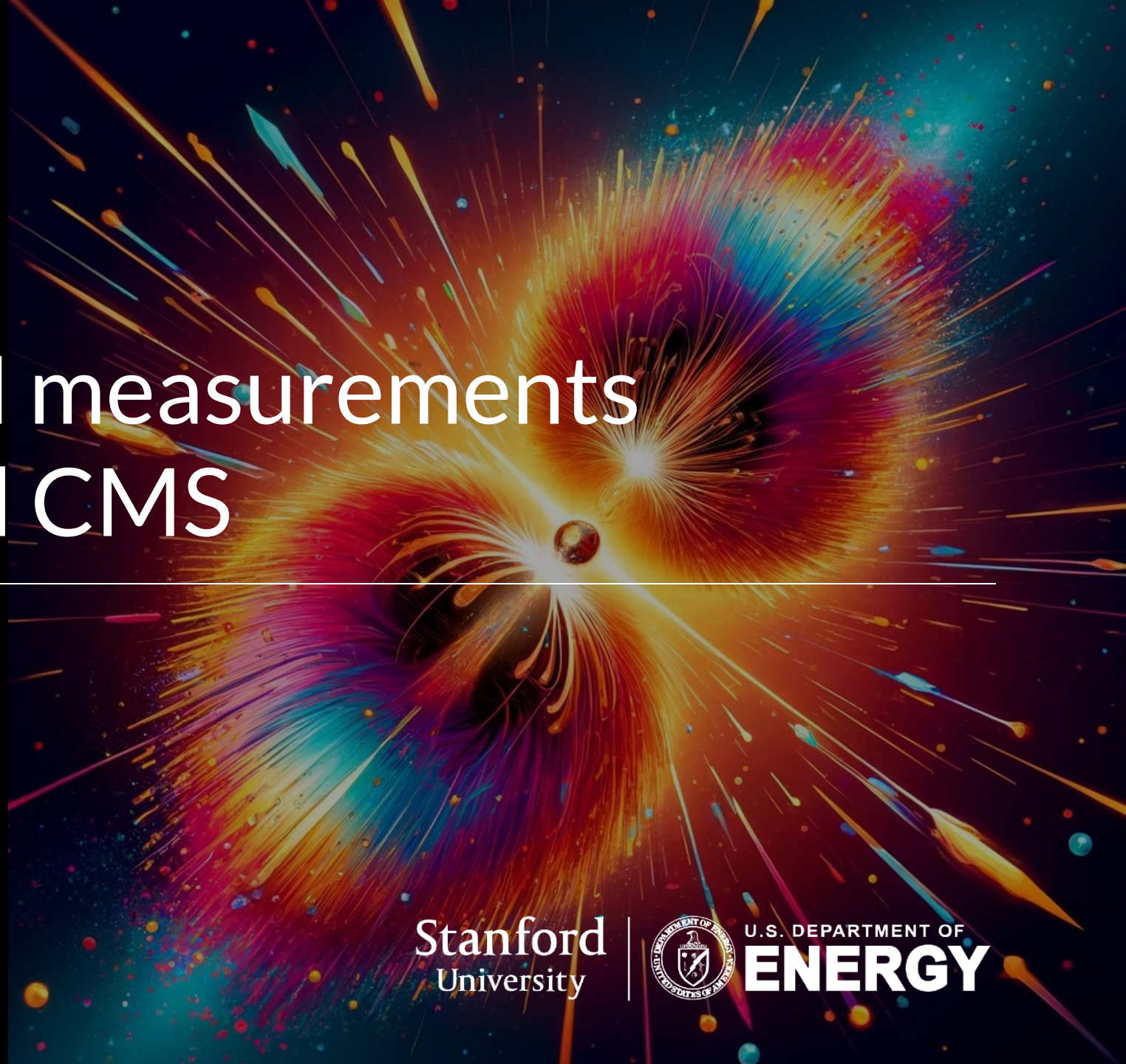
All-hadronic VH measurements from ATLAS and CMS

