 ± 340
			ggF/DY category						
	Merged	HP	24 820 ± 170	463 ± 5	13 890 ± 220	4910 ± 250	2800 ± 400	-	46 900 ± 500
		LP	60 270 ± 240	1095 ± 8	11 050 ± 160	3950 ± 210	1970 ± 250	-	78 300 ± 400
	Resolved		443 500 ± 1800	12 480 ± 40	126 000 ± 1500	16 800 ± 1200	21 200 ± 2800	27 200 ± 1400	647 000 ± 4000
			VBF category						
2-lepton (ZZ)	Merged	HP	0	87 ± 6	0.081 ± 0.009	9.6 ± 1.2	0	-	97 ± 6
		LP	0.133 ± 0.011	170 ± 8	0.85 ± 0.07	9.9 ± 1.2	0.43 ± 0.07	-	181 ± 8
	Resolved		0.272 ± 0.012	1566 ± 29	17.0 ± 0.7	72 ± 10	0.48 ± 0.32	-	1656 ± 31
			ggF/DY category						
	Merged	HP Tag	0.0135 ± 0.0043	85 ± 6	0.283 ± 0.035	21.1 ± 2.3	0.34 ± 0.05	-	107 ± 7
		HP Untag	0.772 ± 0.010	3300 ± 40	4.27 ± 0.08	361 ± 32	0.58 ± 0.11	-	3670 ± 50
		LP Tag	0.0135 ± 0.0043	138 ± 8	0.313 ± 0.034	12.8 ± 1.4	0.30 ± 0.04	-	152 ± 8
		LP Untag	2.341 ± 0.017	5920 ± 50	10.16 ± 0.16	278 ± 26	2.03 ± 0.29	-	6220 ± 60
	Resolved	Tag	-	1323 ± 26	110 ± 10	159 ± 12	4.7 ± 0.8	-	1600 ± 30
		Untag	4.681 ± 0.026	42 750 ± 160	110.6 ± 1.5	1800 ± 100	13.4 ± 2.0	-	44 650 ± 190

Spin-0/2 regions

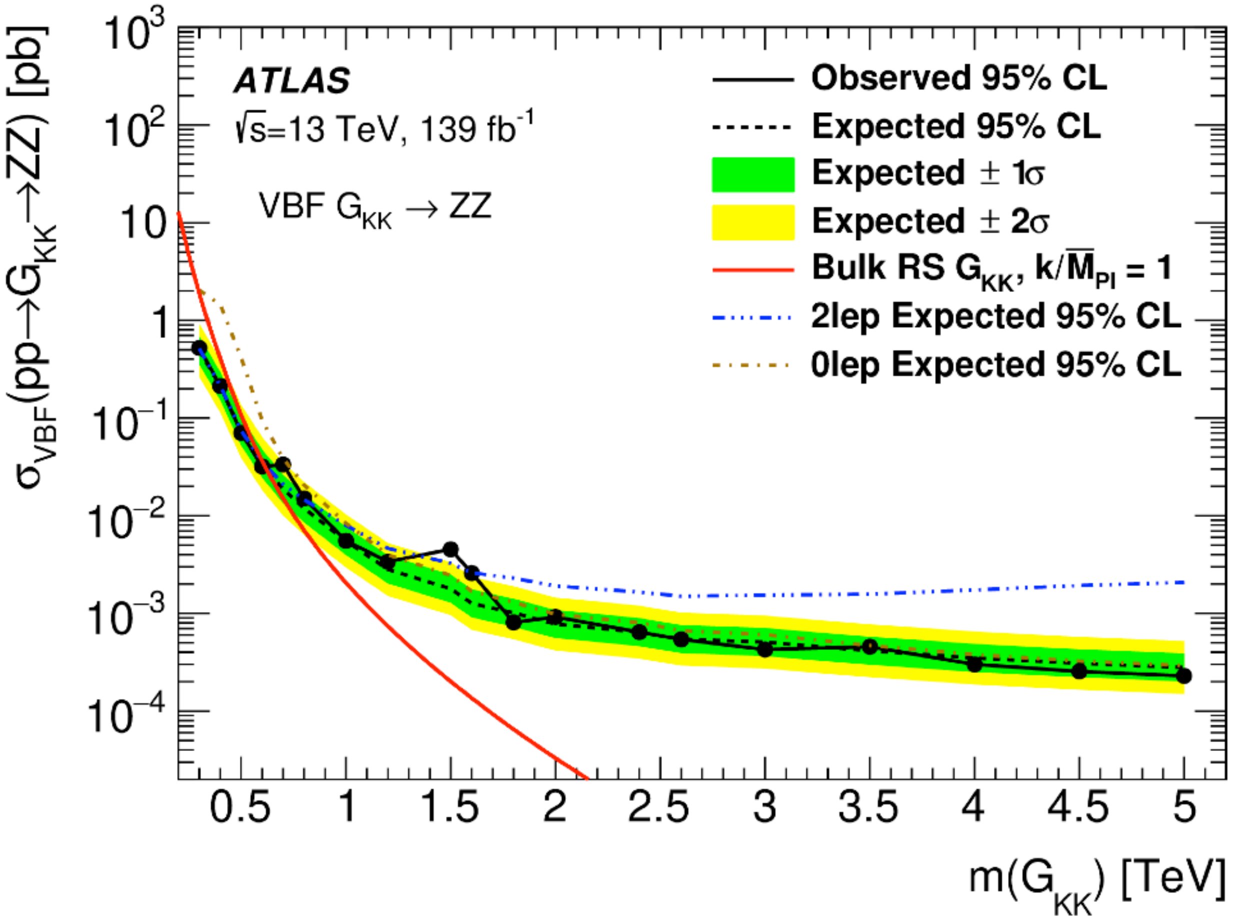




VBF spin-2



ATLAS
√s=13 TeV, 139 fb⁻¹
VBF G_KK → WW
Observed 95% CL
Expected 95% CL
Expected ± 1σ
Expected ± 2σ
Bulk RS G_KK, k/M̄_Pl = 1
σ_VBF(pp→G_KK→WW) [pb]
m(G_KK) [TeV]



