ATLAS Boosted Higgs

LHC Higgs Working group

Andrea Sciandra(UC Santa Cruz), Zhi Zheng (SLAC)

15/11/2023

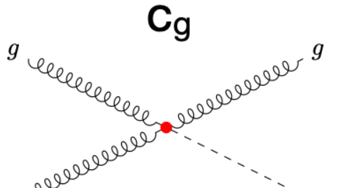
Why Higgs in Boost region

Increased impact expected from new physics

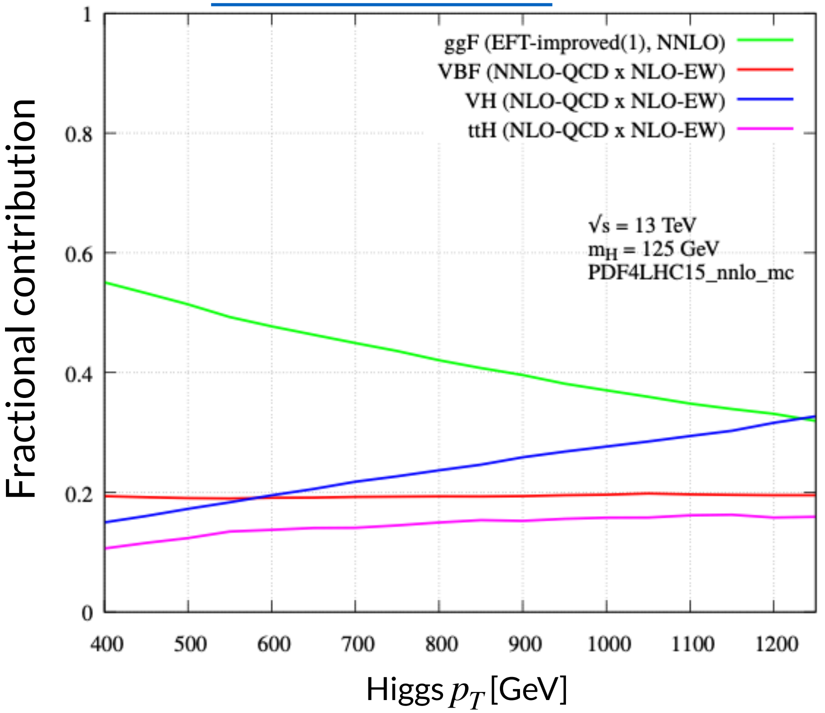
- Probe beyond standard model (BSM)

Increased interests in understanding dynamic properties of the Higgs

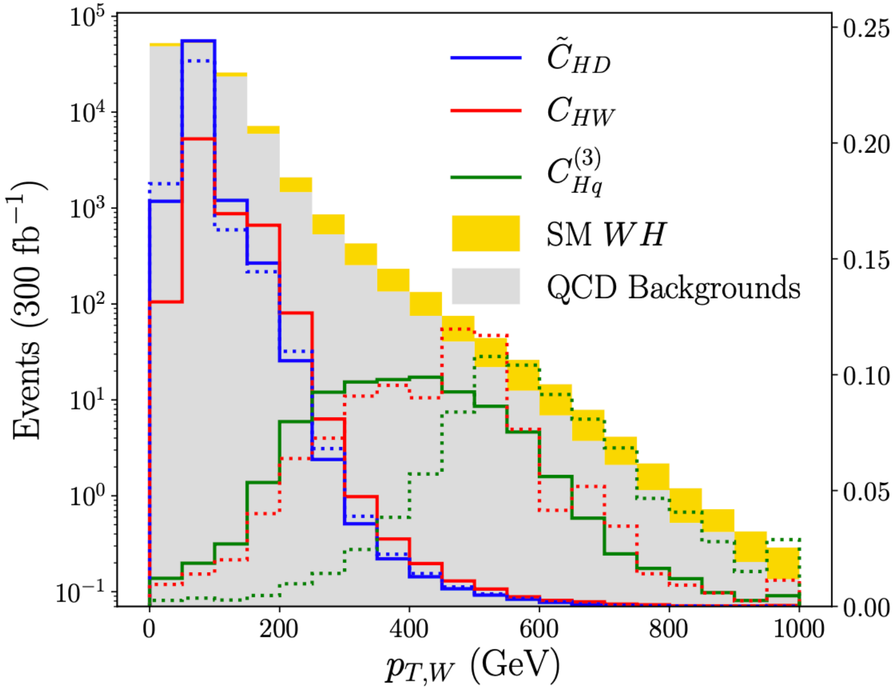
- All production modes contribution similarly toward p_T^H ~ 1 TeV



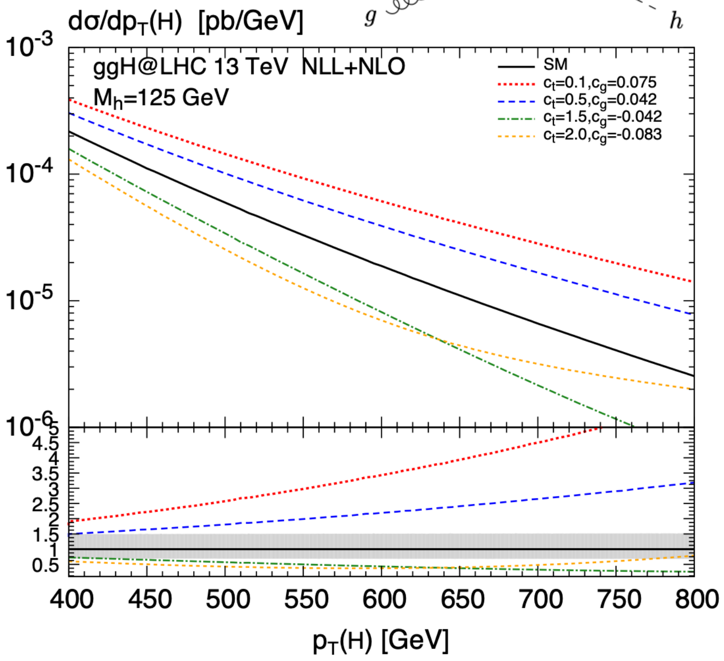
arXiv:2005.07762



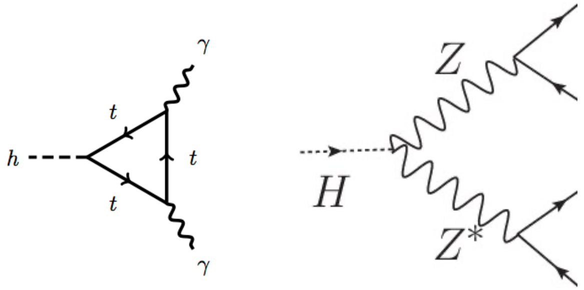
JHEP11(2019)034



JHEP03(2017)115



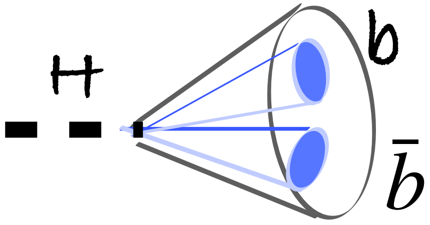
Boosted Higgs @ ATLAS



Clean signal

Measurement of the total and differential Higgs boson production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector by combining the $H \rightarrow ZZ$ and $H \rightarrow \gamma\gamma$ decay channels

[JHEP 05 (2023) 028]



Large Branching fraction

Constraints on Higgs boson production with large transverse momentum using $H \rightarrow b\bar{b}$ decays in the ATLAS detector

[Phys. Rev. D 105 (2022) 092003]

Measurement of high-momentum Higgs boson production in association with a vector boson in the qqbb final state with the ATLAS detector

[ATLAS-CONF-2023-067]

Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$TeV with the ATLAS detector

[Phys. Lett. B 816 (2021) 136204]

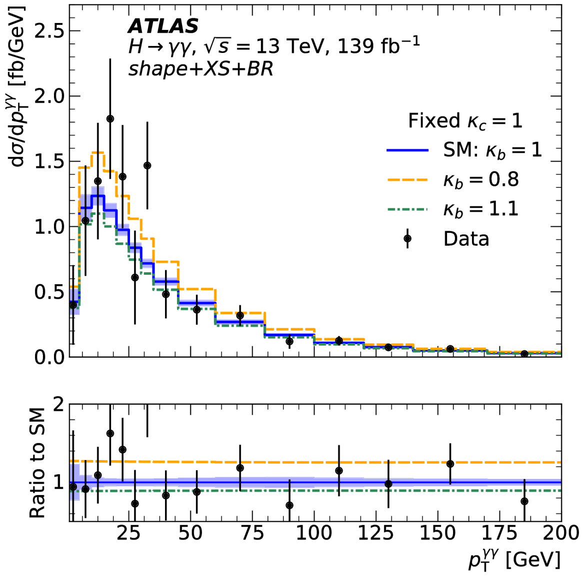
H→ ZZ and H→ $\gamma\gamma$ combination

Phys. Lett. B 816 (2021) 136204

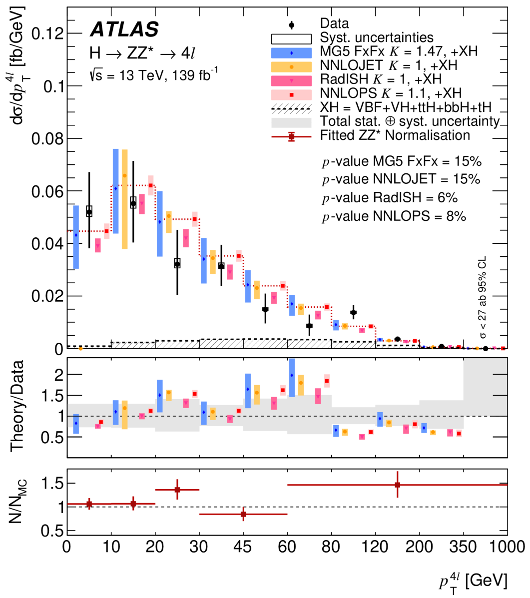
Combine H→ ZZ (JHEP 08 (2022) 027) and H→ $\gamma\gamma$ (Eur. Phys. J. C 80 (2020) 942) results

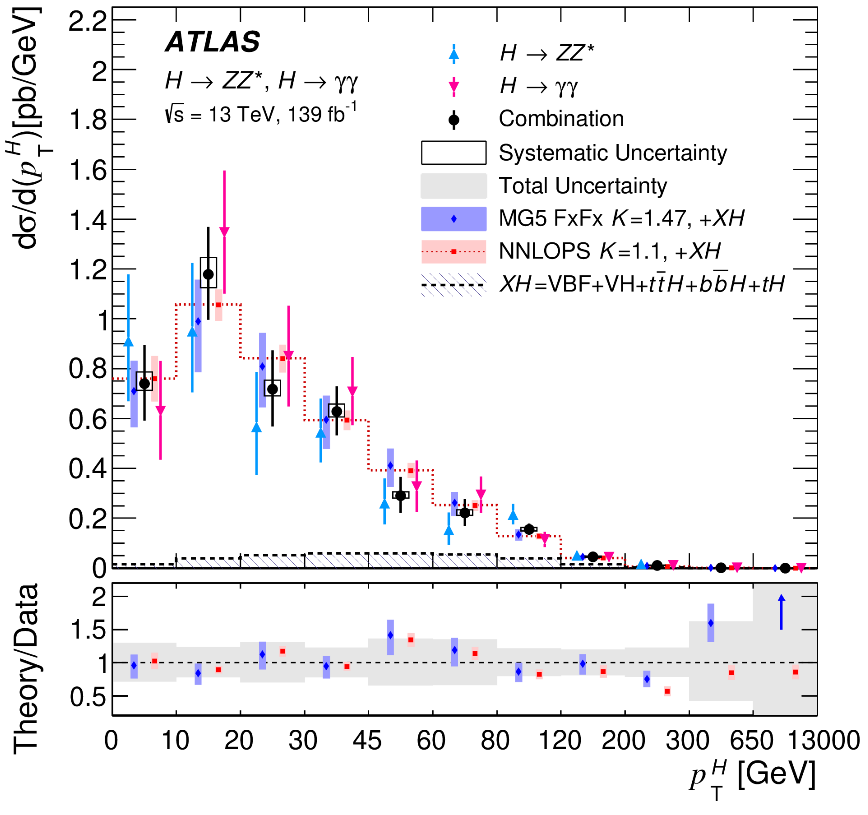
Compare with results from individual channels, total uncertainty is lowered by 20%-40%

JHEP 08 (2022) 027



Eur. Phys. J. C 80 (2020) 942





H→ ZZ and H→ $\gamma\gamma$ combination

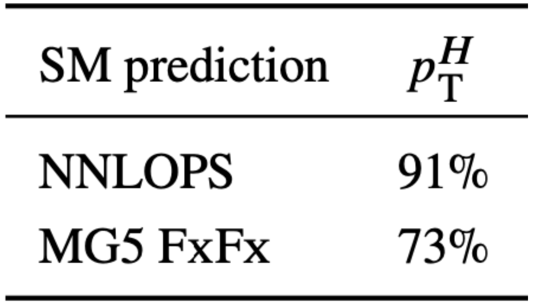
Phys. Lett. B 816 (2021) 136204

Combine H→ ZZ (JHEP 08 (2022) 027) and H→ $\gamma\gamma$ (Eur. Phys. J. C 80 (2020) 942) results