Zhi ZHENG

12/04/2025

LHCHWG 2025



SLAC NATIONAL ACCELERATOR LABORATORY

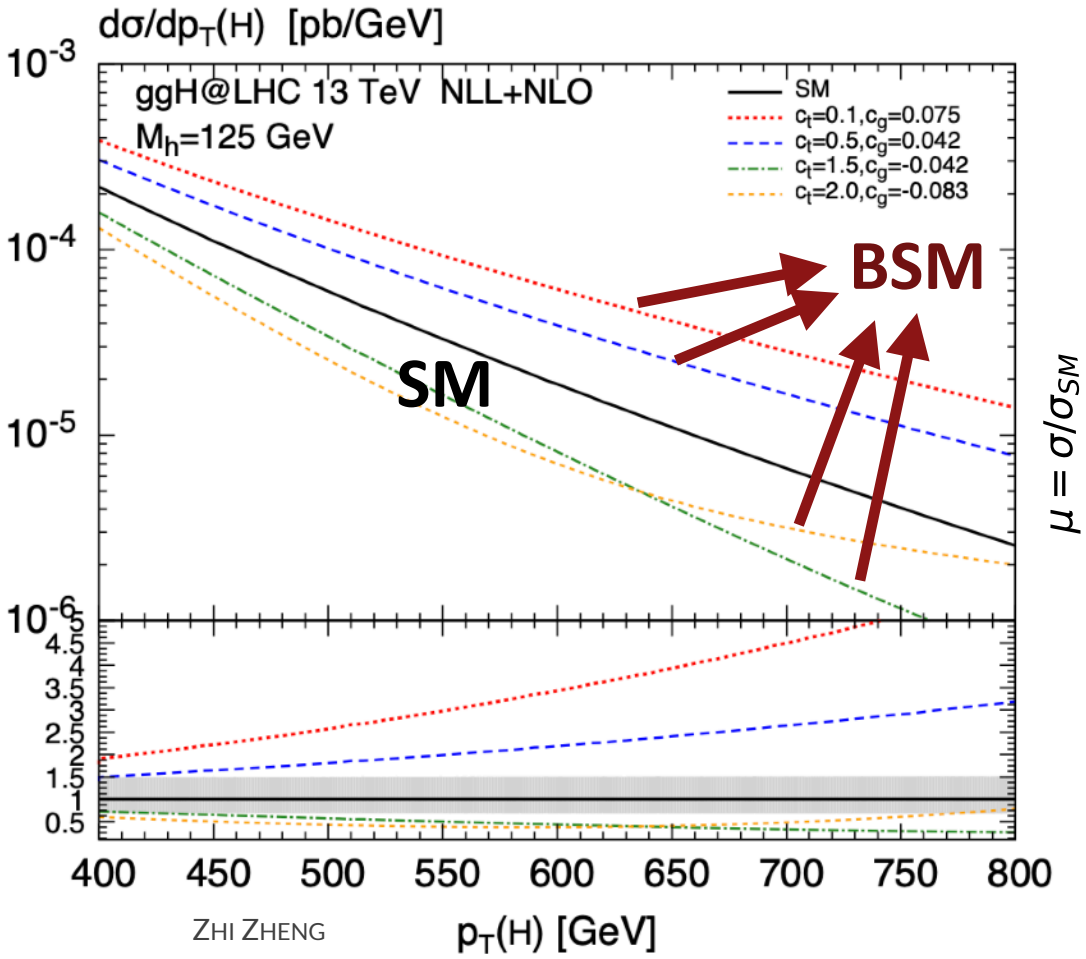
Stanford University

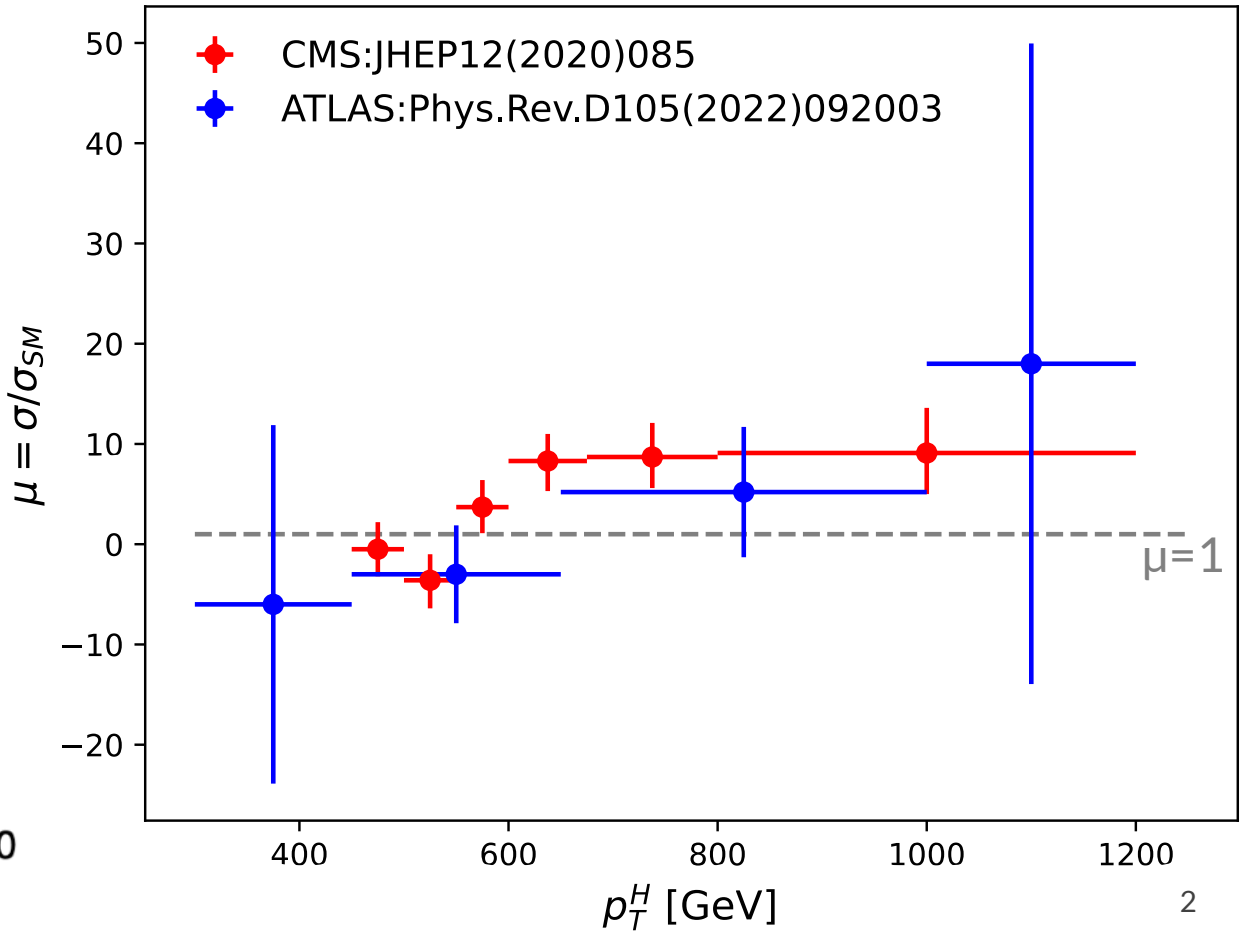
U.S. DEPARTMENT OF ENERGY

Higgs Production at High p_T

New physics contributions can modify Higgs production at high momentum

H(bb) is a special tool to probe high momentum H production thanks to **boosted b-tagging techniques**



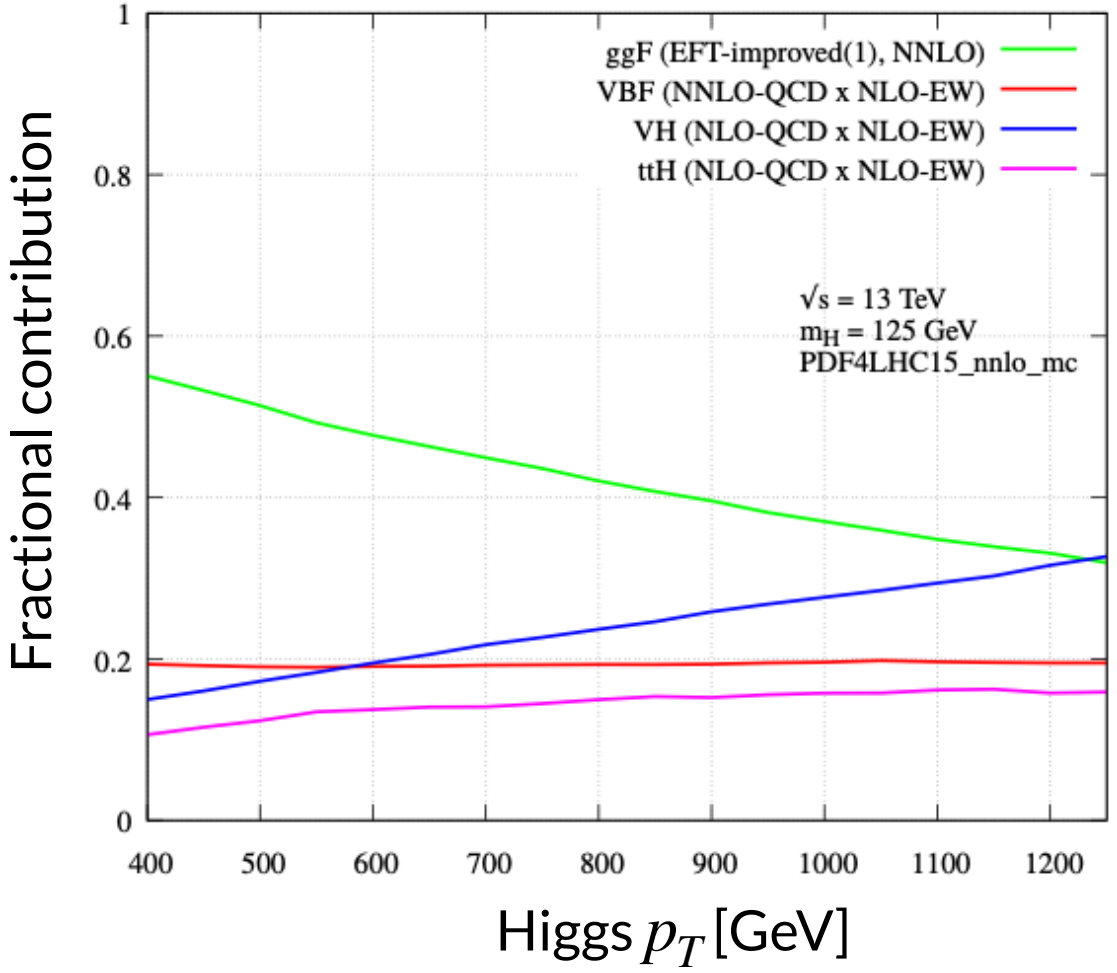


Higgs in the Boost Region: Different Production

All production modes contribute similarly toward p_T^H ~ 1 TeV

Understanding each production mechanism in high p_T regime for the Higgs boson is important

arXiv:2005.07762



ATLAS vs CMS

ATLAS EXPERIMENT 2024

Phys. Rev. Lett. 132 (2024) 131802

Study of High-Transverse-Momentum Higgs Boson Production in Association with a Vector Boson in the *qqbb* Final State with the ATLAS Detector

G. Aad *et al.*[*]
(ATLAS Collaboration)

(Received 14 December 2023; accepted 26 February 2024; published 28 March 2024)

This Letter presents the first study of Higgs boson production in association with a vector boson ($V = W$ or Z) in the fully hadronic $qqbb$ final state using data recorded by the ATLAS detector at the LHC in proton-proton collisions at $\sqrt{s} = 13$ TeV and corresponding to an integrated luminosity of 137 fb^{-1}. The vector bosons and Higgs bosons are each reconstructed as large-radius jets and tagged using jet substructure techniques. Dedicated tagging algorithms exploiting b-tagging properties are used to identify jets consistent with Higgs bosons decaying into $b\bar{b}$. Dominant backgrounds from multijet production are determined directly from the data, and a likelihood fit to the jet mass distribution of Higgs boson candidates is used to extract the number of signal events. The VH production cross section is measured inclusively and differentially in several ranges of Higgs boson transverse momentum: 250–450, 450–650, and greater than 650 GeV. The inclusive signal yield relative to the standard model expectation is observed to be $\mu = 1.4^{+1.0}_{-0.9}$ and the corresponding cross section is $3.1 \pm 1.3(\text{stat})^{+1.8}_{-1.4}(\text{syst})$ pb.

CMS 2025

CMS-PAS-HIG-24-017

Available on the CERN CDS information server

CMS PAS HIG-24-017

CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch

2025/03/25

Search for a boosted Higgs boson decaying to bottom quark pairs in association with a hadronically decaying W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