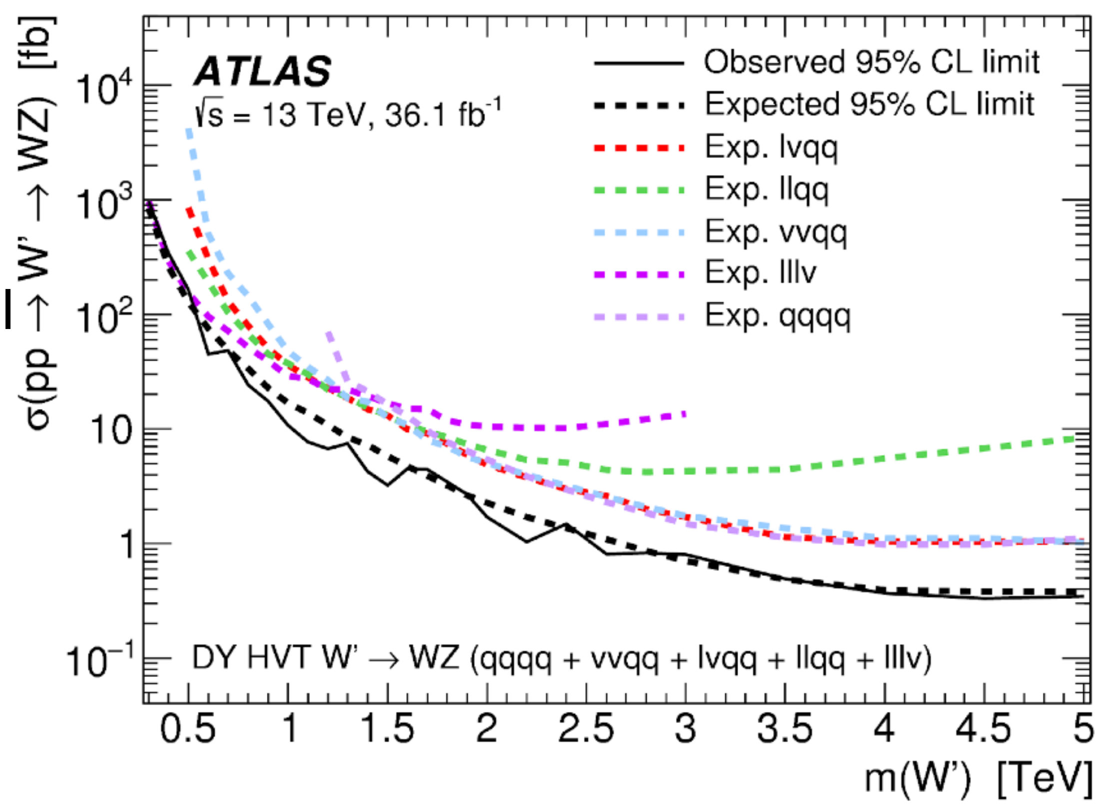
ATLAS
√s=13 TeV, 139 fb⁻¹
VBF G_KK → ZZ
Observed 95% CL
Expected 95% CL
Expected ± 1σ
Expected ± 2σ
Bulk RS G_KK, k/M̄_Pl = 1
2lep Expected 95% CL
0lep Expected 95% CL
σ_VBF(pp→G_KK→ZZ) [pb]
m(G_KK) [TeV]

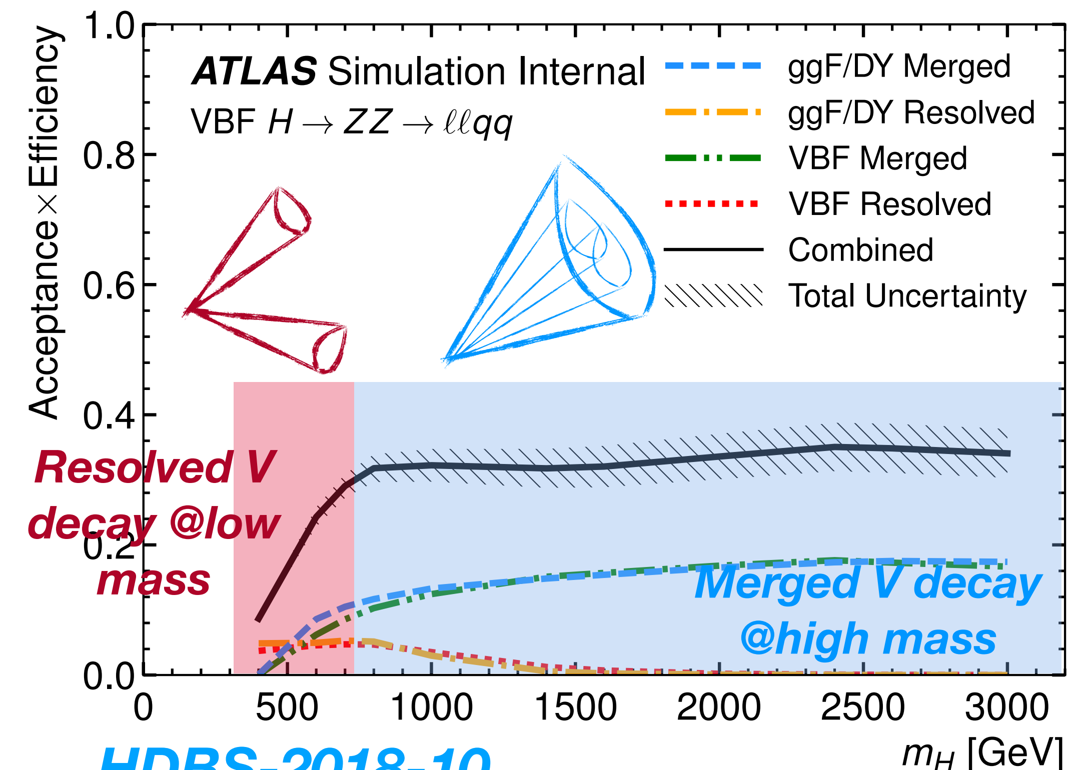
Resonant VV search: semi-leptonic

Search for new physics decaying into diboson in semi-leptonic final states.

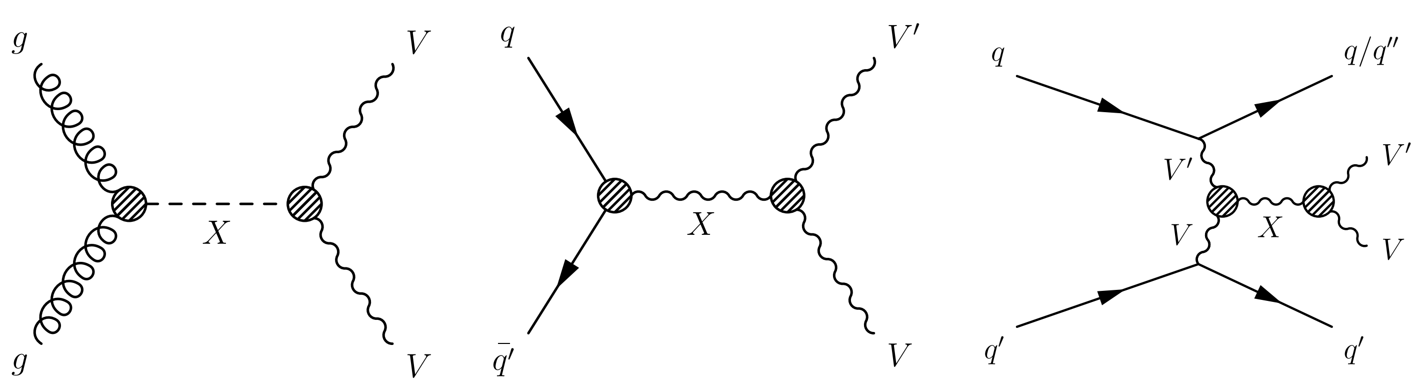
- 3 leptons channel going to be combined:
  - ZV —> vvqq, WV —> lvqq, ZV —> llqq
  - good compromise between the low BR of the full leptonic and the more background of the full hadronic channels.
- 2 categories based on the production mechanism are studied:
  - VBF vs ggF/DY
- 2 different techniques to reconstruct the V—>qq decay dedicated to lower (***resolved***) and higher (***merged***) mass regions.
- Interpretations
  - Spin-0: Randall Sundrum (RS) Radions
  - Spin-1: HeavyVectorTriplet, W', Z'
  - Spin-2: Randall Sundrum (RS) Gravitons