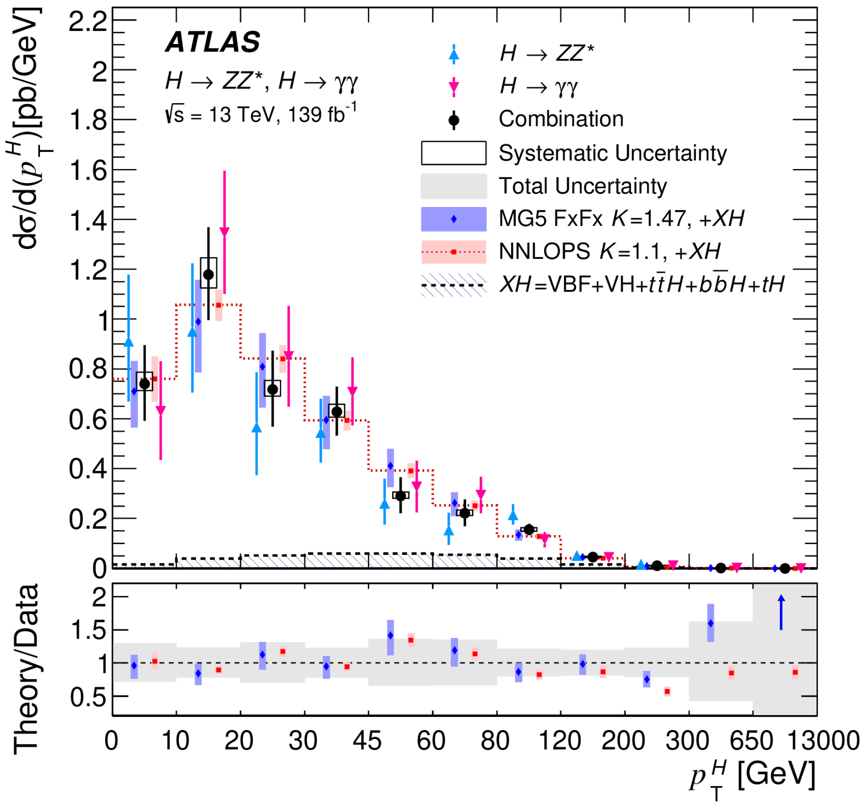
Compare with results from individual channels, total uncertainty is lowered by 20%-40%

Limited by Statistics in both channels

The prediction based on NNLOPS simulation of gluon-gluon fusion events is lightly favored over the MG5 FxFx simulation

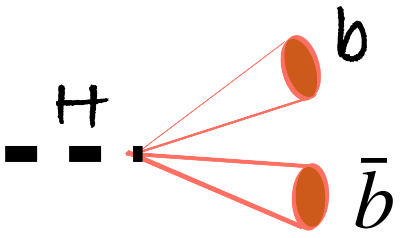


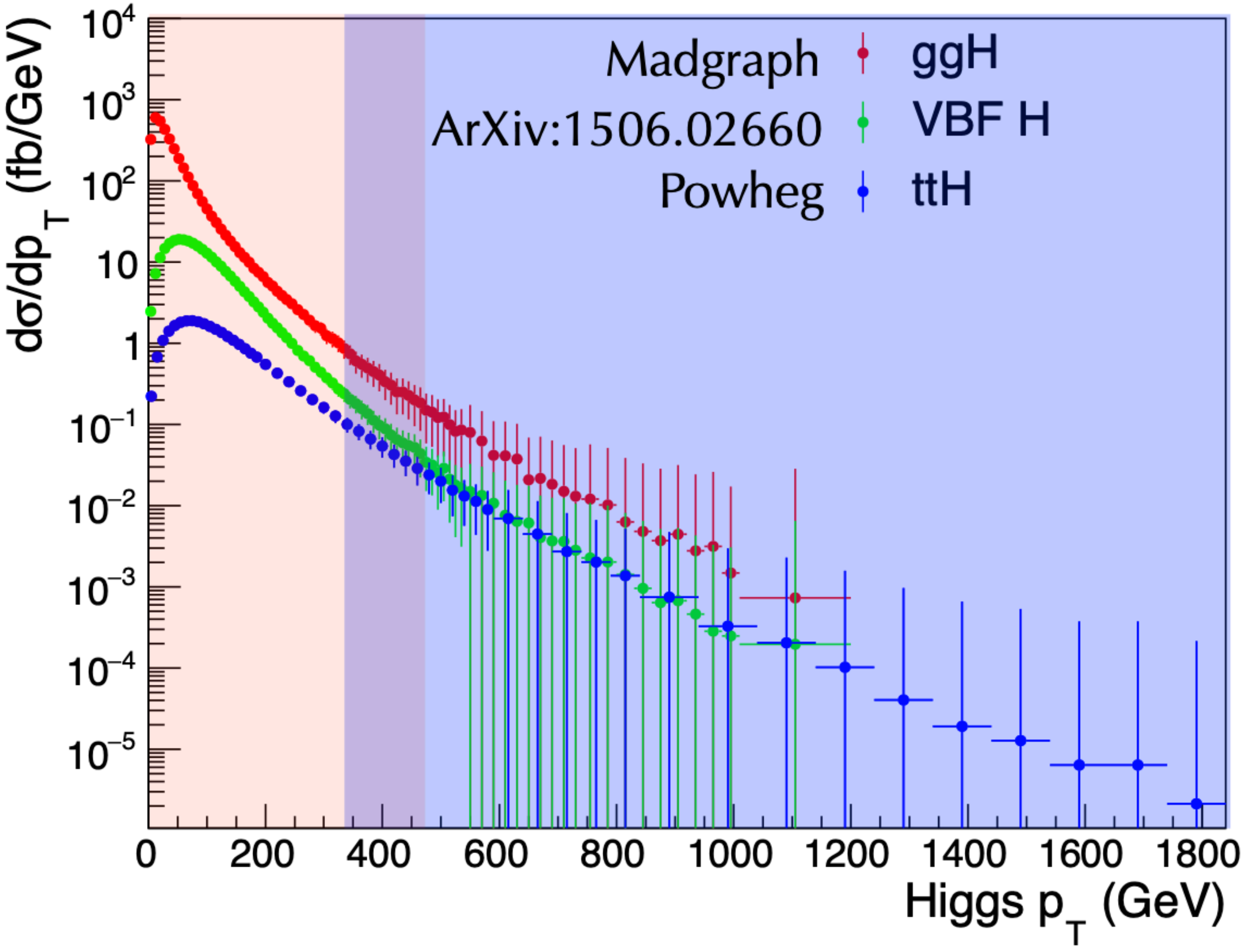
SM prediction	p_T^H
NNLOPS	91%
MG5 FxFx	73%

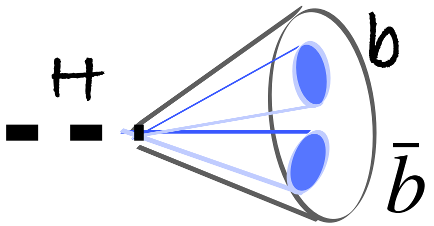


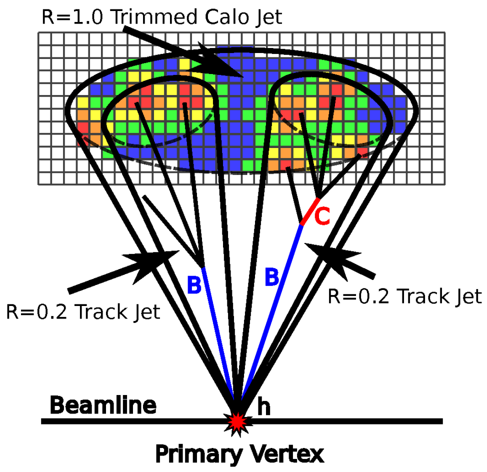
All-Hadronic Higgs in Boost region

High p_T Higgs decay product clustered inside large-redius calorimeter jet









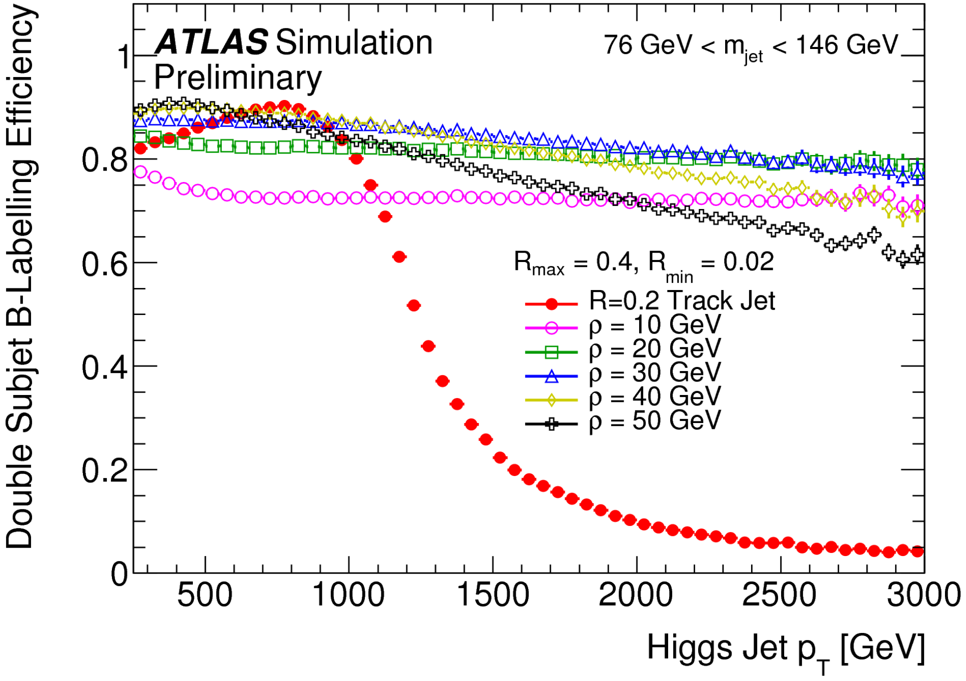
All-Hadronic Higgs in Boost region

Advancement of **novel jets substructure** enabled searches for $H \to bb$ in hadronic final states despite the large irreducible QCD background

Double-b-tagging

Used in Boost all had Higgs

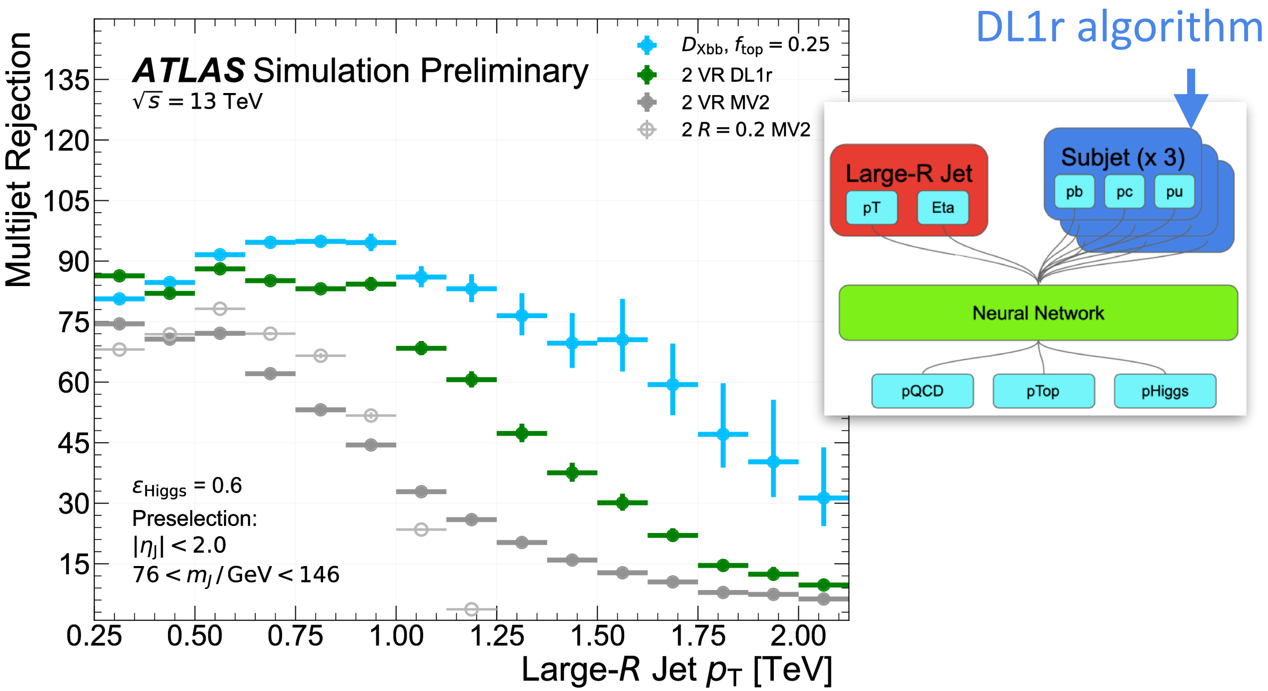
Standard b-tagging algorithm for VR track jets large-R jets with 2b-tagged VR track jets