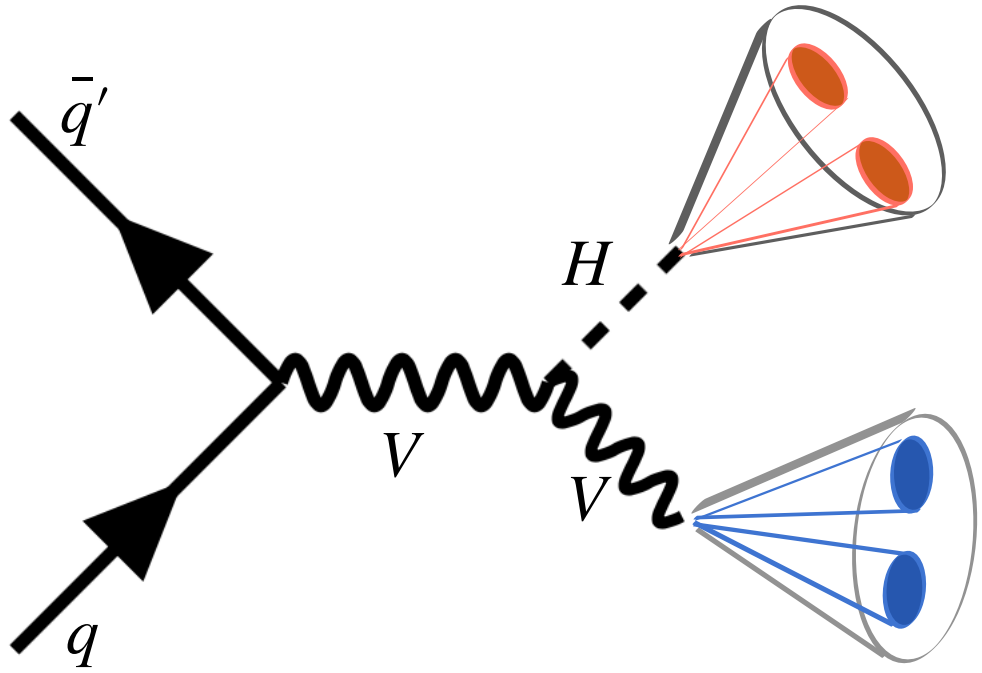
VH fully hadronic, an easy topology

Fully hadronic trigger selection

Two large radius jets

Substructure and b-tagging to identify H and W/Z candidates

Data-driven background estimate



V(→qq)H: Event Selection

Single large-R (R=1 anti-k_t) jet trigger with mass and p_T requirements

At least **two large-R jets** p_T > 250 GeV & $|\eta|$ <2

- pT-leading jet: p_T> 450 GeV, M_J>50 GeV, $|\eta|$< 2
- Second pT-leading jet: p_T> 250 GeV, M_J>40 GeV, $|\eta|$< 2

Event with isolated charge leptons are rejected

H_T trigger or Single large-R (R=0.8 anti-k_t) jet trigger with mass and pT requirements

Two large-R jets: p_T> 450 GeV, $|\eta|$< 2.5,

$-0.6 < \rho < -2.1$ $[\rho = 2ln(m_{SD}/p_T)]$

Event with isolated charge leptons are rejected

p_T^{miss}<40 GeV

V(→qq)H: Higgs Candidate

Jet with X→bb tagger score:

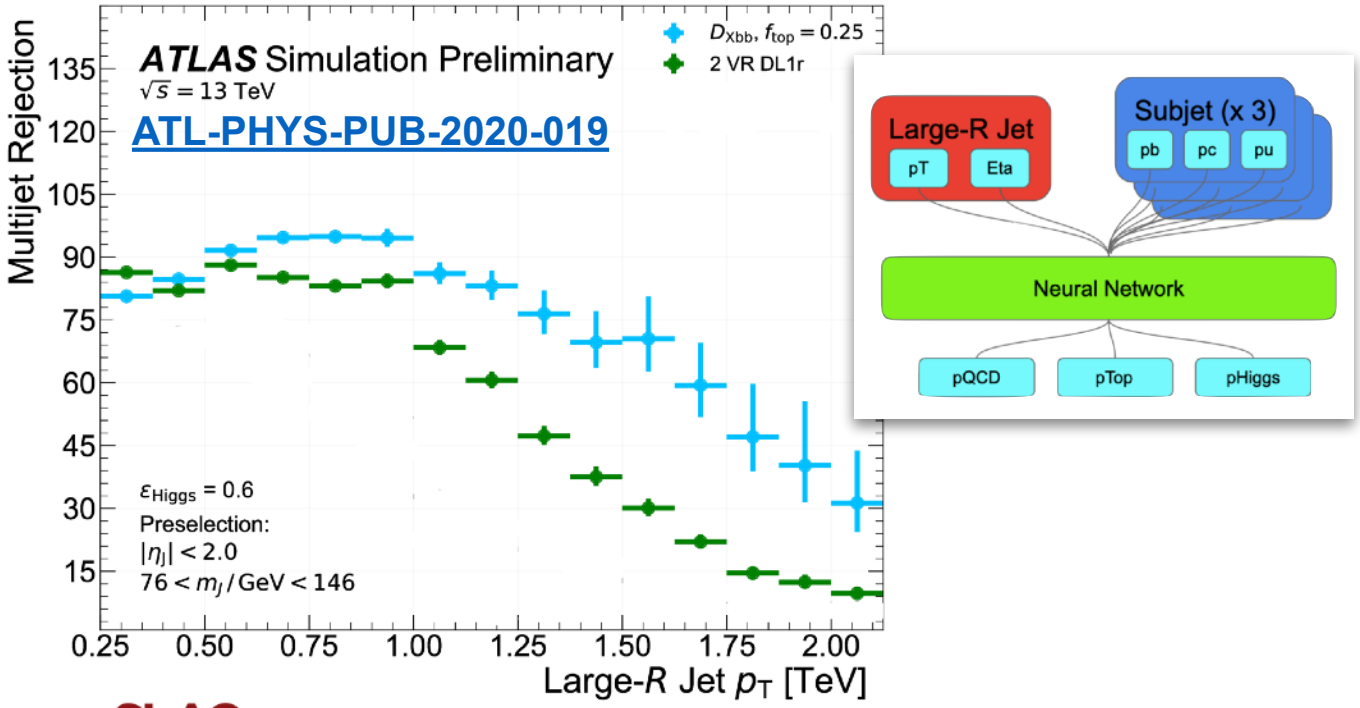
- DNN using DL1r output associated to variable radius track-jets and jet information
- Fixed 60% H→ bb efficiency, multijet mistag rate ~ 1%

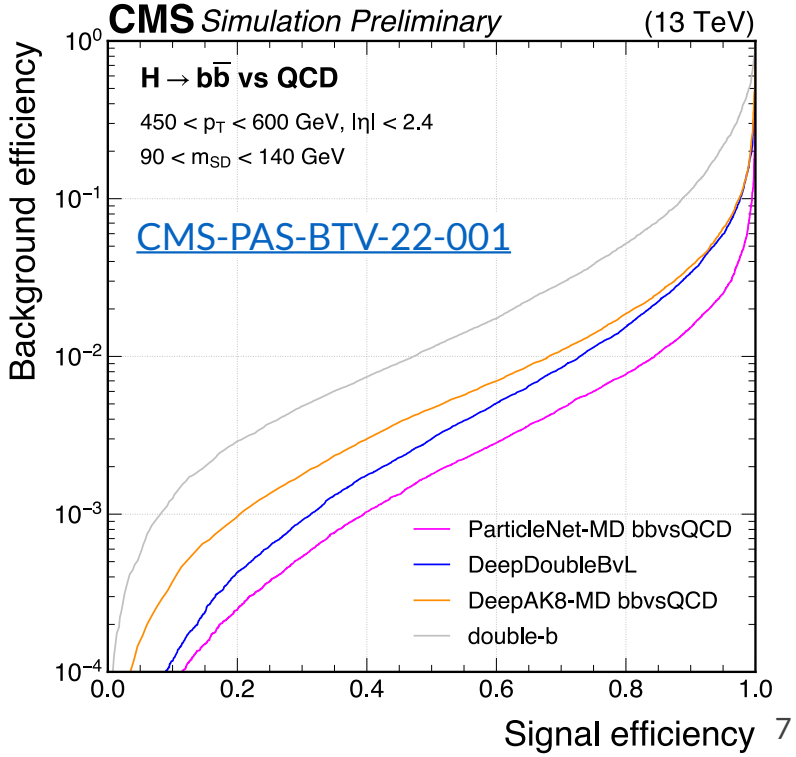
Jet with higher ParticleNet-MD D(bb vs qq)

$$\mathcal{D}(\mathrm{bb\ vs\ qq}) = \frac{p(X \to b\bar{b})}{p(X \to b\bar{b}) + p(X \to c\bar{c}) + p(X \to q\bar{q})}.$$