EXOT-2017-31



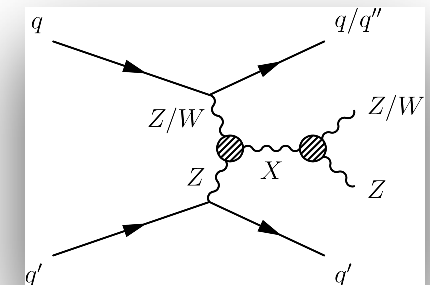


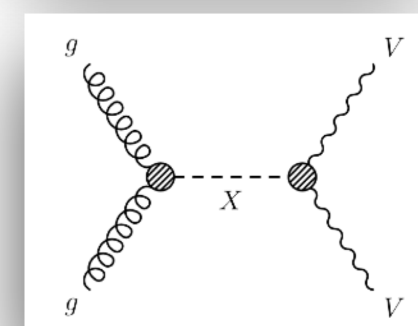
HDBS-2018-10



Resonant VV analysis in a nutshell

Machine Learning approach here





Main analyser in this analysis (HDBS-2018-10) during my PhD:

- trigger studies (TRIG-2018-05)
- analysis optimisation
- MC modelling
- ML approaches

Z —> νν
W —> lν
Z —> ll

Leptonic V selection

Binary classification problem performed with cut based approach in previous paper, R&D to development of new techniques

- June 2019, HDBS & Exotics Workshop
- November 2019, 4th ATLAS ML Workshop
- CDS. 2743418 thesis

VBF/ggF classification

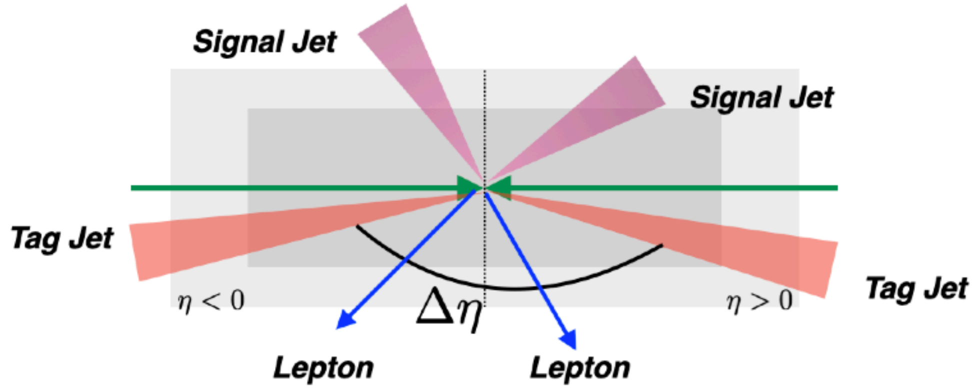
Merged Selection → **Resolved Selection** → **Control Regions**

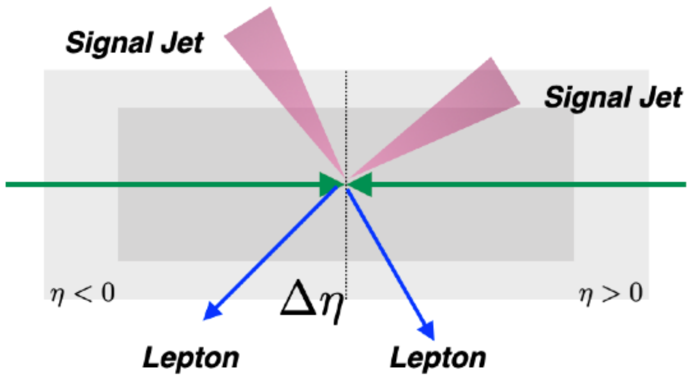
Merged Selection → **Resolved Selection** → **Control Regions**

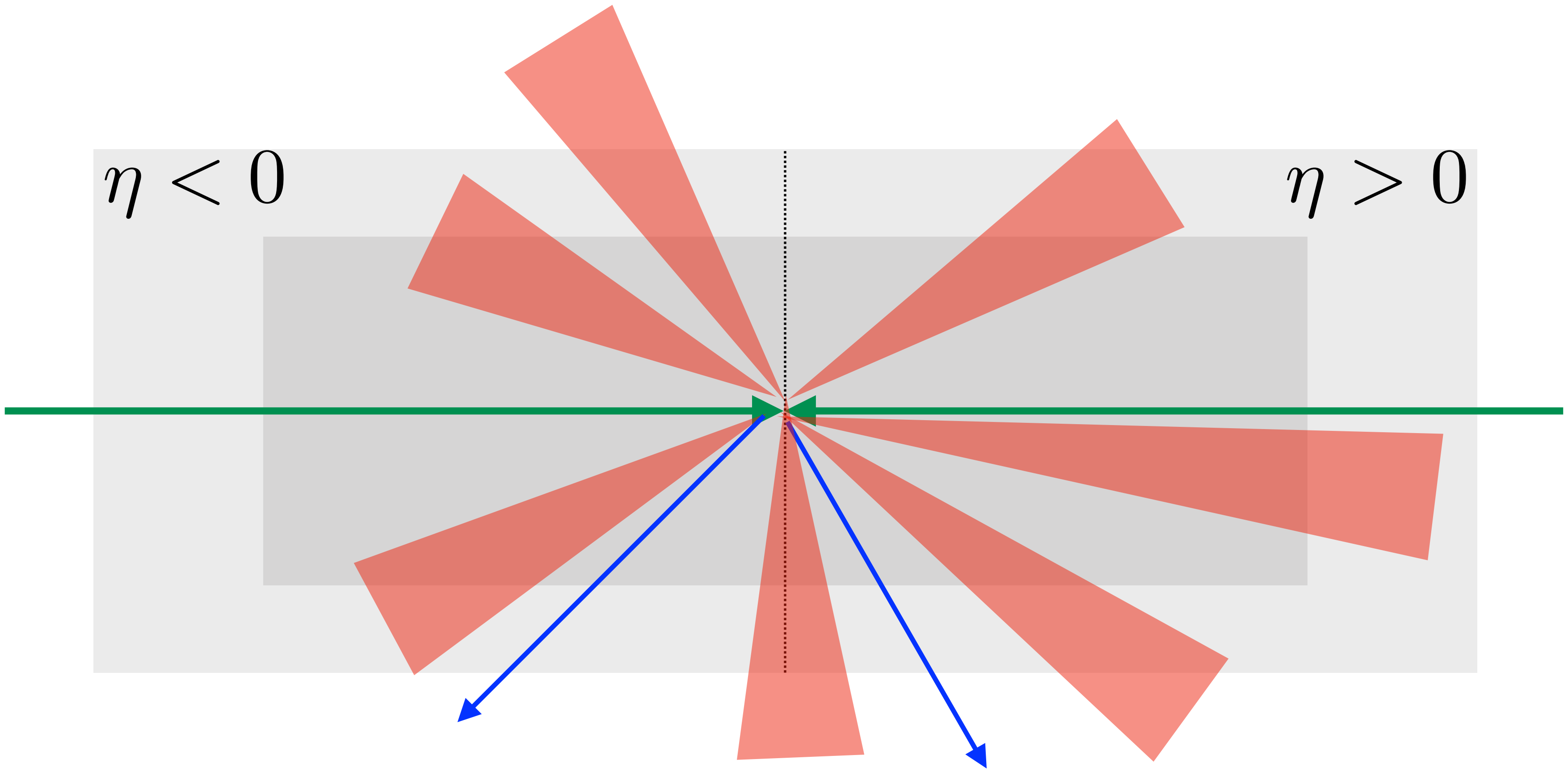
VBF vs ggF topologies

How does the detector see these events?