Neural networks based tagging

Used in VH → bbqq

Training NN with DL1r out pt for each track jets



Boosted All-Had H→ bb

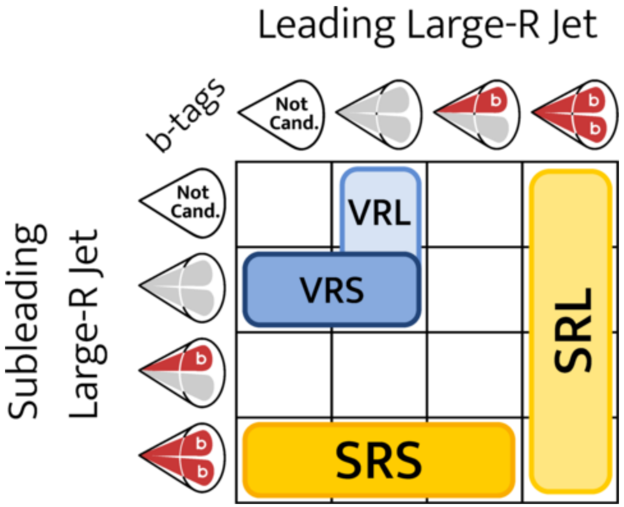
Phys. Rev. D 105 (2022) 092003

Large-R jet trigger, $p_T > 450$ GeV, $m > 60$ GeV

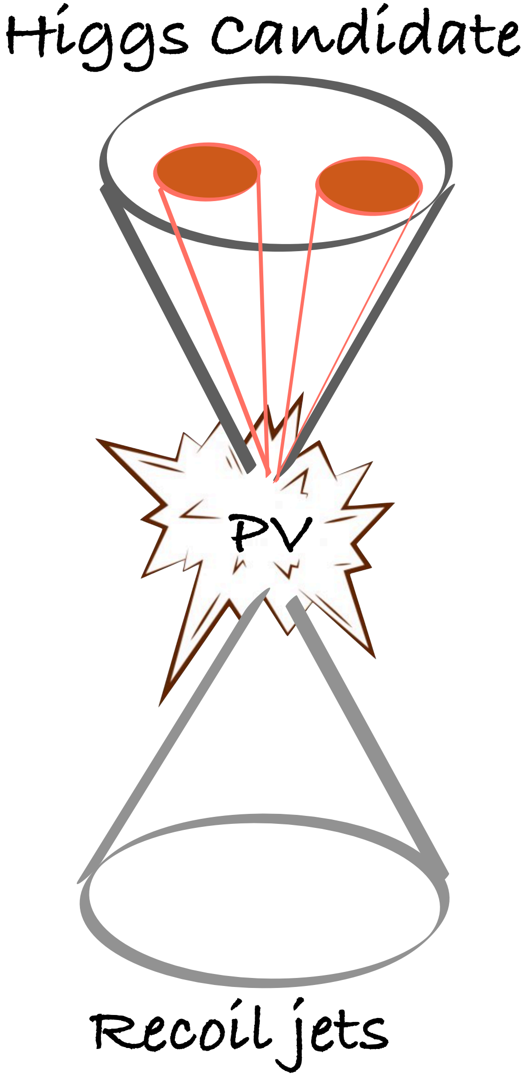
At least 1 additional jet, $p_T > 200$ GeV

Higgs Candidate jets:

- $p_T > 250$ GeV, $m > 60$ GeV, $\eta < 2$
- Boosted: 2m/p_T < 1
- 2 VR track jets with MV2 b-tagging algorithm at 77% WP

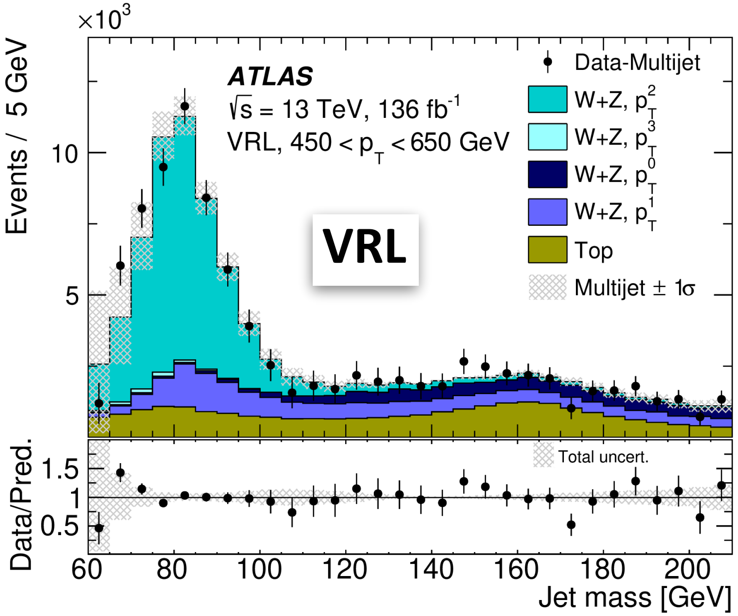


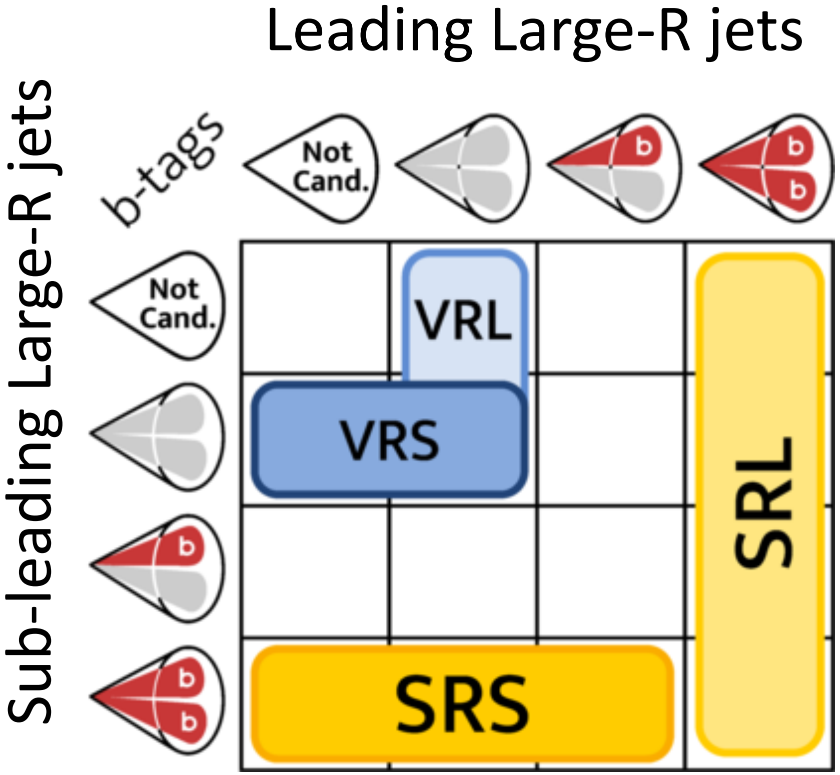
Region	Candidate jet p_T [GeV] SRL	SRS
Inclusive	>450	>250
Fiducial	>450	>450
Differential	450–650, 650–1000, > 1000	250–450, 450–650, 650–1000

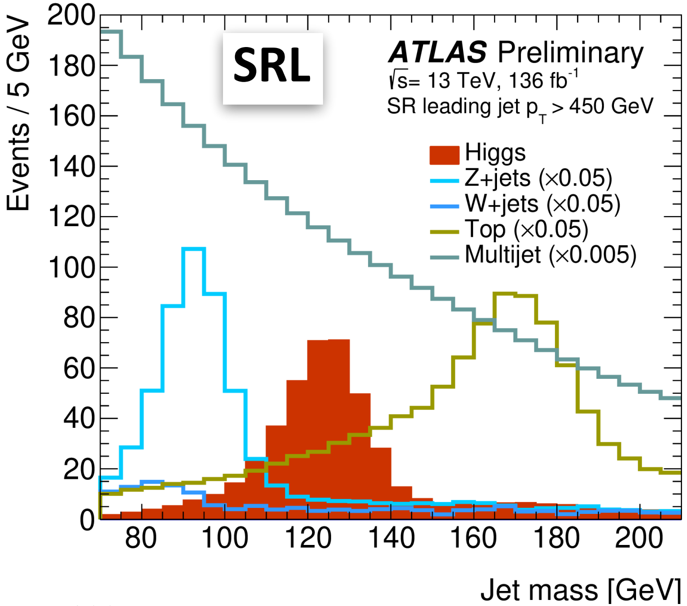


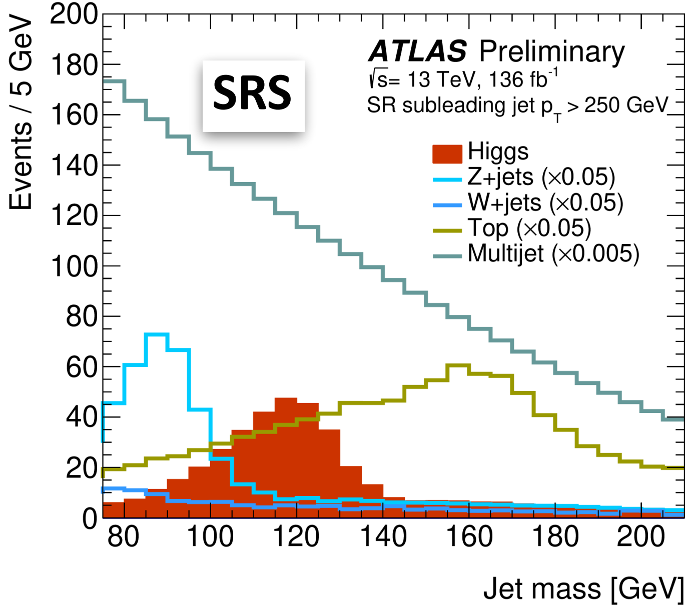
Boosted All-Had H→ bb

Phys. Rev. D 105 (2022) 092003









Boosted All-Had H→ bb: Composition

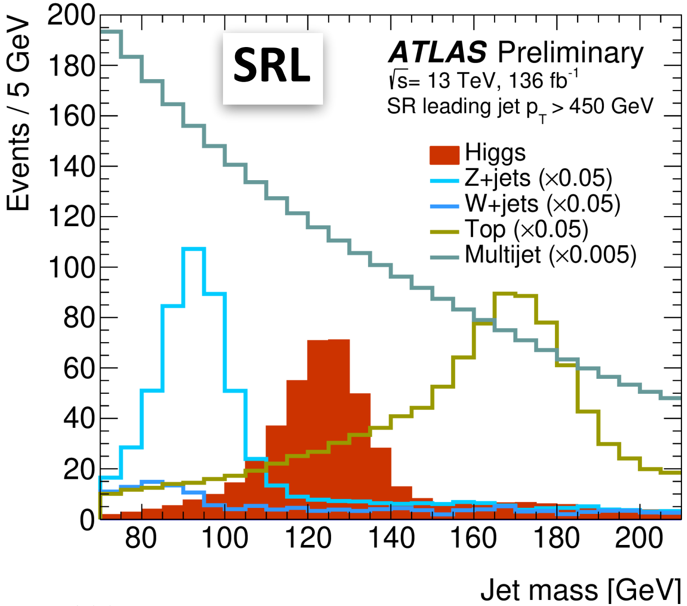
[Phys. Rev. D 105 (2022) 092003]

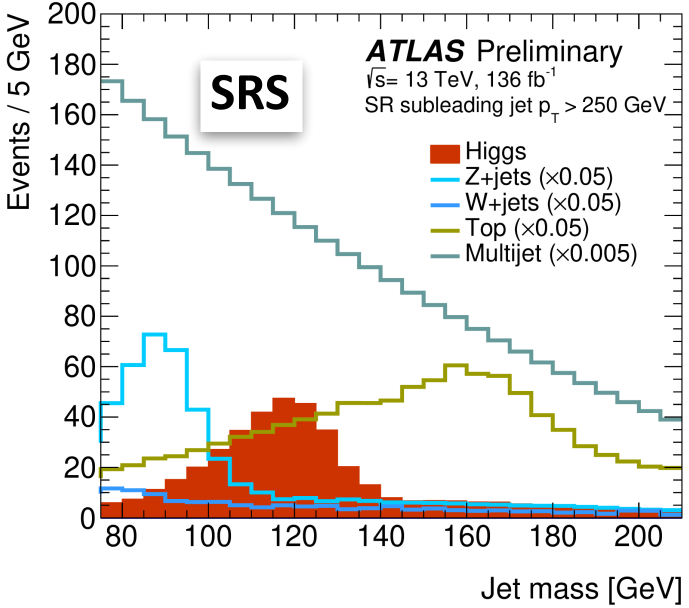
Higgs signal: ggF, VBF, VH, ttH

- Corrections for NLO electroweak effects applied to VBF, VH and ttH
- ggF, VBF, VH cross section (Xsec) from MC compatible with [LHCXSWG], ttH Xsec scale to LHCXSWG

Process	Jet p_T Range [GeV] 250–450	450–650	650–1000	> 1000
		SRL		
ggF	–	0.56	0.50	0.39
VBF	–	0.17	0.16	0.17
VH	–	0.14	0.18	0.25
$t\bar{t}H$	–	0.13	0.16	0.19
		SRS		
ggF	0.28	0.46	0.43	–
VBF	0.07	0.19	0.21	–
VH	0.26	0.24	0.26	–
$t\bar{t}H$	0.39	0.11	0.10	–

ggF dominant





Boosted All-Had H→ bb: Results

Phys. Rev. D 105 (2022) 092003

Signal extraction: fit to jet mass distribution of m_{bb} with SRL, SRS and $CR_{t\bar{t}}$

Result	μ_H	μ_Z	$\mu_{t\bar{t}}$
Expected	1.0 ± 3.2	1.00 ± 0.17	1.00 ± 0.07
Observed	0.8 ± 3.2	1.29 ± 0.22	0.80 ± 0.06