- particleNet-MD using particle-flow candidate & secondary vertex features

➡60% H→bb efficiency, mistag rate ~0.5%



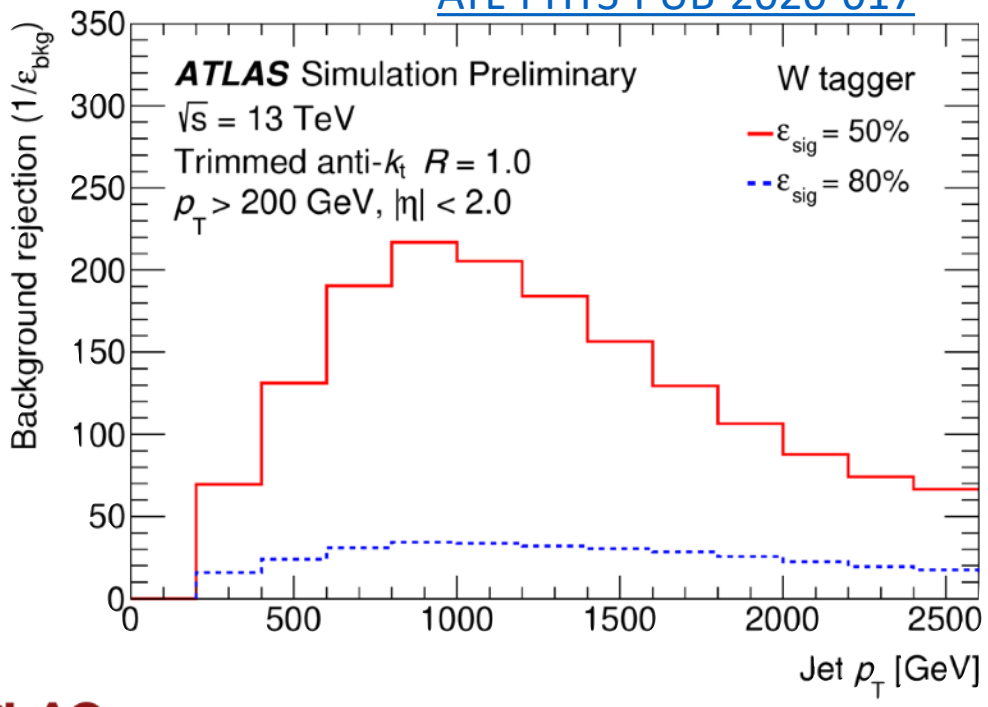


V(→qq)H: V Candidate

Requirements on jet mass, two-prongness & number of tracks

- Signal efficiency of 50% with a multijet mistag rate 0.5-1.4%

ATL-PHYS-PUB-2020-017



Use particleNet-MD: mass-decorrelated tagger for 2-prong hadronic decays

- Signal efficiency 25% and multijet mistag rate 3%

$$PN^{MD}_{QCD} = 1 - (PN^{MD}_{Xbb} + PN^{MD}_{Xcc} + PN^{MD}_{Xqq})$$

$$= 1 - PN^{MD}_{\text{2-prong}}$$

Analysis Strategy: ATLAS

H → $b\bar{b}$ tagger

Pass

SR

VR

60% WP

Fail

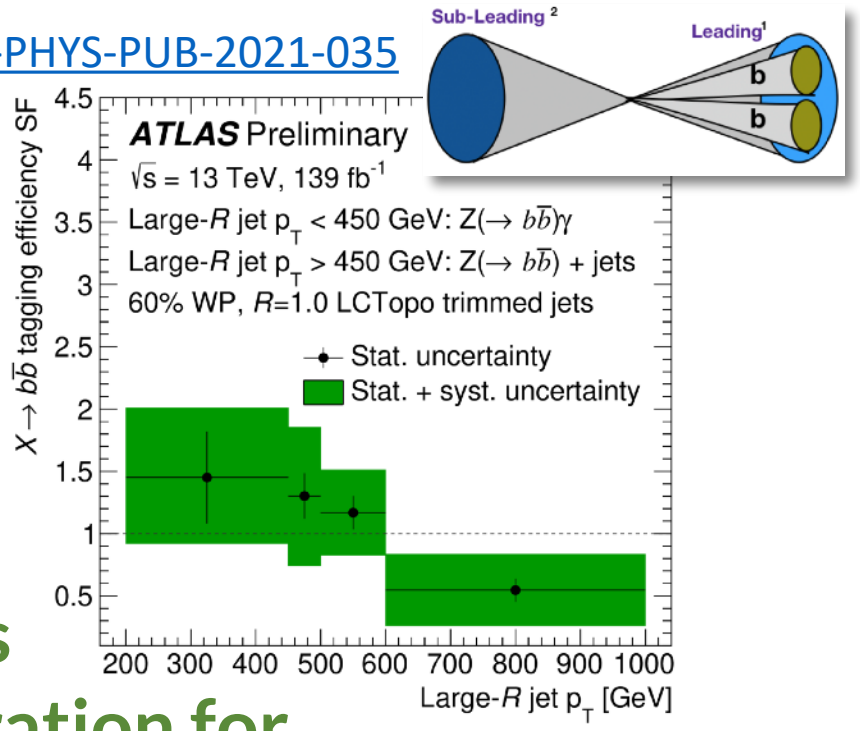
CR

Derive multi-jets background

Validate multi-jet background estimation method

Z+jets Calibration for H→ bb tagger

ATL-PHYS-PUB-2021-035



*SR: Signal Region
*CR: Control Region
*VR: Validation Region

Pass

50% WP

80% WP

Fail

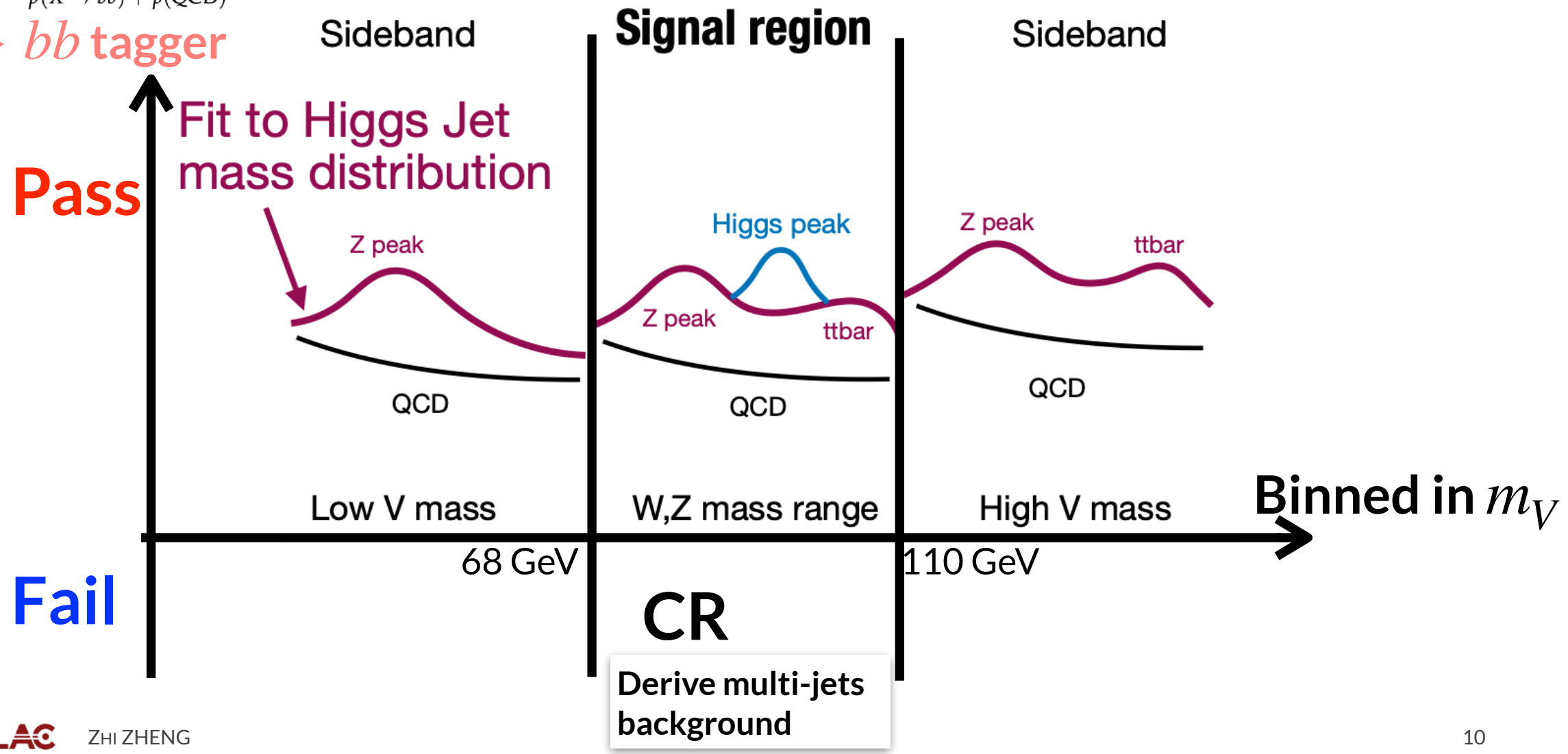
V tagger

Analysis Strategy: CMS

Duc Hoang

$$PNet_{Xbb} = \frac{p(X \to b\bar{b})}{p(X \to b\bar{b}) + p(\mathrm{QCD})}$$

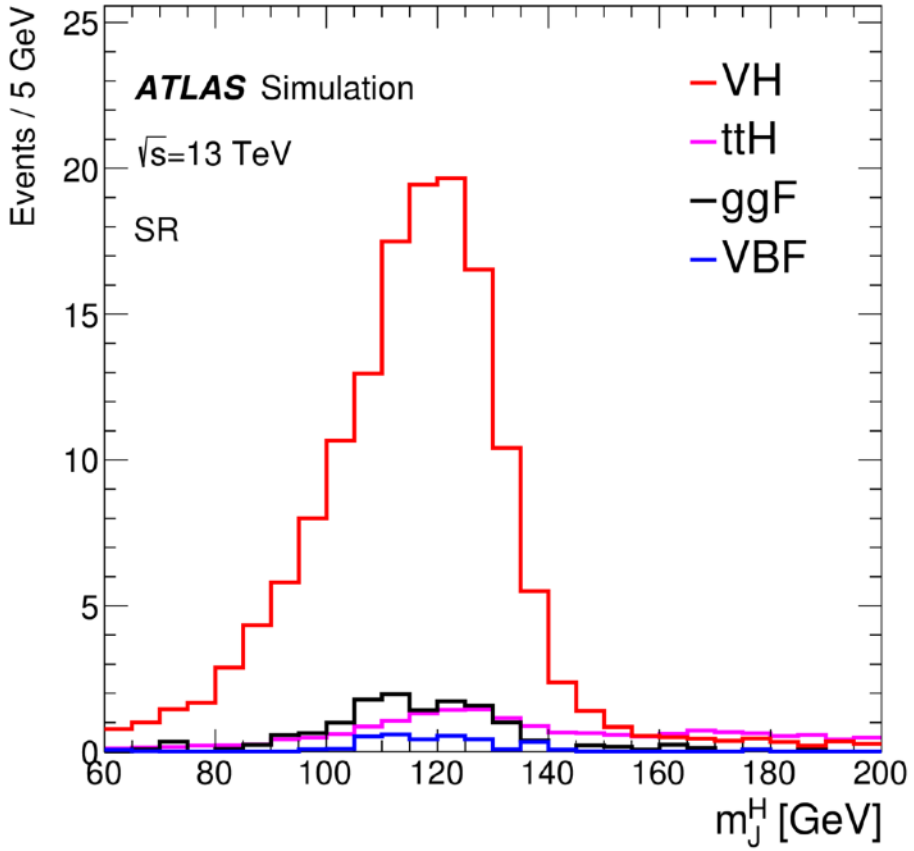
H → bb tagger



Signal Region: Signal Purity

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

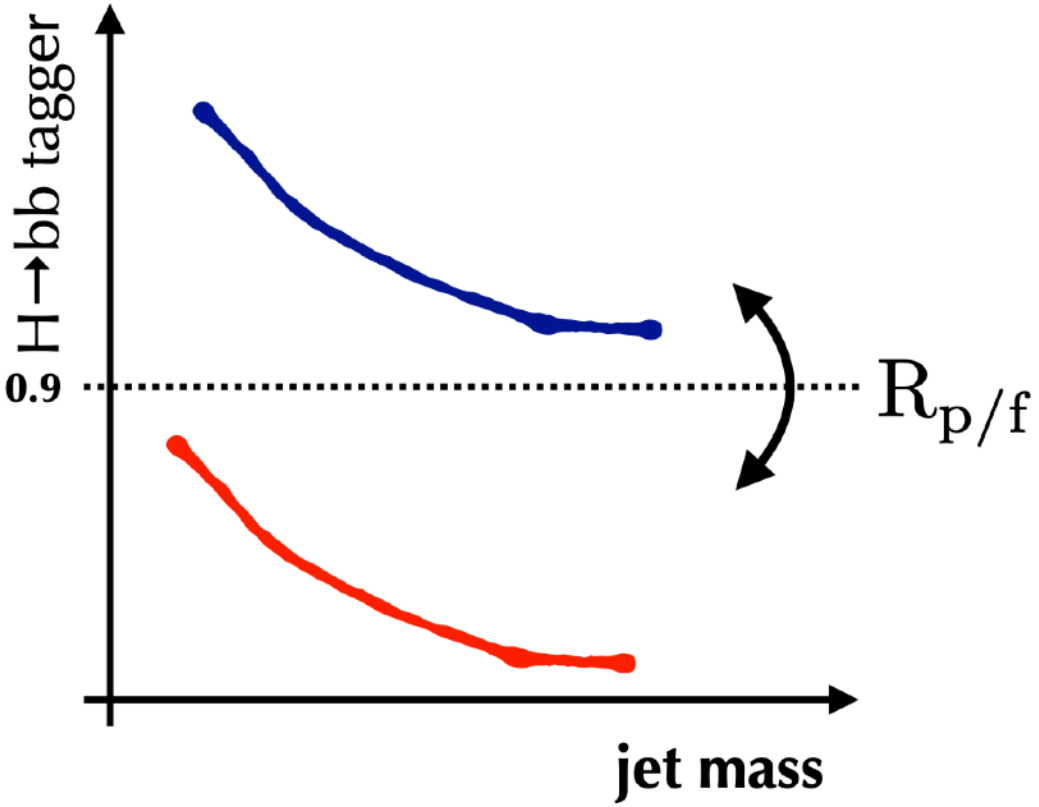