Feynman LO diagrams
- + >LO corrections,
- + PileUp interactions,
- + underlying interactions





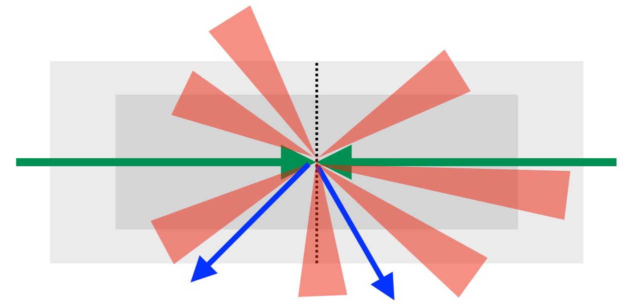


VBF or ggF?

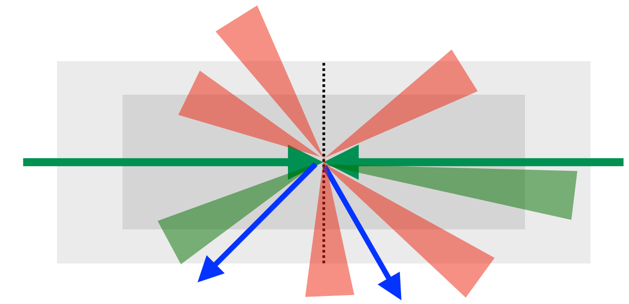
NN approach: DL & input variables

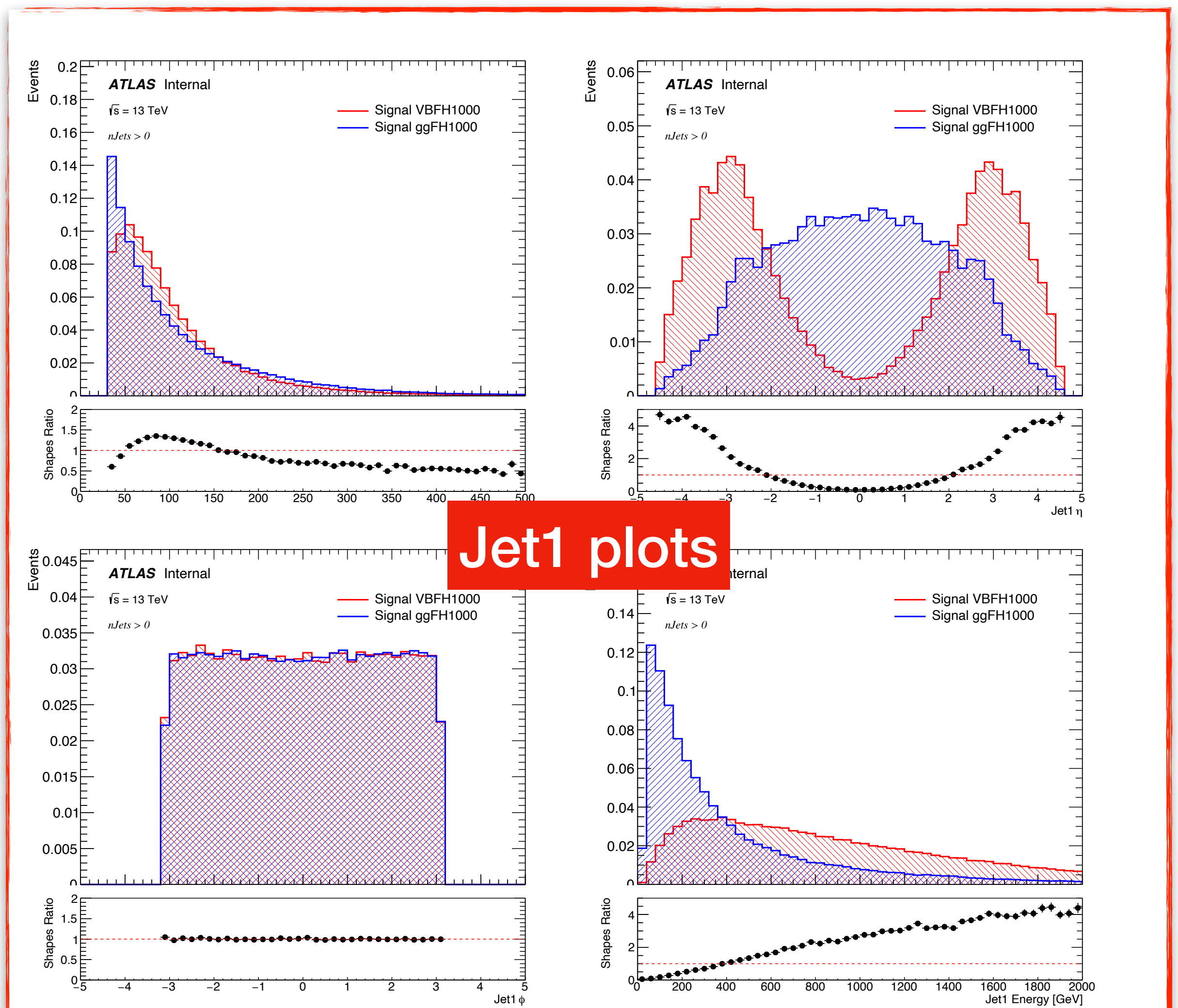
- ***High-level approach:***
    - ***build high-level variables starting from the two tag jets***
- ***Low-level approach:***
    - ***use all the jets and the 4-momentum components***