Fiducial region ($p_T^H > 450$ GeV, $y_H < 2$): $\sigma_H < 115(128)$ fb

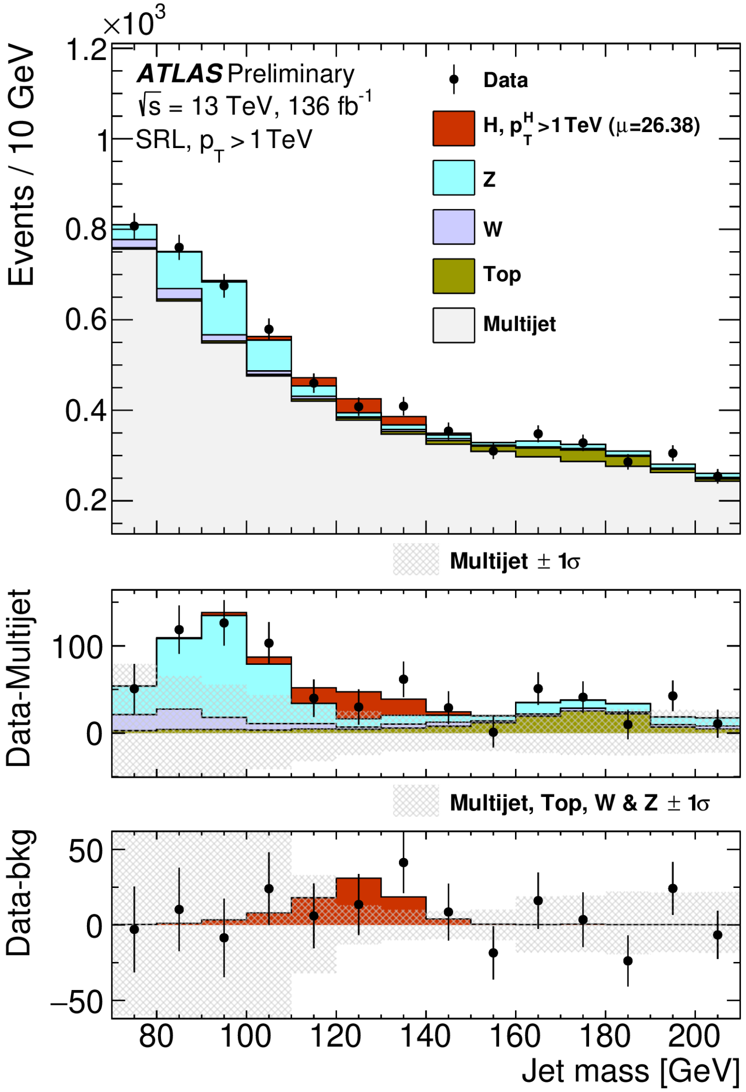
- SM prediction: 18.4 fb

Uncertainty Contribution	$p_{\mathrm{T}}^H > 450$ GeV
Total	3.5
Statistical	2.6
Systematic	2.3
Jet systematic uncertainties	2.2
Modeling and theory systs.	0.8
Flavor-tagging systs.	0.2

Limited by **data statistics**

Leading source of systematic uncertainty:

- Jet mass resolution and mass scale



Boosted All-Had H→ bb: Results—Differential

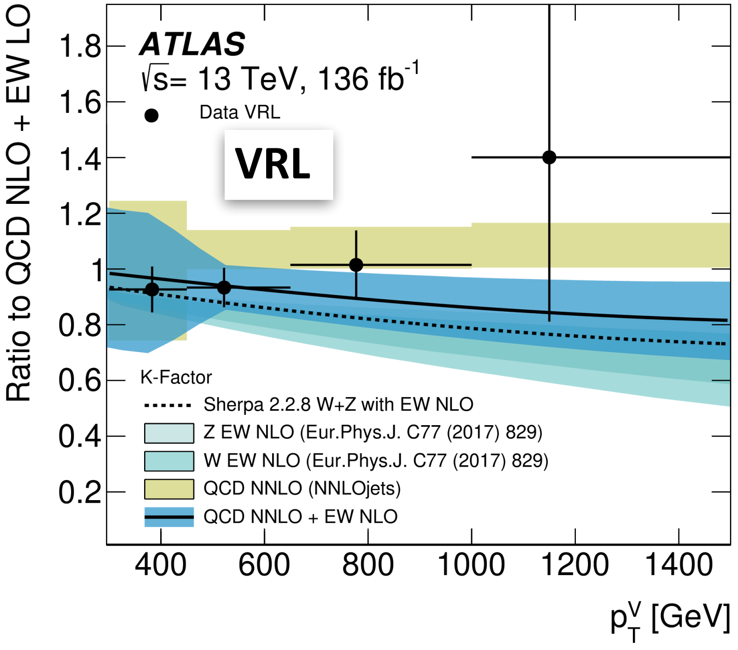
Phys. Rev. D 105 (2022) 092003

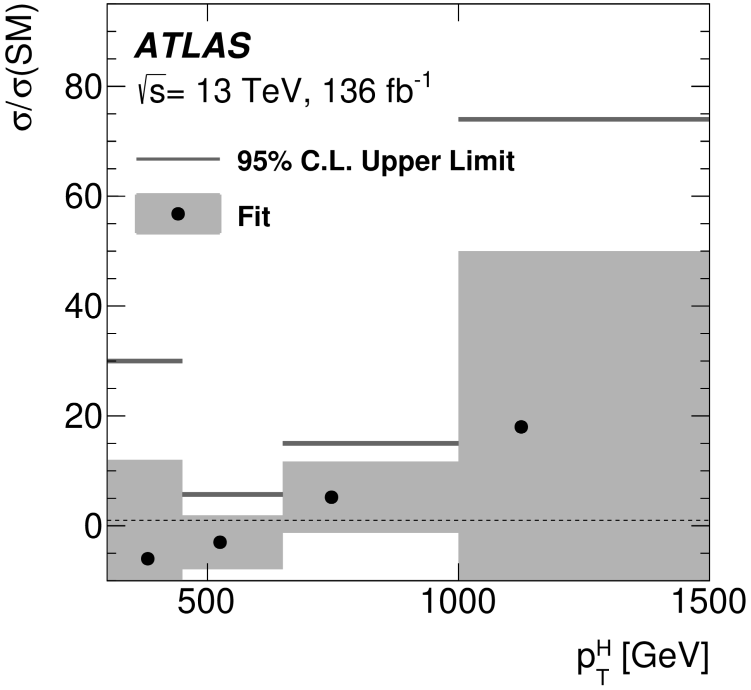
Signal extraction: fit to jet mass distribution of m_{bb} with SRL, SRS and $CR_{t\bar{t}}$

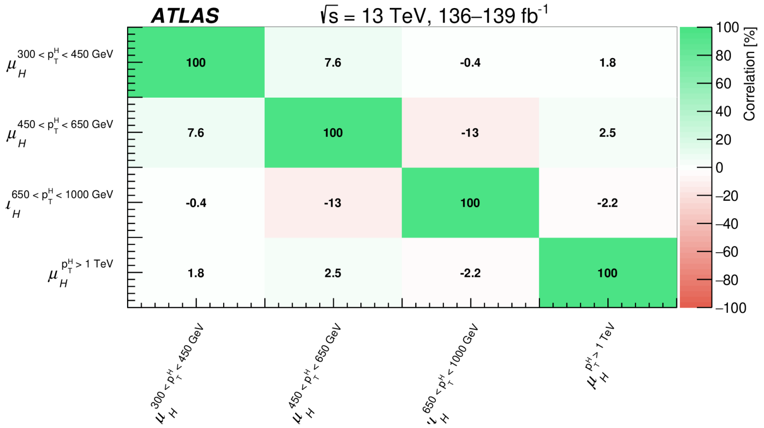
Observed $\sigma(p_T^H > 1\ \mathrm{TeV})$ =2.3 ± 3.9(stat.)± 1.3(syst.)± 0.5 (the.)fb

- SM prediction: 0.13 fb

Volume	p_{T}^{H} [GeV]	$\lvert y_H \rvert$
Fiducial	>450	< 2
STXS	300–450, 450–650, 650–1000, >1000	< 2







	$\mu_H^{300 < p_T^H < 450\ \mathrm{GeV}}$	$\mu_H^{450 < p_T^H < 650\ \mathrm{GeV}}$	$\mu_H^{650 < p_T^H < 1000\ \mathrm{GeV}}$	$\mu_H^{p_T^H > 1\ \mathrm{TeV}}$
$\mu_H^{300 < p_T^H < 450\ \mathrm{GeV}}$	100	7.6	-0.4	1.8
$\mu_H^{450 < p_T^H < 650\ \mathrm{GeV}}$	7.6	100	-13	2.5
$\mu_H^{650 < p_T^H < 1000\ \mathrm{GeV}}$	-0.4	-13	100	-2.2
$\mu_H^{p_T^H > 1\ \mathrm{TeV}}$	1.8	2.5	-2.2	100

Boosted All-Had H→ bb: Results—Differential

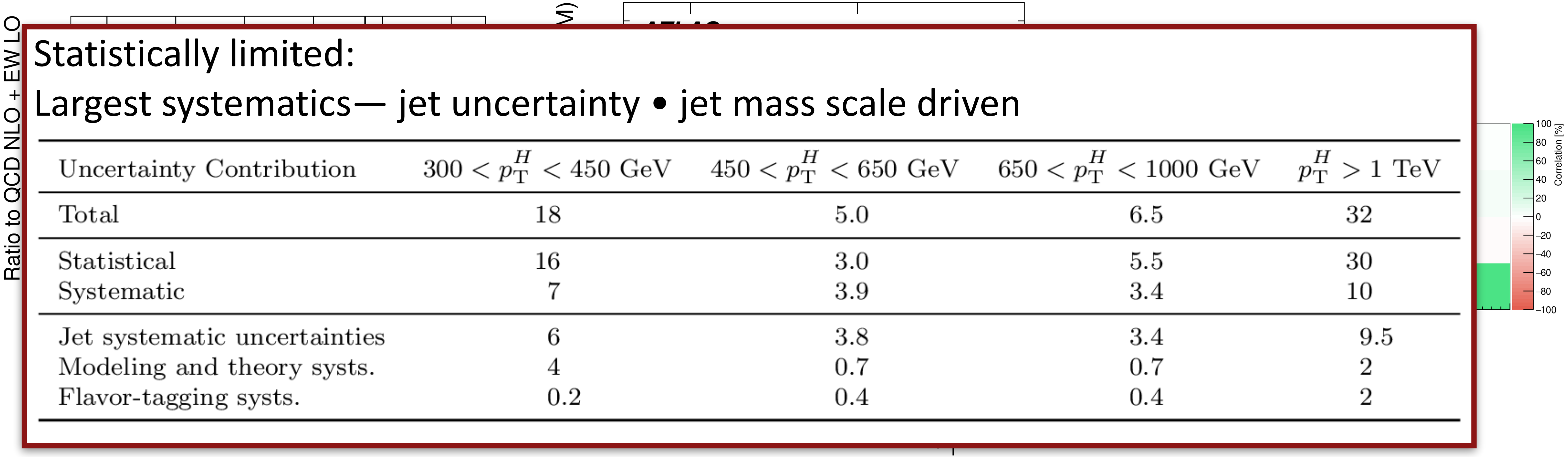
Phys. Rev. D 105 (2022) 092003

Signal extraction: fit to jet mass distribution of m_{bb} with SRL, SRS and $CR_{t\bar{t}}$

Observed $\sigma(p_T^H > 1\ \text{TeV})$ =2.3 ± 3.9(stat.)±1.3(syst.)±0.5 (the.)fb

- SM prediction: 0.13 fb

Volume	p_T^H [GeV]	$\lvert y_H \rvert$
Fiducial	>450	< 2
STXS	300–450, 450–650, 650–1000, >1000	< 2



Statistically limited:

Largest systematics— jet uncertainty • jet mass scale driven

Uncertainty Contribution	$300 < p_T^H < 450$ GeV	$450 < p_T^H < 650$ GeV	$650 < p_T^H < 1000$ GeV	$p_T^H > 1$ TeV
Total	18	5.0	6.5	32
Statistical	16	3.0	5.5	30
Systematic	7	3.9	3.4	10
Jet systematic uncertainties	6	3.8	3.4	9.5
Modeling and theory systs.	4	0.7	0.7	2
Flavor-tagging systs.	0.2	0.4	0.4	2

VH→qqbb

ATLAS-CONF-2023-067

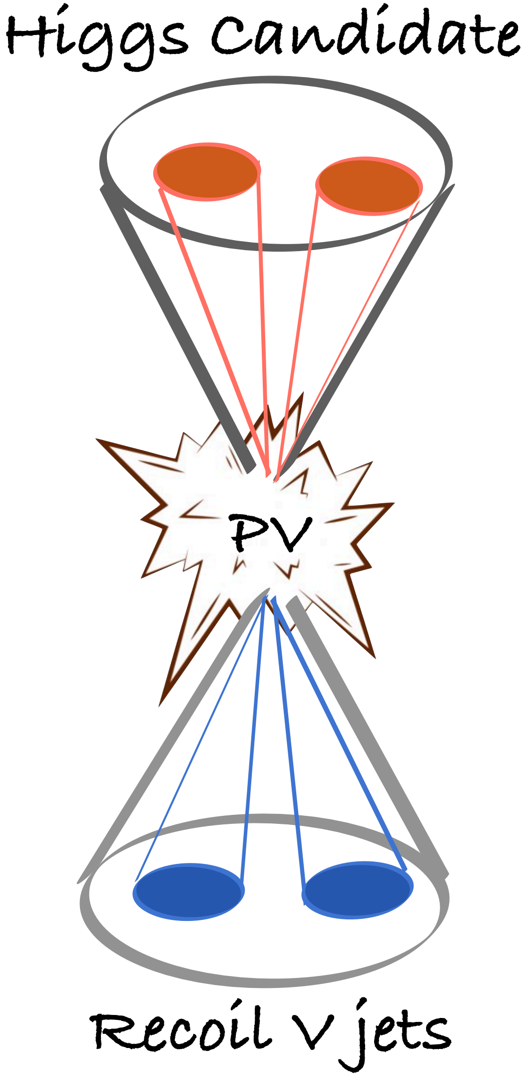
Similar trigger strategy as Hbb inclusive analysis