- ATLAS: trimming + muon in-jet correction
- CMS: soft drop mass



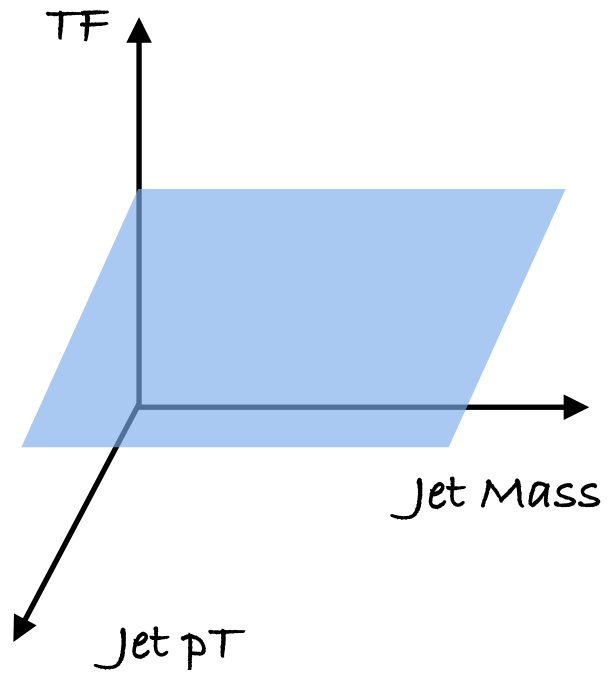
Multijets Background Estimation

Aim to predict the multijet mass distribution in the pass H→ bb tagger region using events in fail H→bb tagger region:

SR = CR ×transfer factor (TF)



TF as a function of Jet mass m and p_T



Multijets Background Estimation

Aim to predict the multijet mass distribution in the pass H→ bb tagger region using events in fail H→bb tagger region:

ATLAS EXPERIMENT

$$N^{\mathrm{SR}}_{\mathrm{QCD}}(p_T, m) = TF(p_T, m) \times N^{\mathrm{CR}}_{\mathrm{QCD}}(p_T, m)$$

$$N^{i}_{\mathrm{SR}} = N^{\mathrm{data},i}_{\mathrm{CR}}\, R^{\mathrm{MC},i}_{\mathrm{SR/CR}}\, TF(\rho^i)$$

To account for deviations, taylor expand TF in p_T and $\rho = log(m^2/p_T^2)$

Polynomial order determined via Fisher F-test

- 1 st in both p_T and ρ

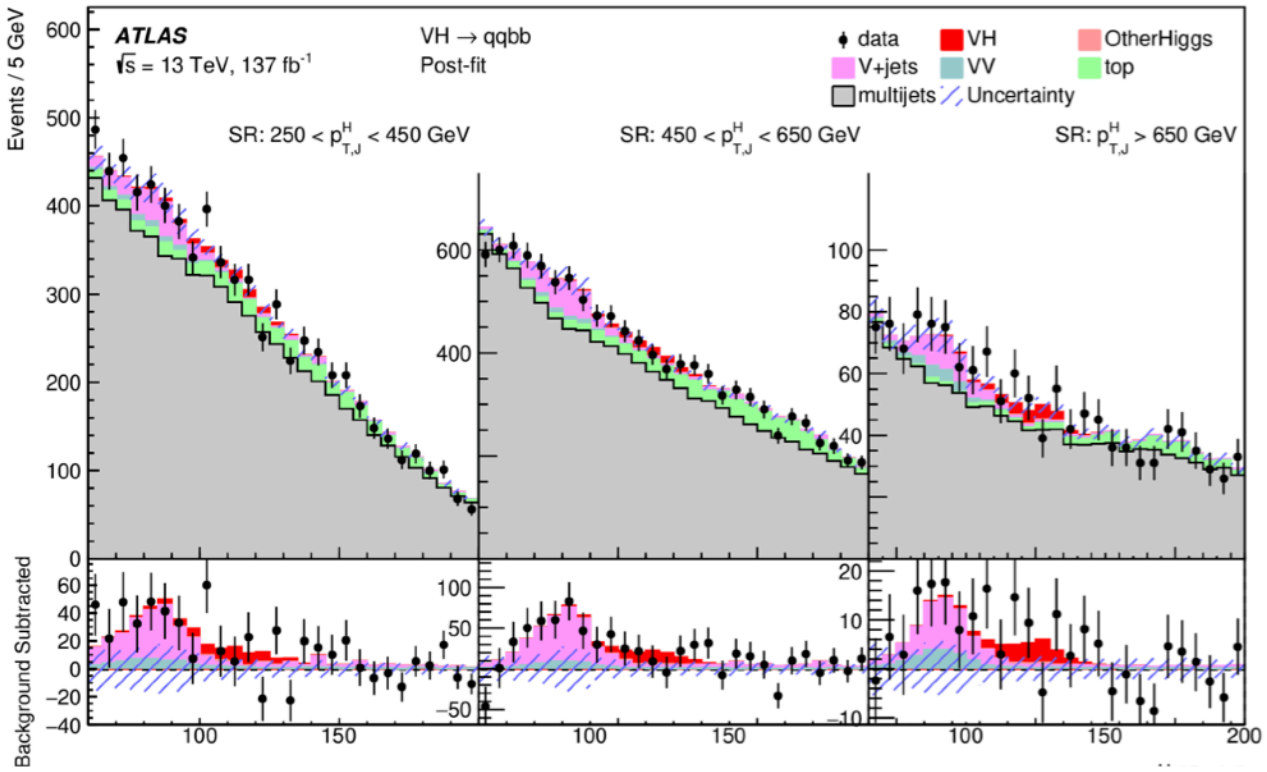
Use first order Bernstein polynomial

$$TF(\rho) = \sum_{k=0}^{n_\rho} a_k \left[b_{k,n_\rho}(\rho) \right]$$

$$b_{\nu,n}(x) = \binom{n}{\nu} x^{\nu} (1-x)^{n-\nu}$$

$$N^{\mathrm{QCD}}_{\mathrm{pass}}(p_{T_i}, m_j) = \sum_{k,l} \alpha_{kl} \rho^{kl}_{ij}\, p_{T_i} \times N^{\mathrm{QCD}}_{\mathrm{fail}}(p_T, m)$$