All jets used for low-level approach



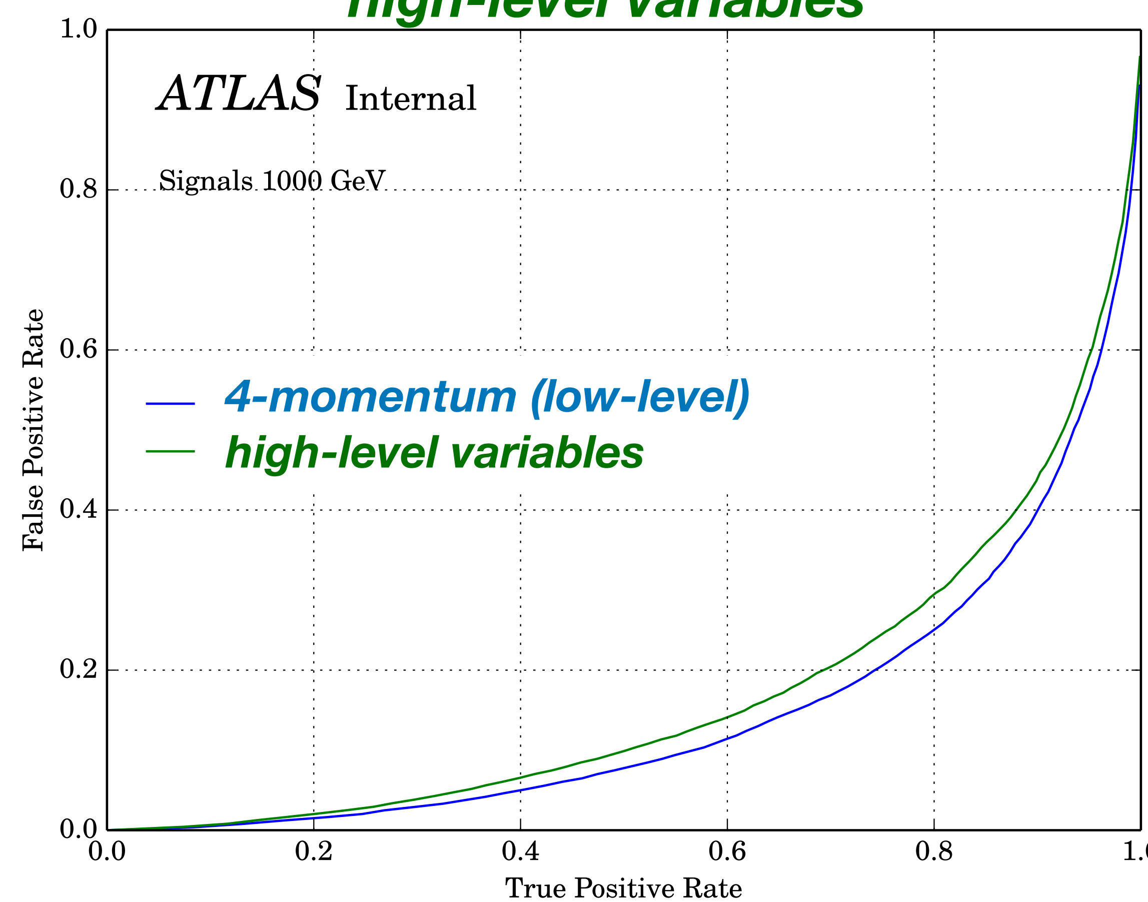
2 jets tagged for the high-level approach





xN jets

4-momentum (low-level) better than high-level variables

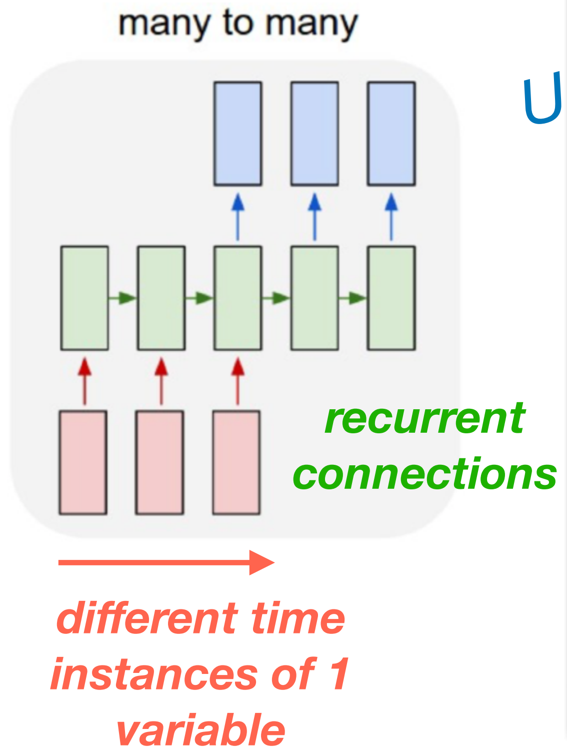


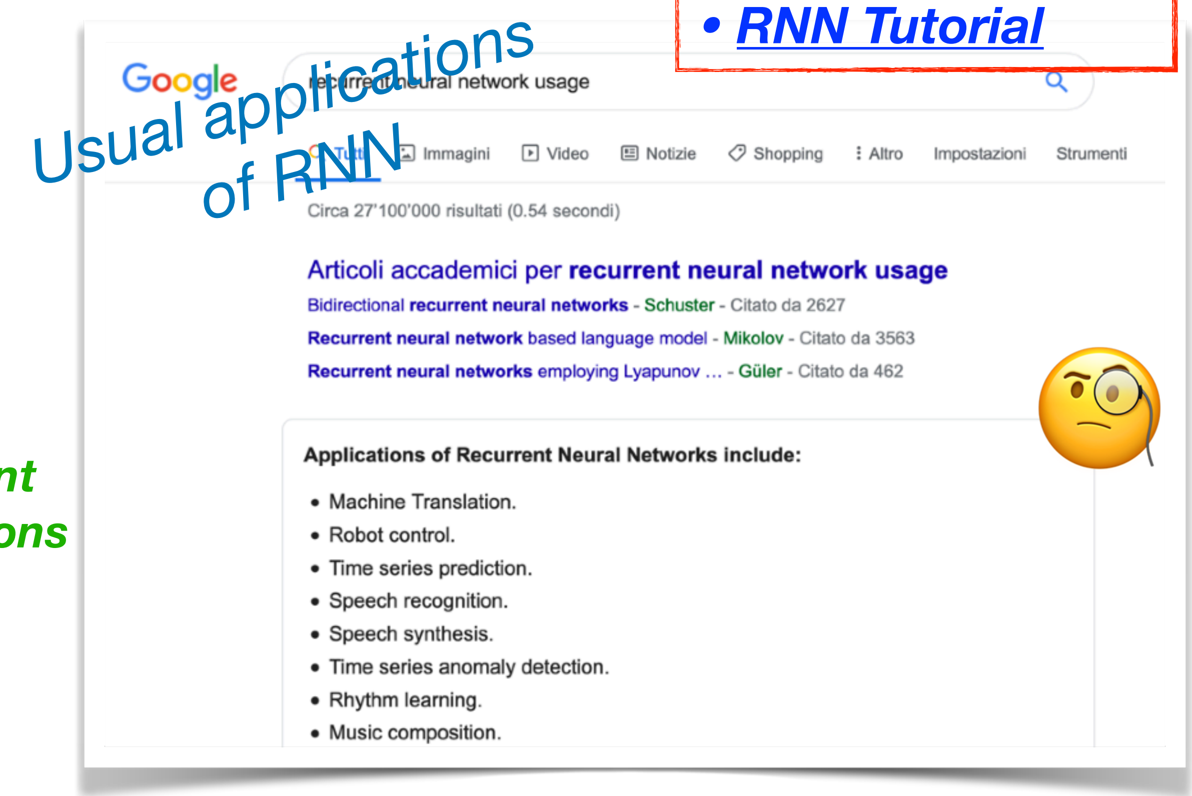
General deep learning result, arXiv:1402.4735