Signal Region:

- At least of the two p_T-leading jets must pass H→$b\bar{b}$ tagger requirements
    - The jet pass H→$b\bar{b}$ tagger (NN based) is Higgs candidate
    - If both, jet with larger mass is Higgs candidate
- Other jet must satisfy V-tagger requirements

Events are split according to Higgs-candidate p_T ($p^H_{T,J}$):

[250,450), [450,650), $\geq$ 650 GeV



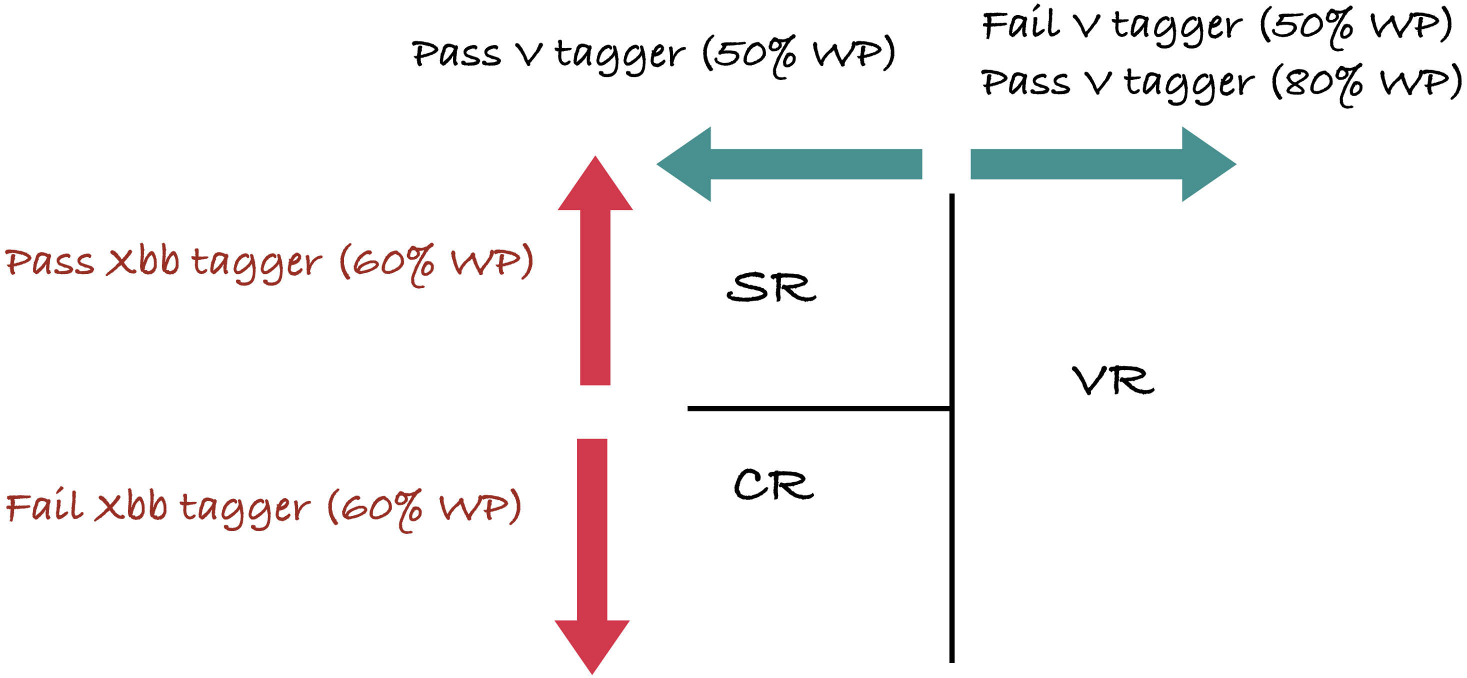
VH→qqbb: Analysis strategy

ATLAS-CONF-2023-067

Higgs candidate jet mass fit (m_J^H) to extract signal in SR

CR used to drive multijet background

Validation region: Validate multijet background



Pass V tagger (50% WP)

Fail V tagger (50% WP)
Pass V tagger (80% WP)

Pass Xbb tagger (60% WP)

Fail Xbb tagger (60% WP)

SR

CR

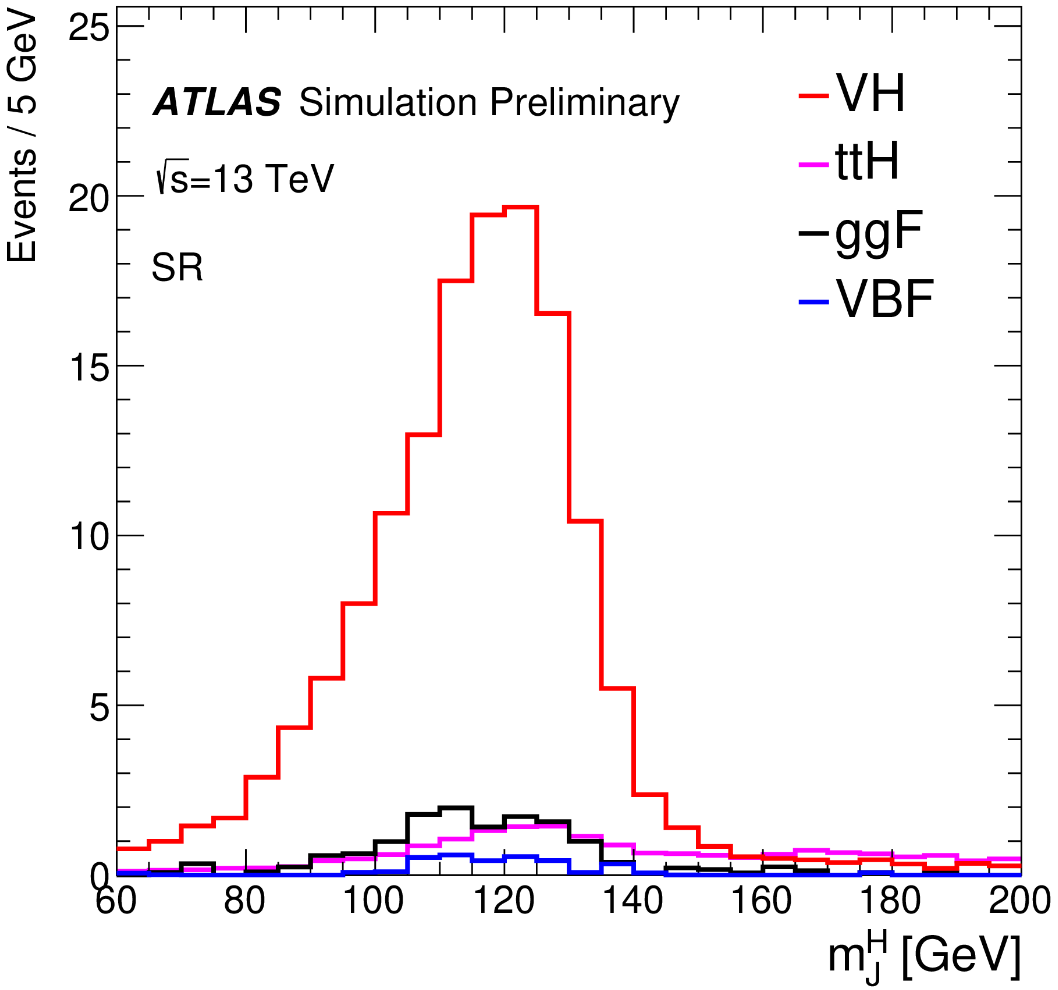
VR

VH→qqbb: Signal & Background composition

ATLAS-CONF-2023-067

In SR, **VH** production mechanism dominates: **~ 85%**

- $t\bar{t}H$ (8%), ggF (6%), VBF (1.4%)



VH→qqbb: Signal & Background composition

ATLAS-CONF-2023-067

In SR, **VH** production mechanism dominates: **~ 85%**

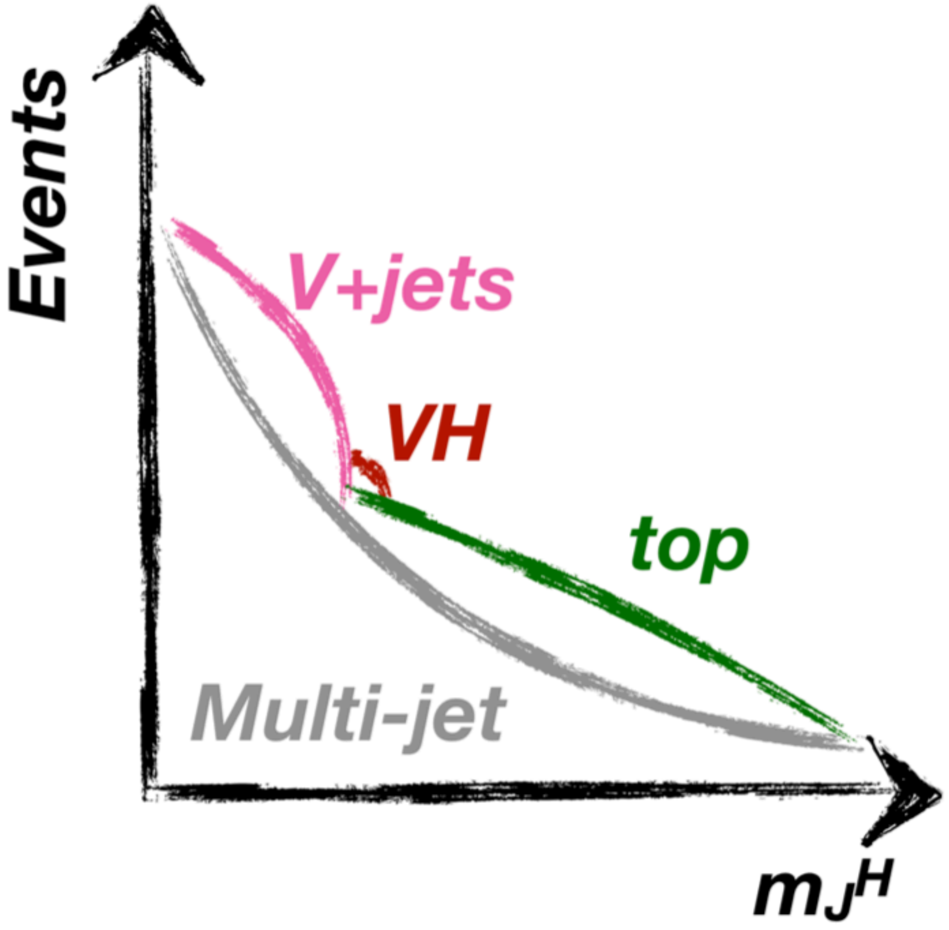
- $t\bar{t}H$ (8%), ggF (6%), VBF (1.4%)

Background dominated by multi-jets production (90%)

- $t\bar{t}$ (5%), V+jets (3.6%), VV(0.7%)

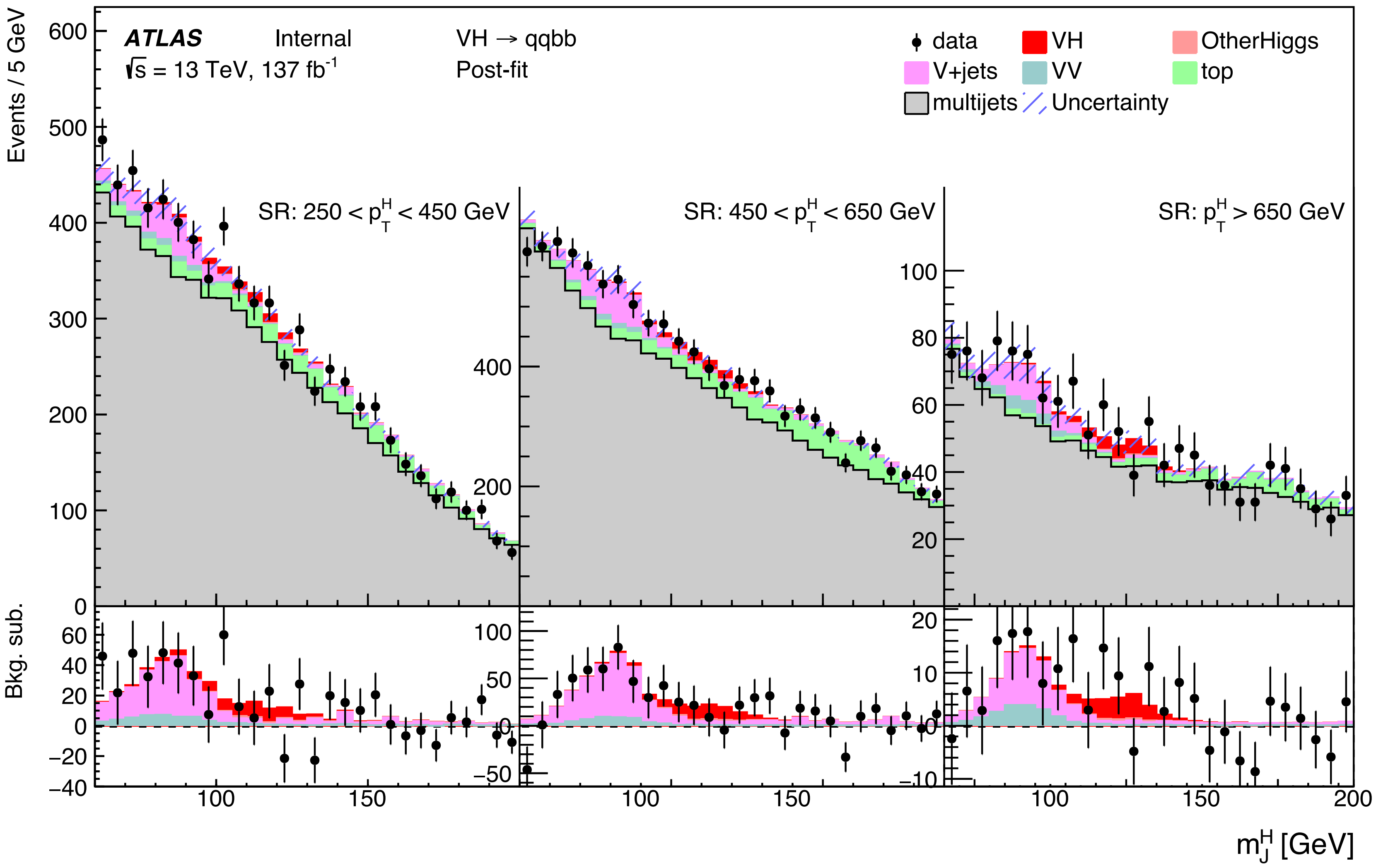
Key is to have full control of multi-jets background estimation