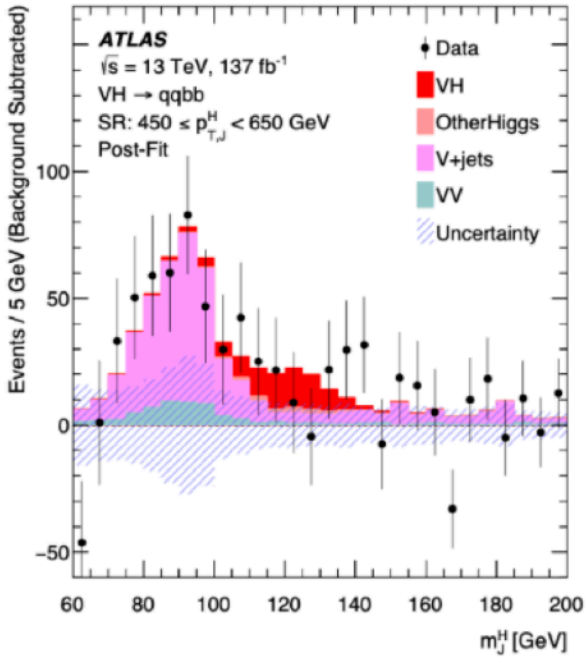
Results: ATLAS - 3 p_T categories



Kinematic region	Observed μ	Observed σ [fb]		Expected σ [fb]
$250 \le p_{\mathrm{T}}^{H} < 450$ GeV, $\lvert y_H \rvert < 2$	$0.8^{+2.2}_{-1.9}$	47^{+125}_{-109}	(<363)	57.0
$450 \le p_{\mathrm{T}}^{H} < 650$ GeV, $\lvert y_H \rvert < 2$	$0.4^{+1.7}_{-1.5}$	2^{+10}_{-9}	(< 24)	5.9
$p_{\mathrm{T}}^{H} \ge 650$ GeV, $\lvert y_H \rvert < 2$	$5.3^{+11.3}_{-3.2}$	6^{+13}_{-4}	(< 43)	1.2

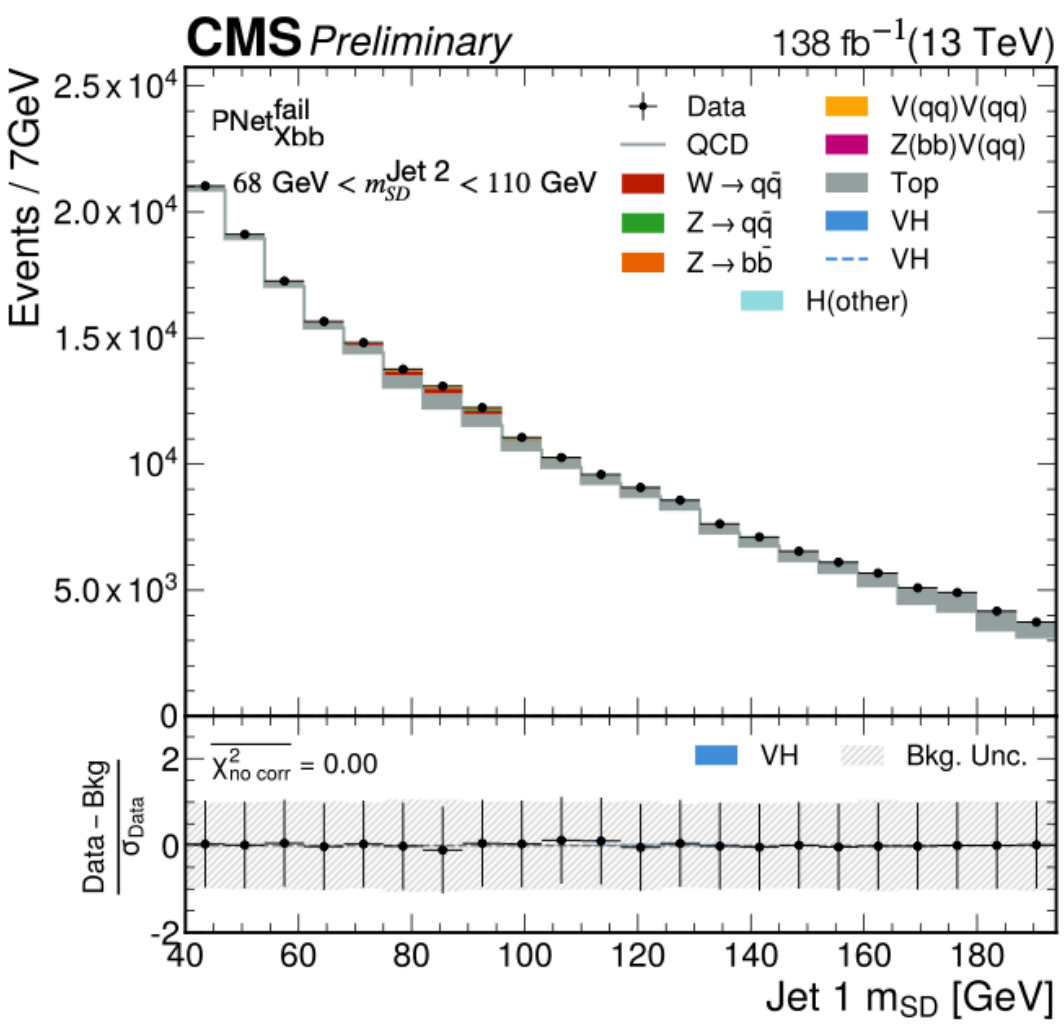
$$\mu_Z = 1.41^{+0.80}_{-0.58}$$

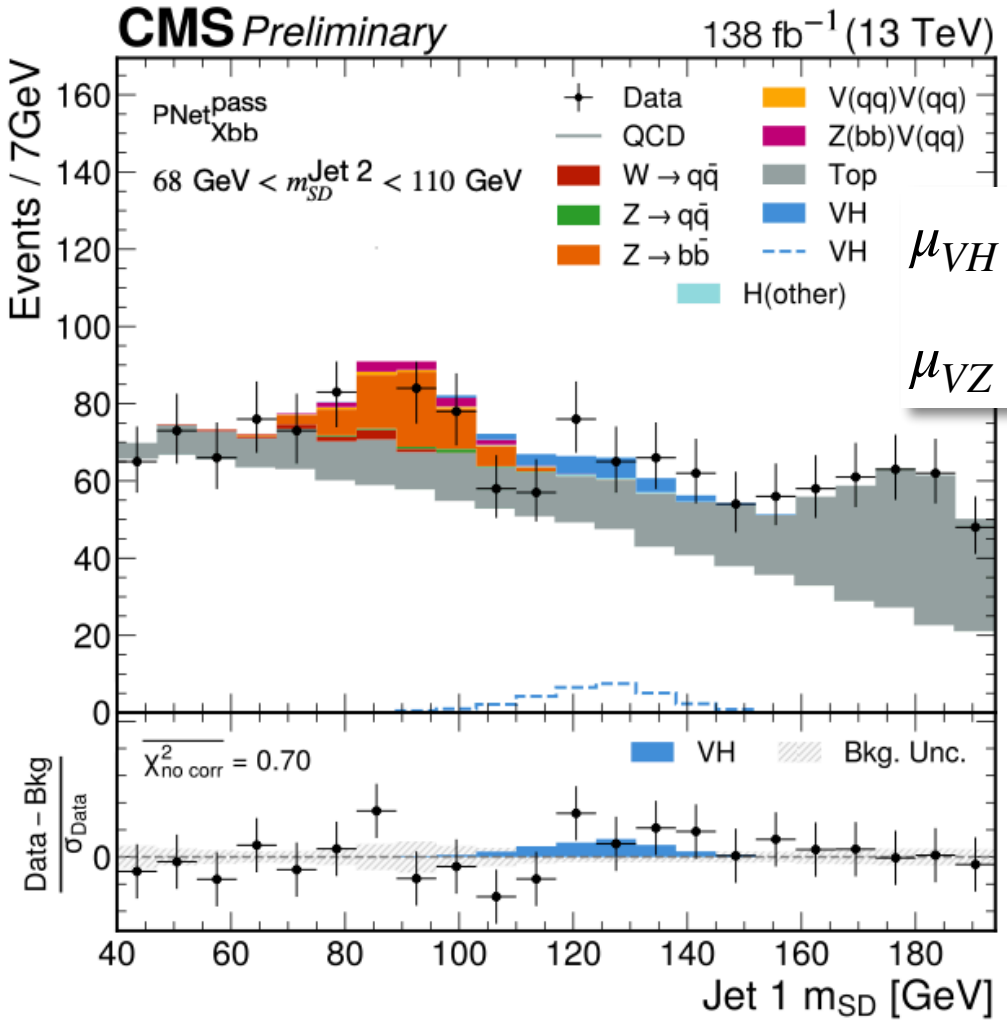
$$\mu_{VH} = 1.39^{+1.02}_{-0.88}$$



*the binning in pT is justified by our jet pT resolution (~80 GeV)