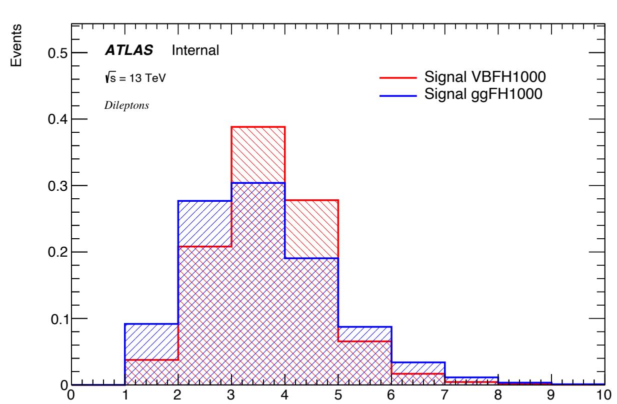
The VBF-RNN architecture

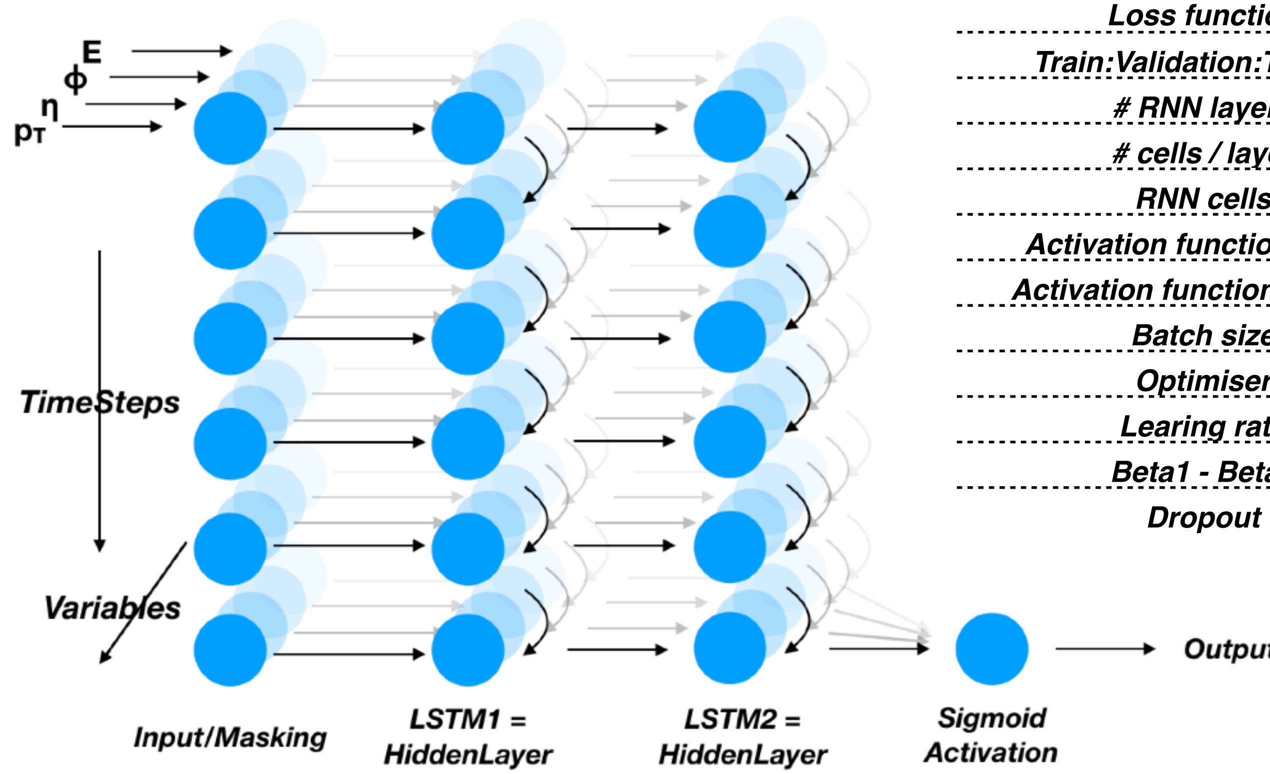
Look at more info:
- *RNN Lecture*
- *RNN paper*
- *RNN Tutorial*

- Main idea: deal with ***variable-length input set*** (1, 2, 3… N jets).
- Which is the best NN architecture?
  - **RNN seem to fit for our problem!**
- An input set (the 4 variables of the 4-momentum) is repeated at different "Time-steps" that could be variables (our number of jets).
- 4 variables used (p_T, η, ϕ, E).