- Two **data-driven** estimation
- Transfer factor methods



VH→qqbb: Inclusive Results - post-fit plots in SR

ATLAS-CONF-2023-067



VH→qqbb: Inclusive Fit results

Observed Z+jets normalization: $\mu_Z = 1.41^{+0.80}_{-0.58}$

Observed VH signal strength: $\mu_{VH} = 1.39^{+1.02}_{-0.88}$

- Observed significance for rejection of null-signal hypothesis 1.7σ (1.2σ expected)
- Corresponding to an observed cross-cross section: $3.3 \pm 1.5(\mathrm{stat})^{+1.9}_{-1.5}(\mathrm{syst})$ pb

Limited by data statistics

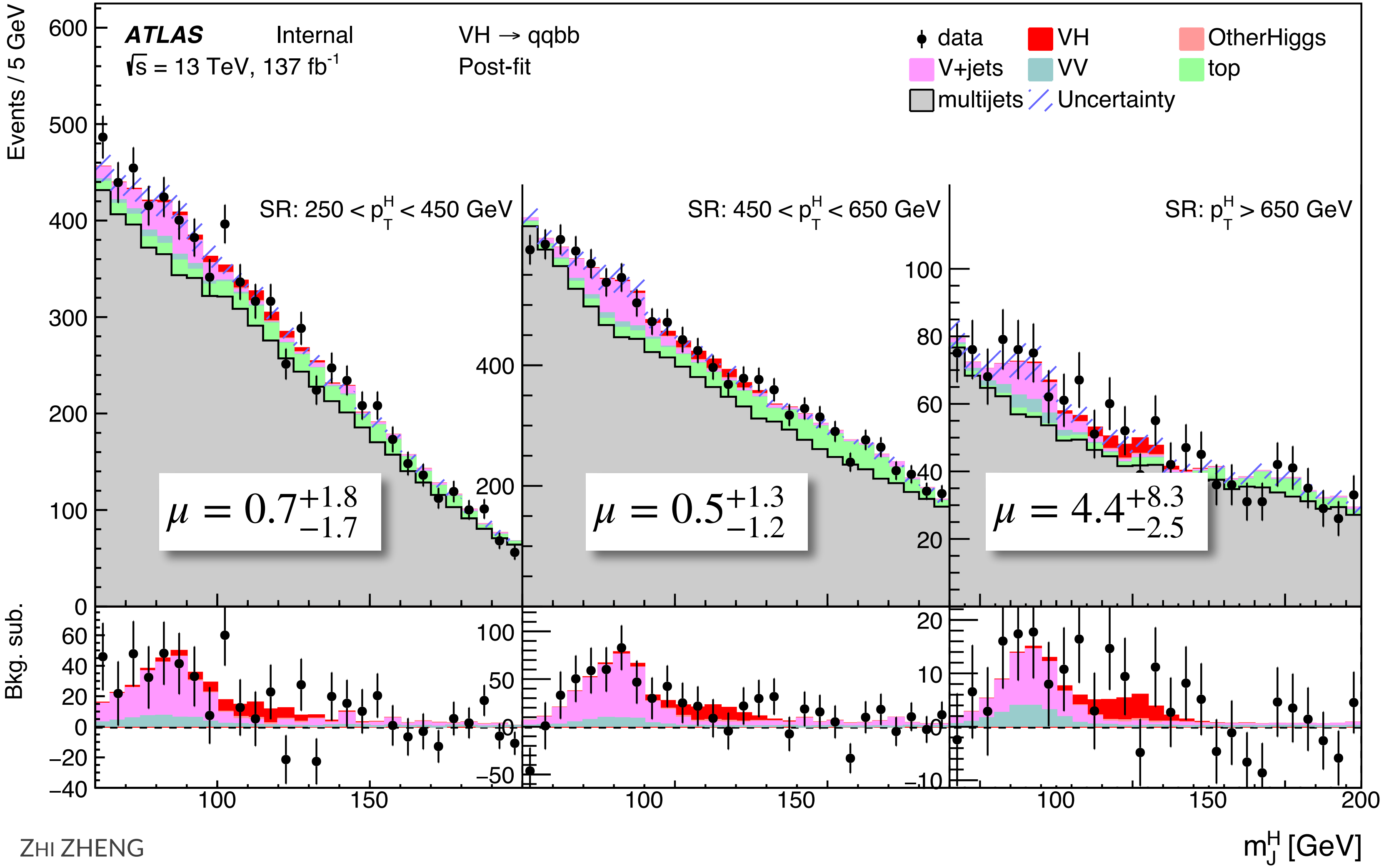
Systematics uncertainties dominate by **shape of multi-jet data-driven estimate** & **H→ $b\bar{b}$ tagger scale factors**

Uncertainty source	$\delta\mu$
Signal modeling	$^{+0.10}_{-0.02}$
MC statistical uncertainty	$^{+0.13}_{-0.13}$
Instrumental (pileup, luminosity)	$^{+0.012}_{-0.004}$
Large-R jet	$^{+0.13}_{-0.14}$
Top-quark modeling	$^{+0.14}_{-0.15}$
Other theory modeling	$^{+0.050}_{-0.031}$
$H \rightarrow b\bar{b}$ tagging	$^{+0.52}_{-0.23}$
Multijet estimate (TF uncertainty)	$^{+0.52}_{-0.41}$
Multijet modeling (TF vs. BDT)	$^{+0.14}_{-0.18}$
Total systematic uncertainty	$^{+0.80}_{-0.61}$
Signal statistical uncertainty	$^{+0.60}_{-0.60}$
Z+jets normalization	$^{+0.42}_{-0.20}$
Total statistical uncertainty	$^{+0.63}_{-0.63}$
Total uncertainty	$^{+1.02}_{-0.88}$

VH→qqbb: Differential Results

ATLAS-CONF-2023-067

Signal strength resulting from fit to each of the three p_T categories



Summary

Currently with different channels ATLAS is able to measure the different p_T distribution for Higgs

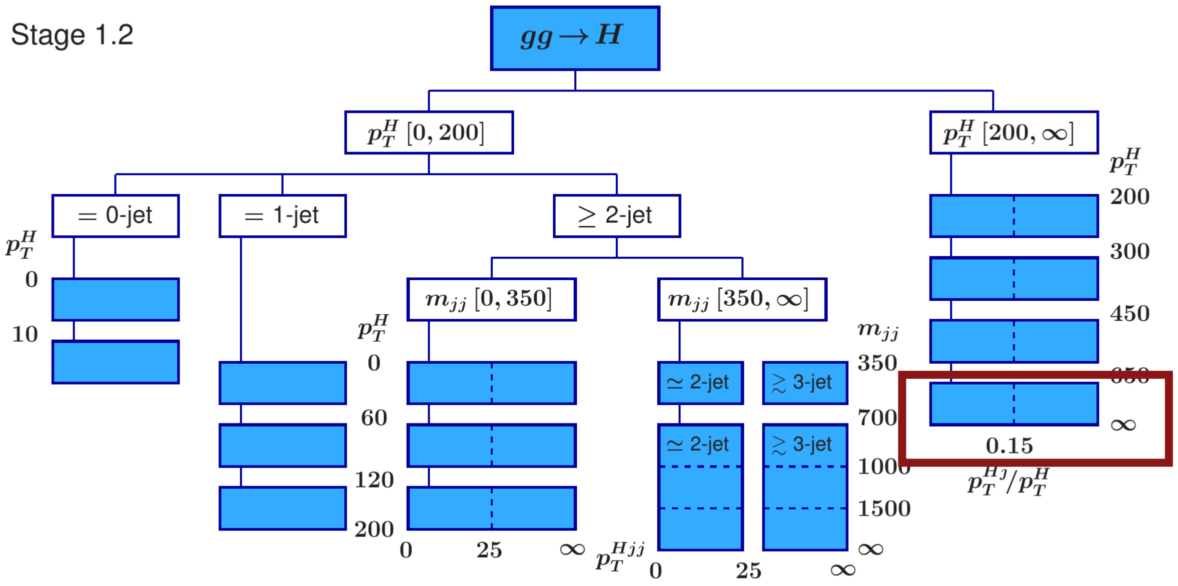
- H→ ZZ and H→ $\gamma\gamma$ Combination [Phys. Lett. B 816 (2021) 136204] → access $p_T^H > 650$ GeV
- All-Hadronic Higgs in Boost region [Phys. Rev. D 105 (2022) 092003] →access $p_T^H > 1$ TeV
- VH→ qqbb [ATLAS-CONF-2023-067] → access $p_T^H > 650$ GeV

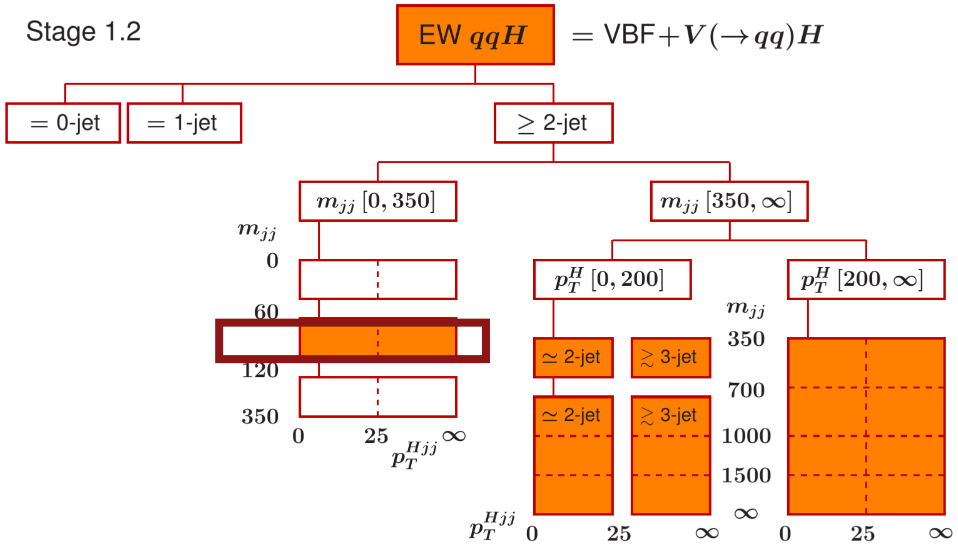
Most of the analysis still limited by statistics

For H→ bb channels improvements are expected to come with better tagger

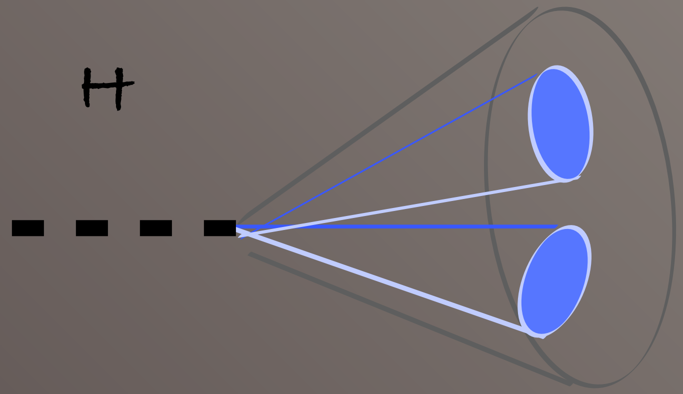
Wishlist

- NLO EW corrections for ggF, partially unknown and expected to be sizable at large transverse momentum!
- Current STXS is not ideal for boosted bb analysis
  - We have ability to reach 1 TeV for ggF
  - Introducing additional p_T^H bins in $p_T^H > 200$ GeV would help the VH all hadronic analysis



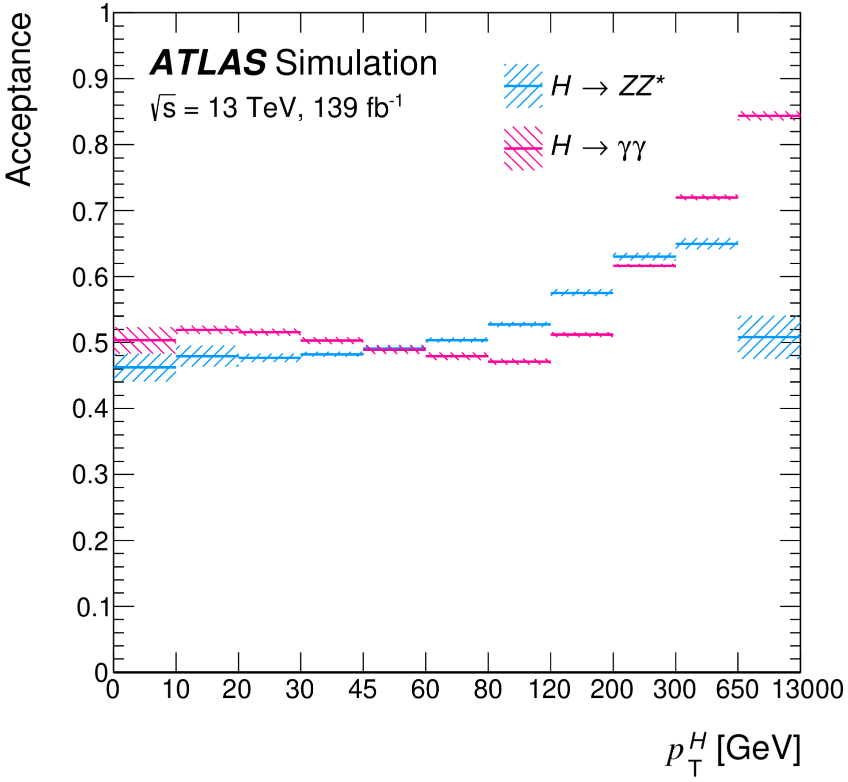


Back up



H➝ ZZ and H➝ $\gamma\gamma$ combination

[Phys. Lett. B 816 (2021) 136204]