Results: CMS- inclusive in p_T but split in data taking conditions

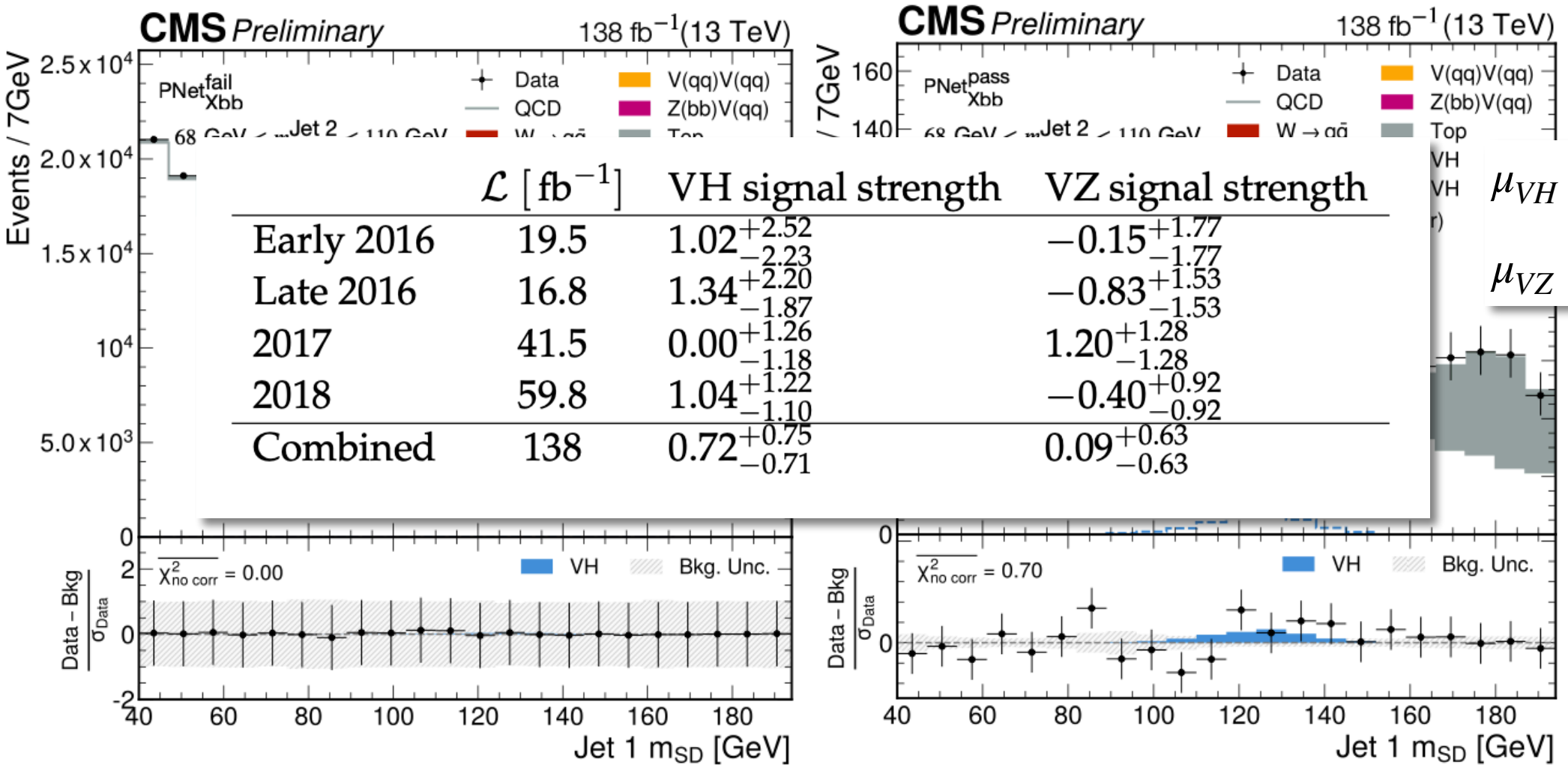




$$\mu_{VH} = 0.72^{+0.75}_{-0.71}$$

$$\mu_{VZ} = 0.09^{+0.63}_{-0.63}$$

Results: CMS- inclusive in p_T but split in data taking conditions



	$\mathcal{L}$ [fb^{-1}]	VH signal strength	VZ signal strength
Early 2016	19.5	$1.02^{+2.52}_{-2.23}$	$-0.15^{+1.77}_{-1.77}$
Late 2016	16.8	$1.34^{+2.20}_{-1.87}$	$-0.83^{+1.53}_{-1.53}$
2017	41.5	$0.00^{+1.26}_{-1.18}$	$1.20^{+1.28}_{-1.28}$
2018	59.8	$1.04^{+1.22}_{-1.10}$	$-0.40^{+0.92}_{-0.92}$
Combined	138	$0.72^{+0.75}_{-0.71}$	$0.09^{+0.63}_{-0.63}$

$$\mu_{VH} = 0.72^{+0.75}_{-0.71}$$

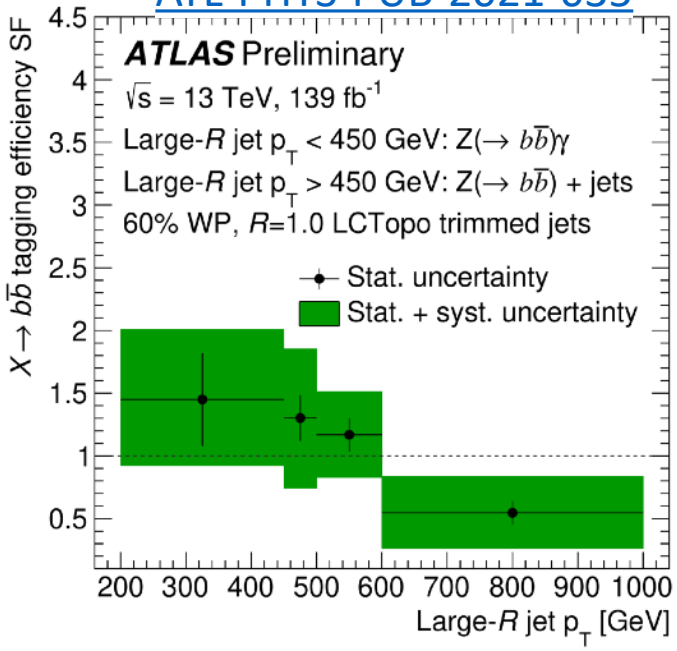
$$\mu_{VZ} = 0.09^{+0.63}_{-0.63}$$

Summary

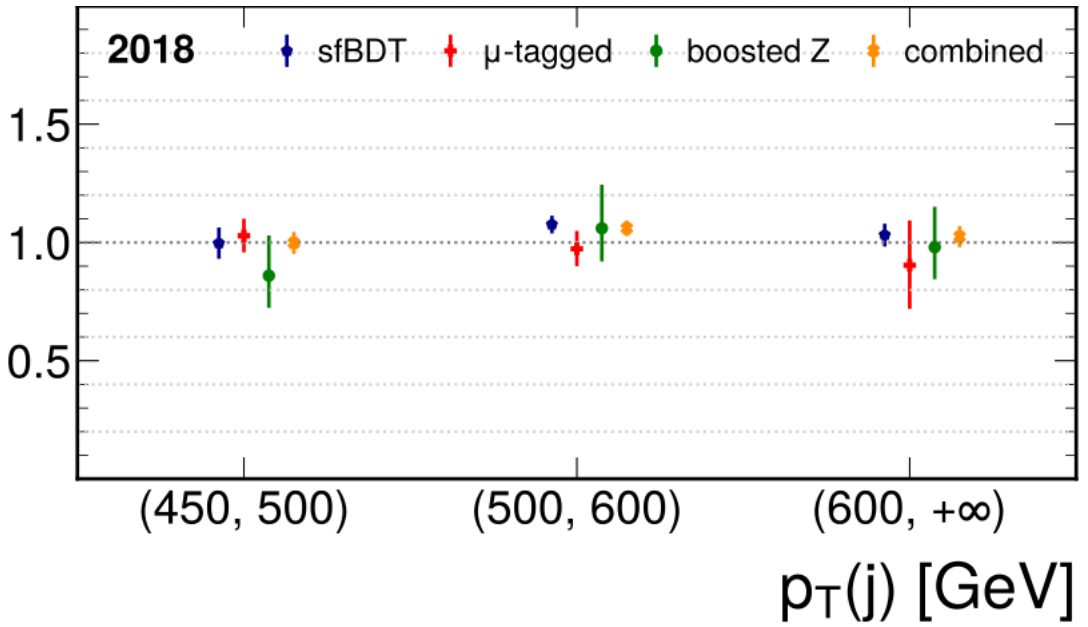
Both ATLAS and CMS have measured V(qq)H production at high p_T