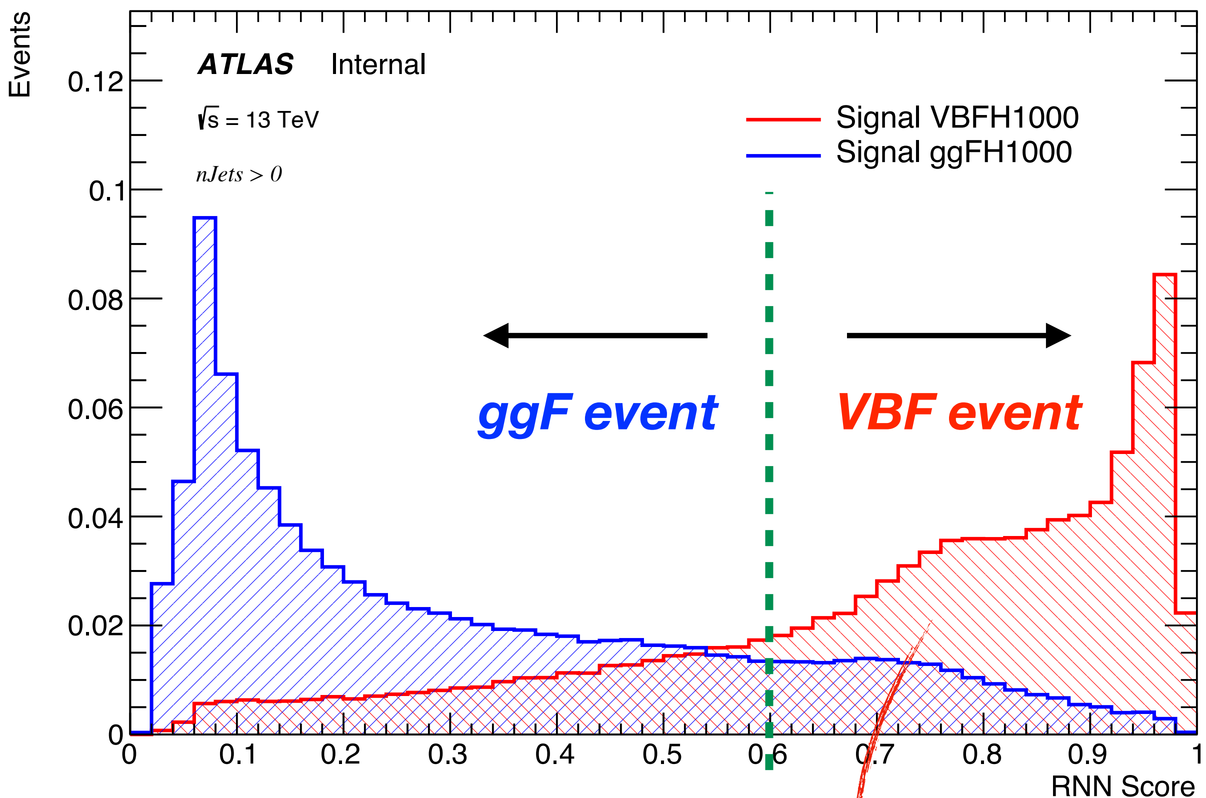


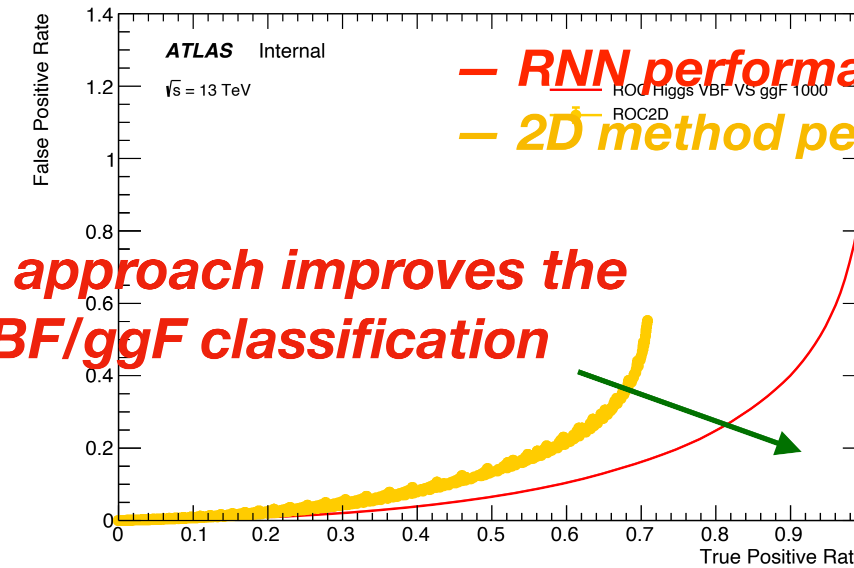
Loss function	binary cross entropy
Train:Validation:Test (%)	56:14:30
# RNN layers	2
# cells / layer	25
RNN cells	LSTM
Activation function (RNN)	tanh
Activation function (output)	sigmoid
Batch size	512
Optimiser	Adam
Learing rate	0,001
Beta1 - Beta2	0,9 - 0,999
Dropout	0,3

Keras

Theano

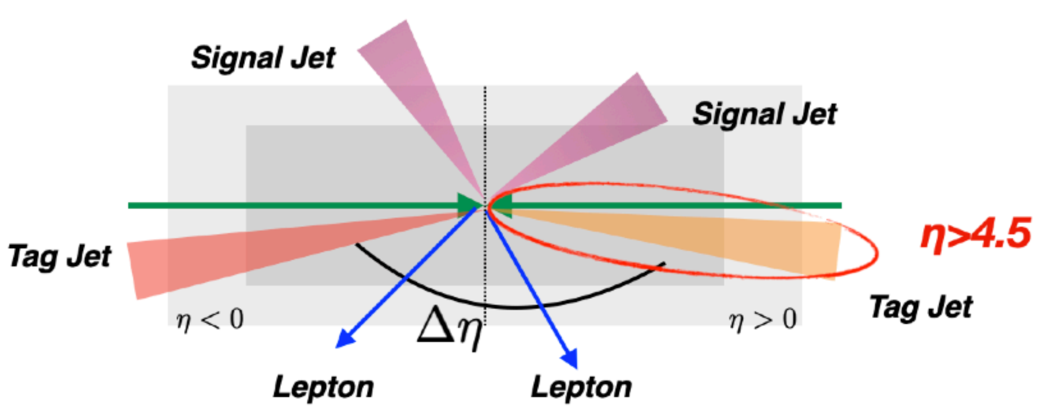
RNN approach: performances

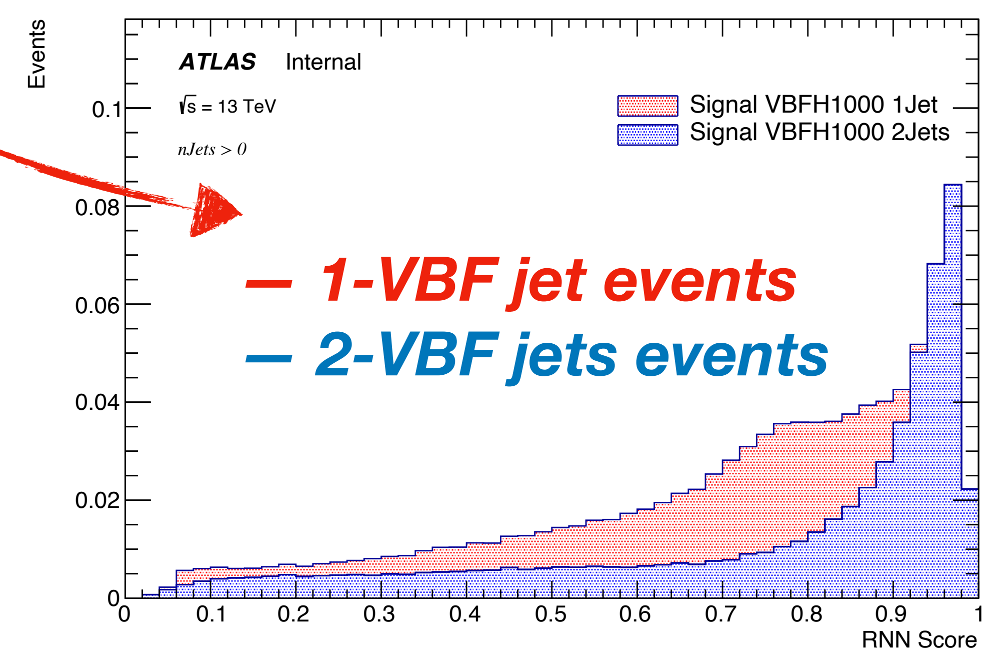




	Std Categ	NNScore > 0.6	NNScore > 0.7	NNScore > 0.8
VBF H 1000	37,8%	67,2%	57,6%	42,7%
ggF H 1000	6,1%	14,8%	9,6%	4,6%

What happens if 1 VBF jet is outside η-acceptance?