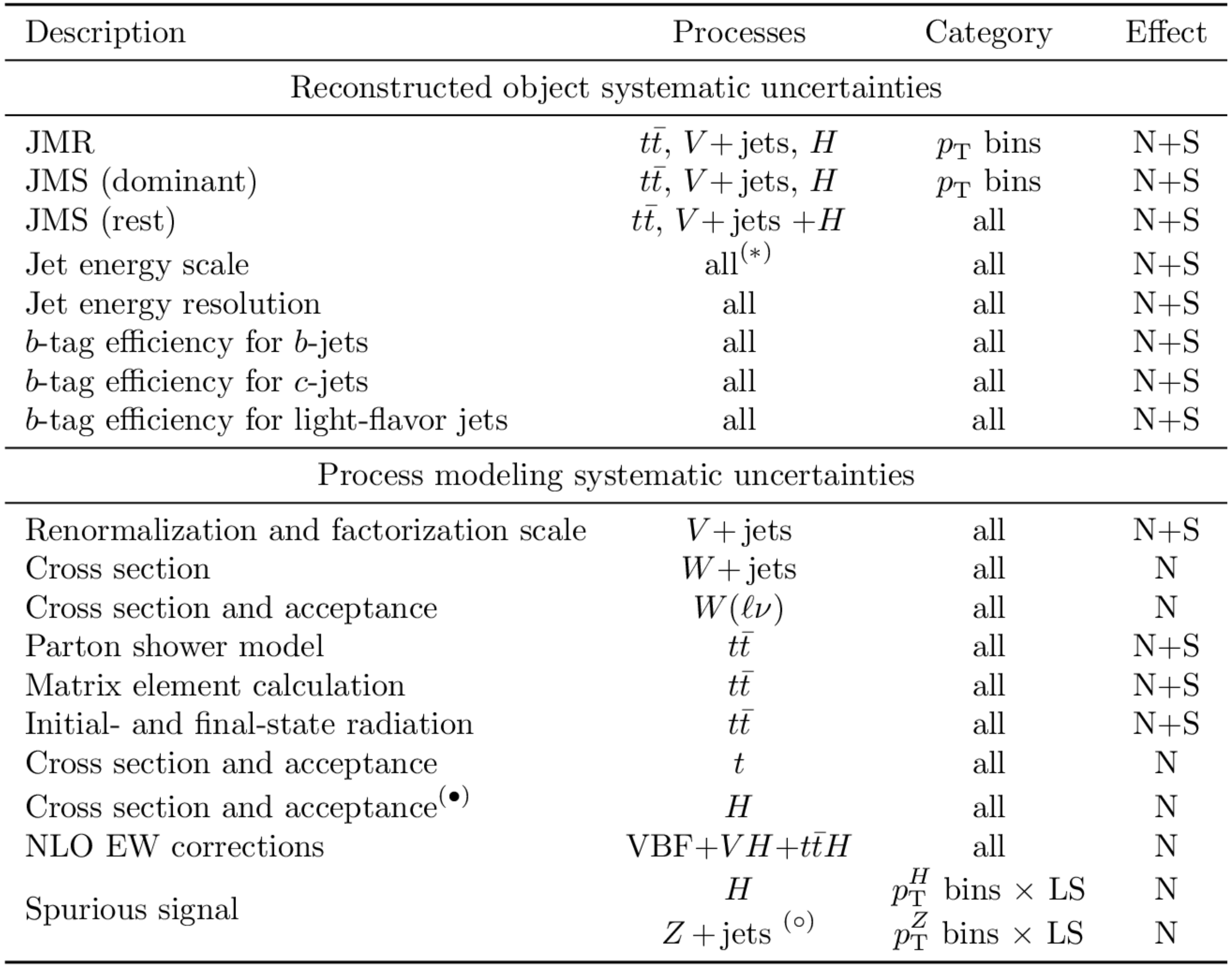
Limited by Statistics from both channels

The acceptance drops for H➝ ZZ due to lepton separation requirement ($\Delta R(l_i,l_j) > 0.1$)

The acceptance for H➝ $\gamma\gamma$ increase because of pT requirement for photon

Boosted All-Had H→ bb: Systematics

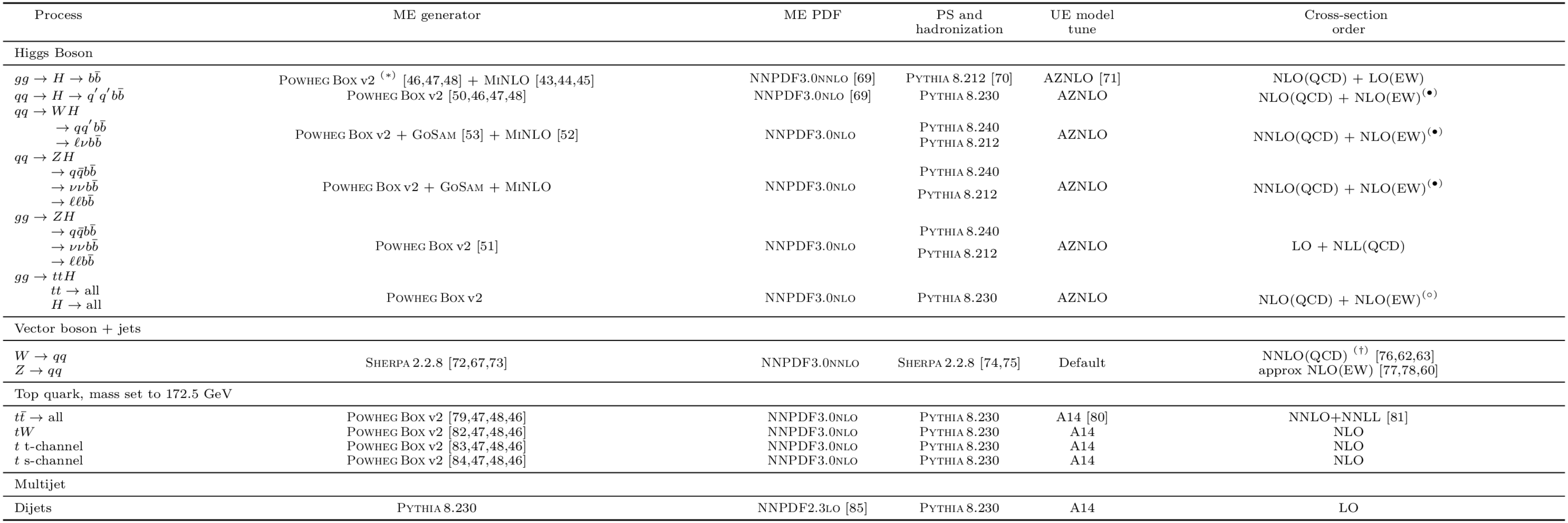
Phys. Rev. D 105 (2022) 092003



Description	Processes	Category	Effect
Reconstructed object systematic uncertainties			
JMR	$t\bar{t}$, $V+$jets, H	p_{T} bins	N+S
JMS (dominant)	$t\bar{t}$, $V+$jets, H	p_{T} bins	N+S
JMS (rest)	$t\bar{t}$, $V+$jets $+H$	all	N+S
Jet energy scale	all$^{(*)}$	all	N+S
Jet energy resolution	all	all	N+S
b-tag efficiency for b-jets	all	all	N+S
b-tag efficiency for c-jets	all	all	N+S
b-tag efficiency for light-flavor jets	all	all	N+S
Process modeling systematic uncertainties			
Renormalization and factorization scale	$V+$jets	all	N+S
Cross section	$W+$jets	all	N
Cross section and acceptance	$W(\ell\nu)$	all	N
Parton shower model	$t\bar{t}$	all	N+S
Matrix element calculation	$t\bar{t}$	all	N+S
Initial- and final-state radiation	$t\bar{t}$	all	N+S
Cross section and acceptance	t	all	N
Cross section and acceptance$^{(\bullet)}$	H	all	N
NLO EW corrections	VBF+VH+$t\bar{t}H$	all	N
Spurious signal	H	p_{T}^{H} bins × LS	N
	$Z+$jets $^{(\circ)}$	p_{T}^{Z} bins × LS	N

Boosted All-Had H→ bb: MC

Phys. Rev. D 105 (2022) 092003



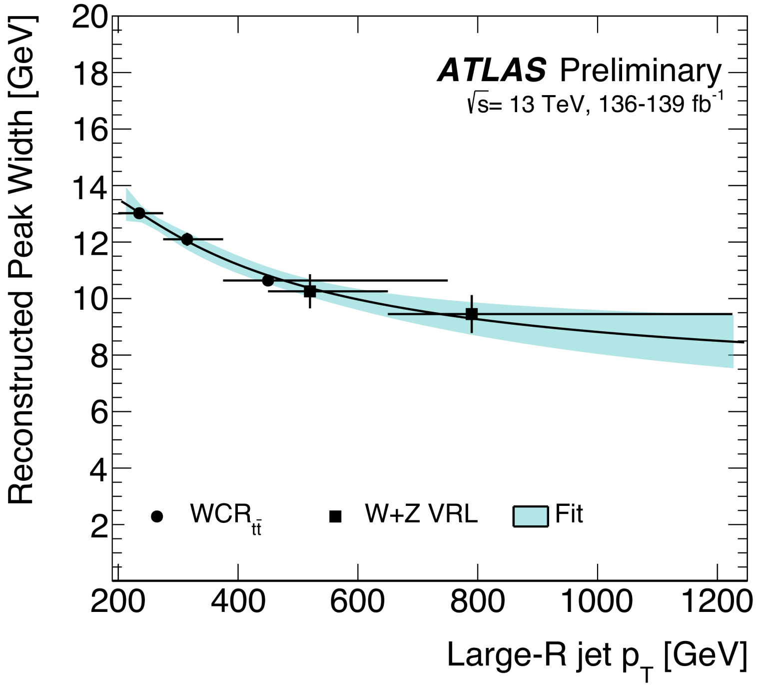
Process	ME generator	ME PDF	PS and hadronization	UE model tune	Cross-section order
Higgs Boson					
$gg \to H \to b\bar{b}$	Powheg Box v2 $^{(*)}$ [46,47,48] + MiNLO [43,44,45]	NNPDF3.0nnlo [69]	Pythia 8.212 [70]	AZNLO [71]	NLO(QCD) + LO(EW)
$qq \to H \to q'q'b\bar{b}$	Powheg Box v2 [50,46,47,48]	NNPDF3.0nlo [69]	Pythia 8.230	AZNLO	NLO(QCD) + NLO(EW)$^{(\bullet)}$
$qq \to WH$	Powheg Box v2 + GoSam [53] + MiNLO [52]	NNPDF3.0nlo		AZNLO	NNLO(QCD) + NLO(EW)$^{(\bullet)}$
$\to qq'b\bar{b}$			Pythia 8.240		
$\to \ell\nu b\bar{b}$			Pythia 8.212		
$qq \to ZH$	Powheg Box v2 + GoSam + MiNLO	NNPDF3.0nlo		AZNLO	NNLO(QCD) + NLO(EW)$^{(\bullet)}$
$\to q\bar{q}b\bar{b}$			Pythia 8.240		
$\to \nu\nu b\bar{b}$			Pythia 8.212		
$\to \ell\ell b\bar{b}$					
$gg \to ZH$	Powheg Box v2 [51]	NNPDF3.0nlo		AZNLO	LO + NLL(QCD)
$\to q\bar{q}b\bar{b}$			Pythia 8.240		
$\to \nu\nu b\bar{b}$			Pythia 8.212		
$\to \ell\ell b\bar{b}$					
$gg \to ttH$	Powheg Box v2	NNPDF3.0nlo	Pythia 8.230	AZNLO	NLO(QCD) + NLO(EW)$^{(\circ)}$
$tt \to$ all					
$H \to$ all					
Vector boson + jets					
$W \to qq$ $Z \to qq$	Sherpa 2.2.8 [72,67,73]	NNPDF3.0nnlo	Sherpa 2.2.8 [74,75]	Default	NNLO(QCD) $^{(\dagger)}$ [76,62,63] approx NLO(EW) [77,78,60]
Top quark, mass set to 172.5 GeV					
$t\bar{t} \to$ all	Powheg Box v2 [79,47,48,46]	NNPDF3.0nlo	Pythia 8.230	A14 [80]	NNLO+NNLL [81]
tW	Powheg Box v2 [82,47,48,46]	NNPDF3.0nlo	Pythia 8.230	A14	NLO
t t-channel	Powheg Box v2 [83,47,48,46]	NNPDF3.0nlo	Pythia 8.230	A14	NLO
t s-channel	Powheg Box v2 [84,47,48,46]	NNPDF3.0nlo	Pythia 8.230	A14	NLO
Multijet					
Dijets	Pythia 8.230	NNPDF2.3lo [85]	Pythia 8.230	A14	LO

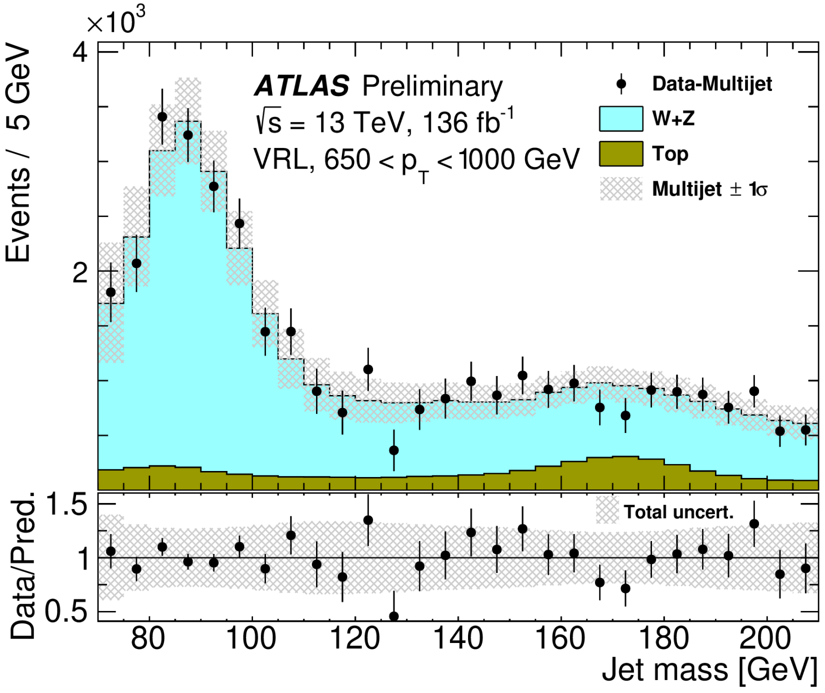
Boosted All-Had H→ bb: W/Z jets

[Phys. Rev. D 105 (2022) 092003]

Interaction between V+j template and QCD analytic function make it impossible to simultaneously extract V+j width and normalization

→ Extraction of dedicated W/Z Jet mass resolution constrains in VRL and VRS





Boosted All-Had H→ bb: Background

Phys. Rev. D 105 (2022) 092003

Background: Dominated with QCD

QCD multijet:

modeled with smooth **analytical model**

W/Z + jets:

modeled with Sherpa 2.2.8 (NLO QCD+NLO EW) + QCD NNLO k-factor by NNLOJet group

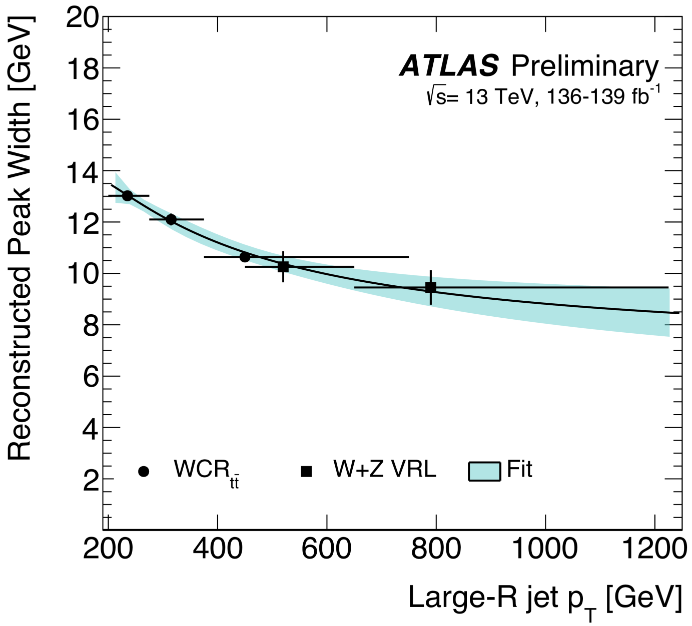
“**Standard candle**” for Higgs extraction

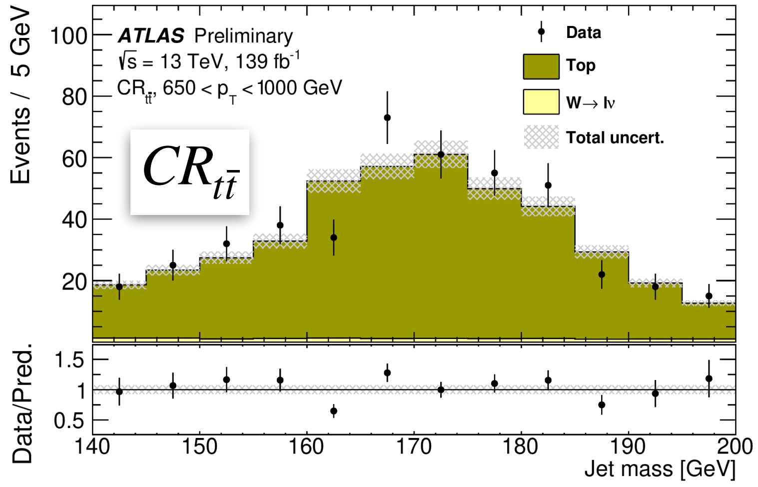
Mass resolution constrains from VRs

Top:

Model by Powheg+Pythia

Normalisation contraient in semileptonic CR





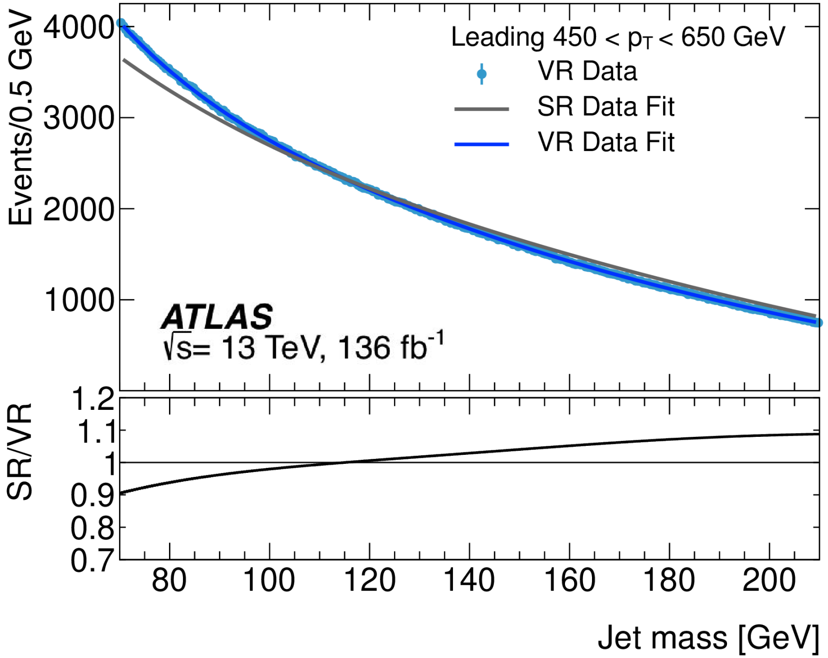
Boosted All-Had H→ bb: QCD

Phys. Rev. D 105 (2022) 092003

$\mathrm{VR_{hyb}}$ used for data-driven validation of QCD fit model

$$\mathrm{VR}^i_{\mathrm{hyb}} = \underbrace{\boxed{(\mathrm{VR}^i - V_{\mathrm{VR}} - \mathrm{Top}_{\mathrm{VR}})}}_{\text{resonance subtracted VR}} \times \overbrace{\boxed{\left(\frac{\mathrm{MJ_{SR}}}{\mathrm{MJ_{VR}}}\right)}}^{\text{residual corrections}} + \underbrace{\boxed{V_{\mathrm{SR}} + t\bar{t}_{\mathrm{SR}} + H_{\mathrm{SR}}}}_{\text{signal injection}}$$

- **Each signal region (ie: p_{TJ} bin) has corresponding validation region**
- **Data in VR split into slices with statistics representing SR**
  - Tests fit model in scenario statistically equal to final SR fits
  - Results from each slice are averaged for final number
- **$\mathrm{MJ_{S/VR}}$ from full fit to signal/val regions**
  - Higgs mass window blinded in SR's
  - $\mathrm{MJ_{VR}}$ is average of 10 slices



The smooth function: **For QCD**

(eq 1) $$f_N\left(x \middle| \vec{\theta}\right) = \theta_0 \exp\left(\sum_{i=1}^{N} \theta_i x^i\right) \qquad x = (m_J - 140\ \mathrm{GeV})/70\ \mathrm{GeV}$$

VH→qqbb

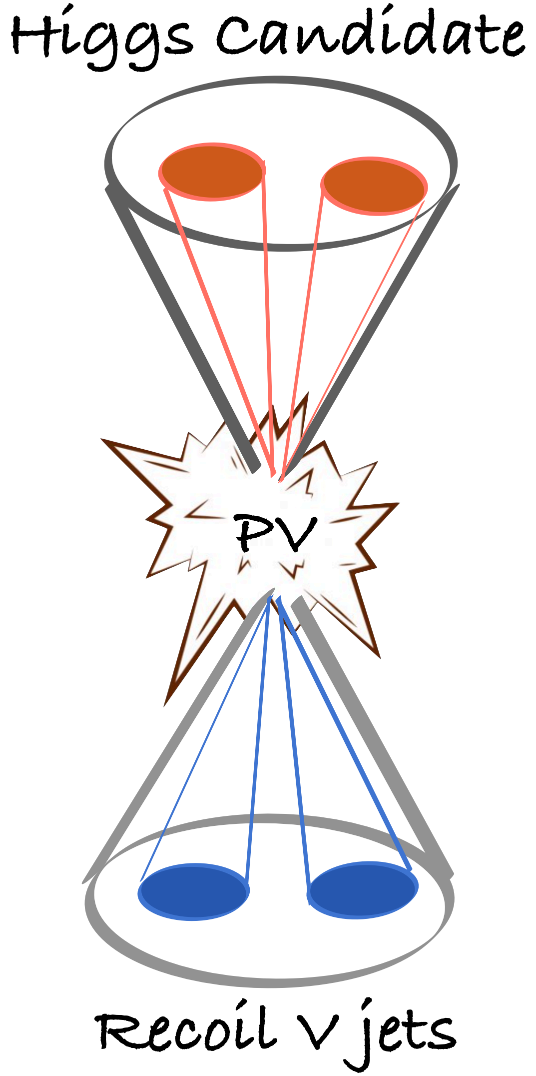
ATLAS-CONF-2023-067

Single large-R (R=1.0 anti-k_t) jet trigger with Mass and p_T threshold

At least two large-R jets p_T > 200 GeV & $|\eta|$ <2

- p_T- leading jet: p_T > 450, M_J > 60 GeV
- Second p_T- leading jet: M_J > 40 GeV

Events with isolated charges leptons are rejected



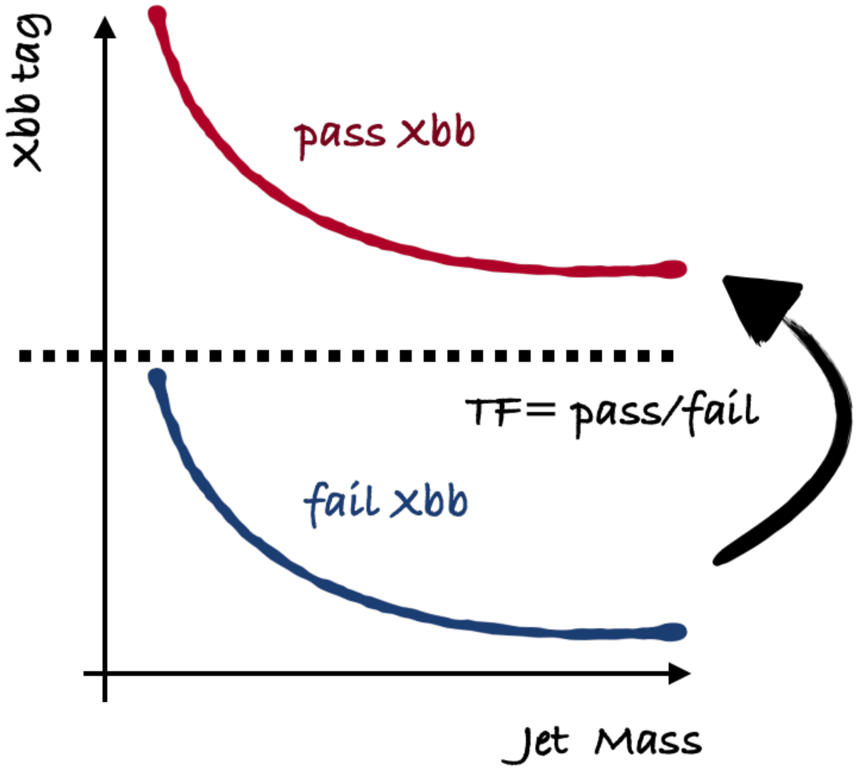
VH→qqbb: multijet background

ATLAS-CONF-2023-067

Aim to predict the multijet mass distribution in the SR using event in CR region:

SR = CR ×transfer factor (TF)

- TF as a function of $\rho = \log(m^2/p_T^2)$ and pT
- $TF(p_T, \rho) = \sum_{k,l} \alpha_{kl}\rho^k p_T^l$, where α_{kl} are polynomial coefficients
- Polynomial order determined via Fisher F-test
  - 1 st in both p_T and ρ



Alternative method: BDT

- Validation TF method and used as systematics