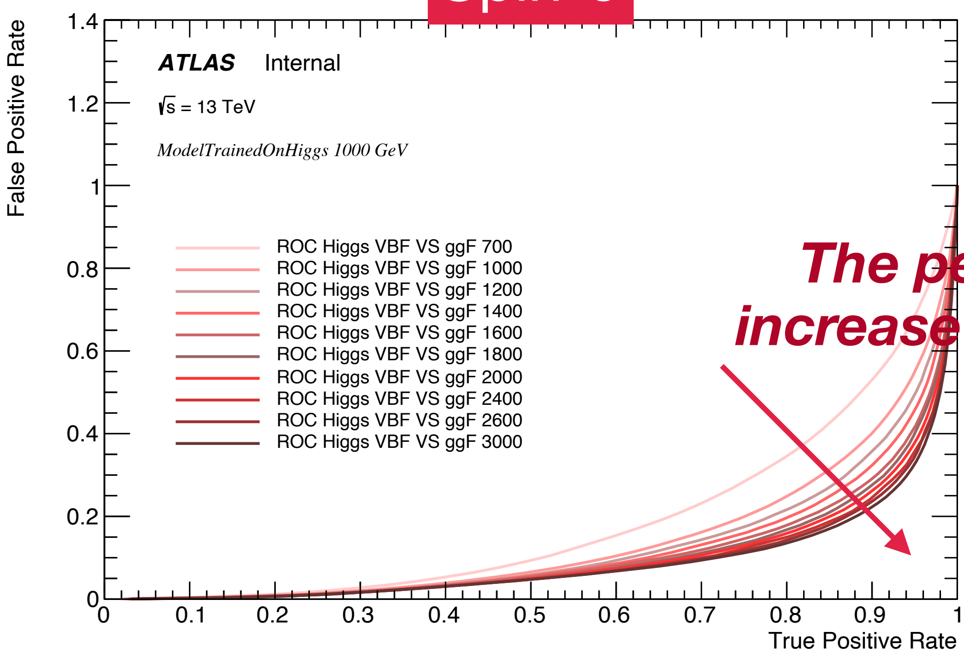




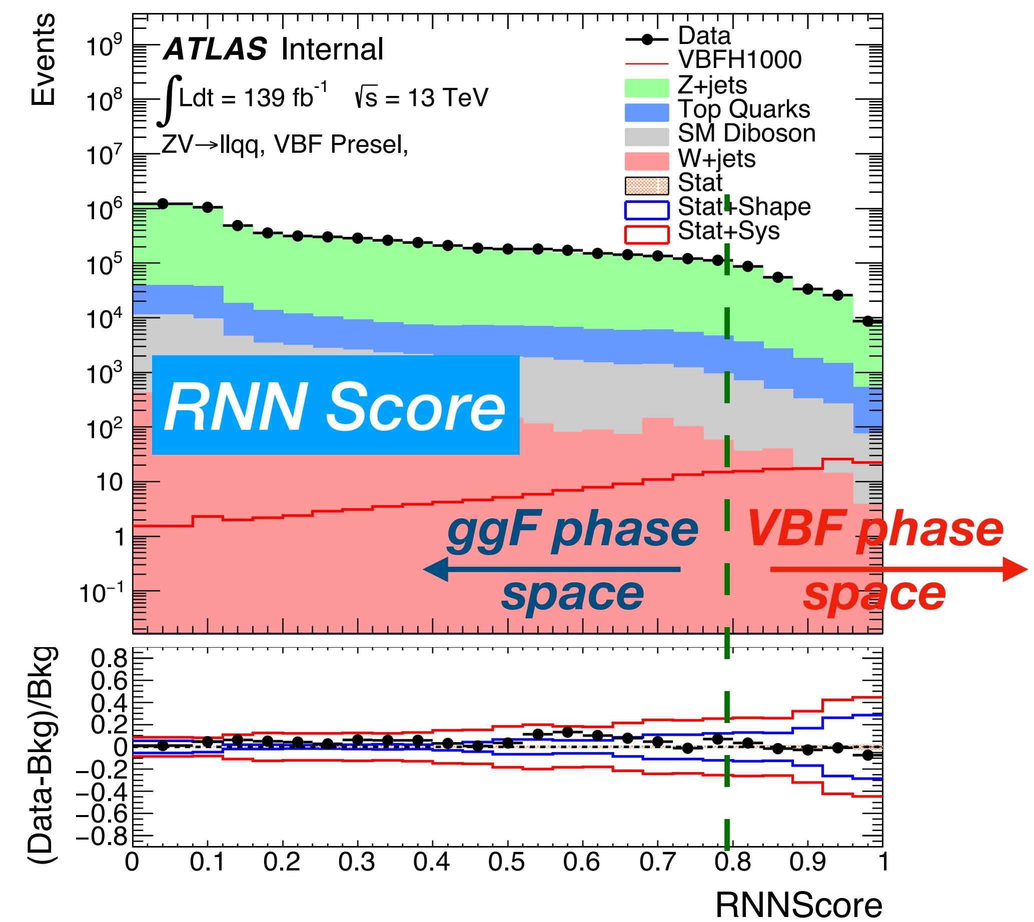
The RNN is able to classify also events with 1 VBF jet that have never used with cut-based approach

RNN: masses, spins, leptons channel, data...

Spin-0



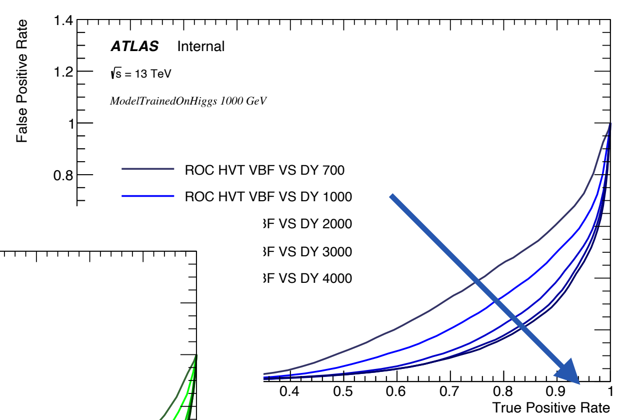
The performances increase with the mass



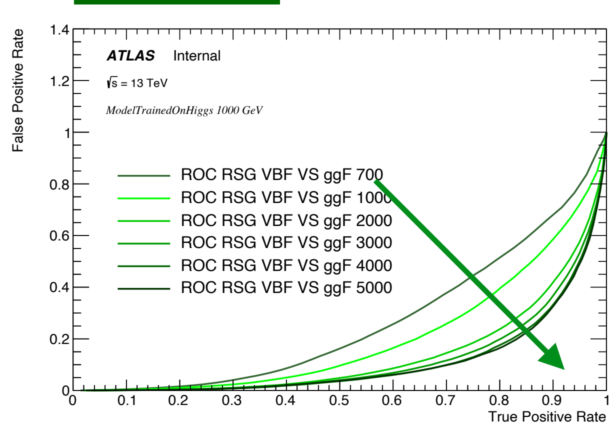
RNN Score

ggF phase space ← → ***VBF phase space***

Spin-1



Spin-2



Different spin models shows consistent behaviour with respect to the spin-0 RNN model

Stable performances for different leptons channels

0/1/2 leptons