- Dominant systematic uncertainty comes from the multijet data-driven background estimate - this can be improved with more data
- H→ $b\bar{b}$ tagger scale factors (**impact: $\delta\mu$=0.5**) is also one of the dominant systematics in ATLAS

ATL-PHYS-PUB-2021-035



CMS-PAS-BTV-22-001

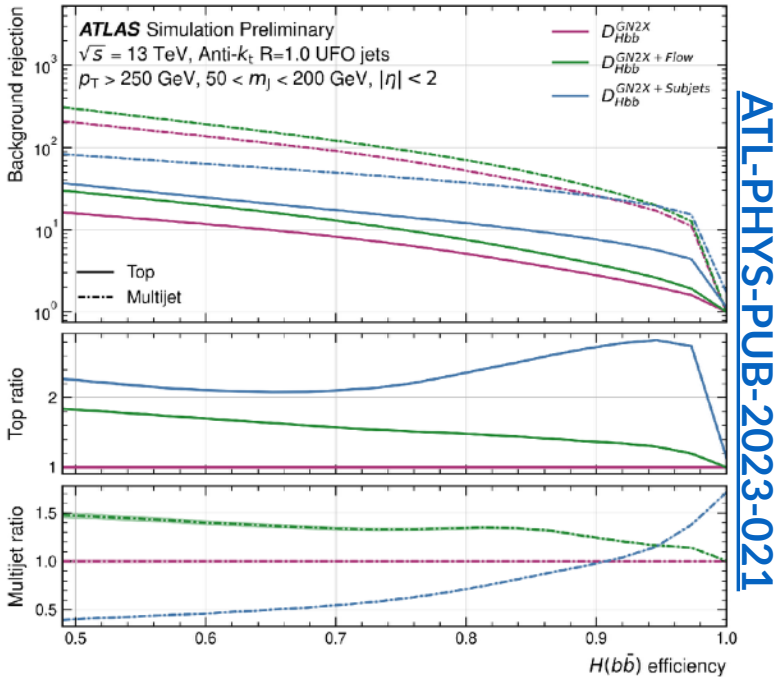


*use sfBDT techniques to calibrate H→$b\bar{b}$ tagger in g → $b\bar{b}$ sample, achieving smaller uncertainty

Summary

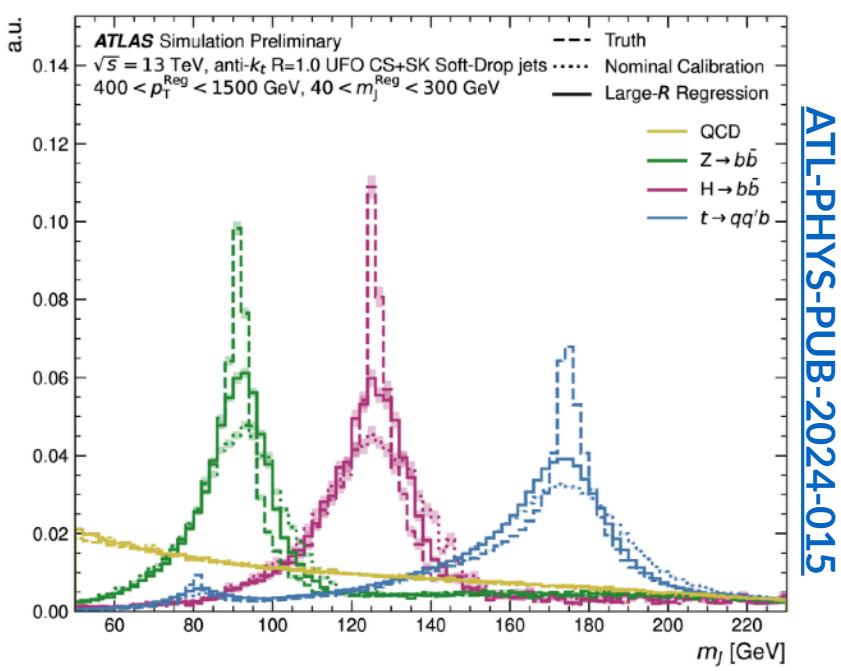
Future analyses will benefit from improved boosted Higgs and vector-boson taggers and from better large-R jet mass resolution.

New H→ $b\bar{b}$ tagger in ATLAS

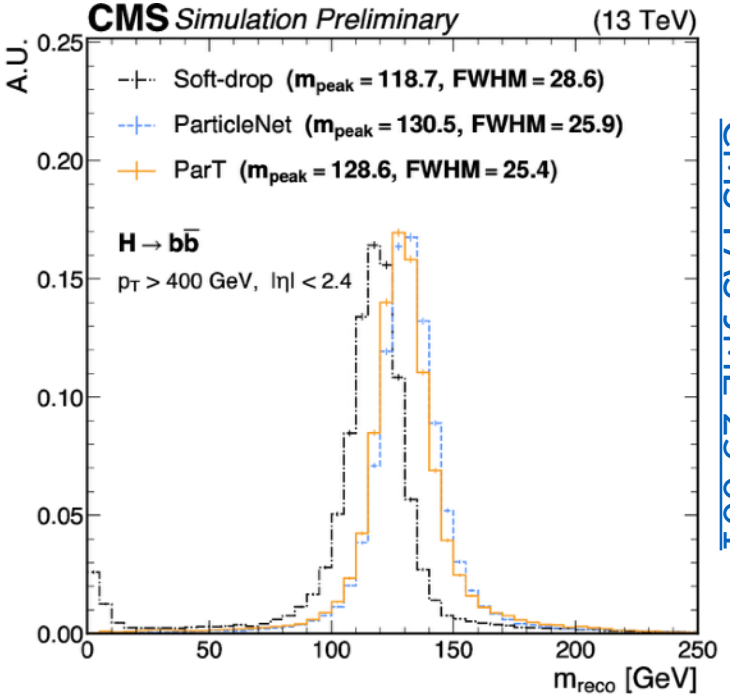


ATL-PHYS-PUB-2023-021

Improvements on Higgs mass resolution in ATLAS and CMS



ATL-PHYS-PUB-2024-015



CMS-PAS-JME-25-001

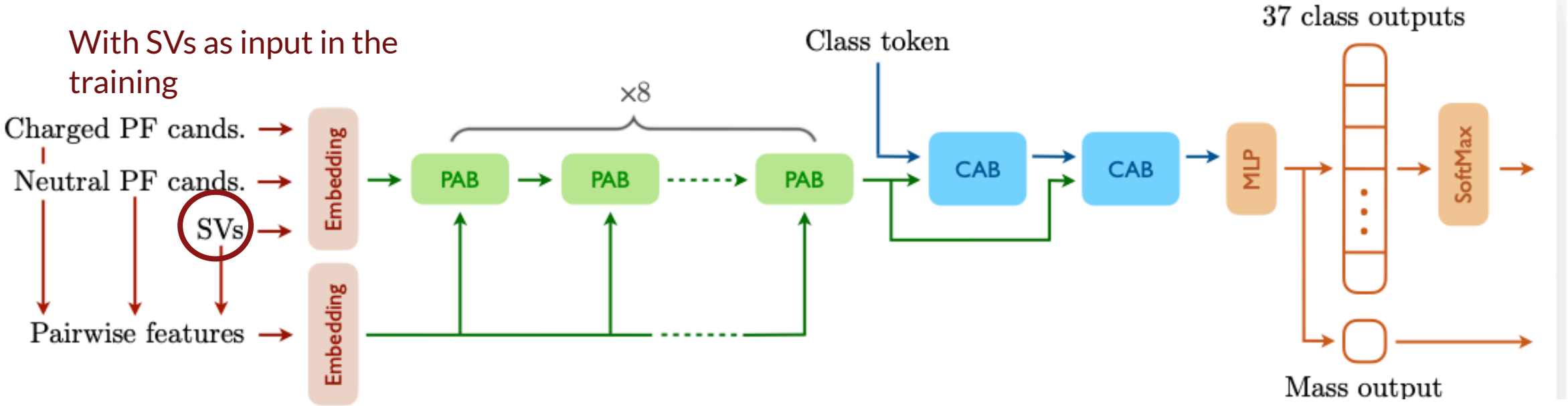
*Achieve 60% H→bb efficiency, mistag rate ~0.5%

Bonus

Input and Output

CMS-PAS-JME-25-001

With SVs as input in the training



$$\mathcal{L}_{\log-\cosh}(m_{\text{true}}, m_{\text{pred}}) = \ln(\cosh(m_{\text{pred}} - m_{\text{true}}))$$

Charged particles, neutral particles, and vertices as inputs

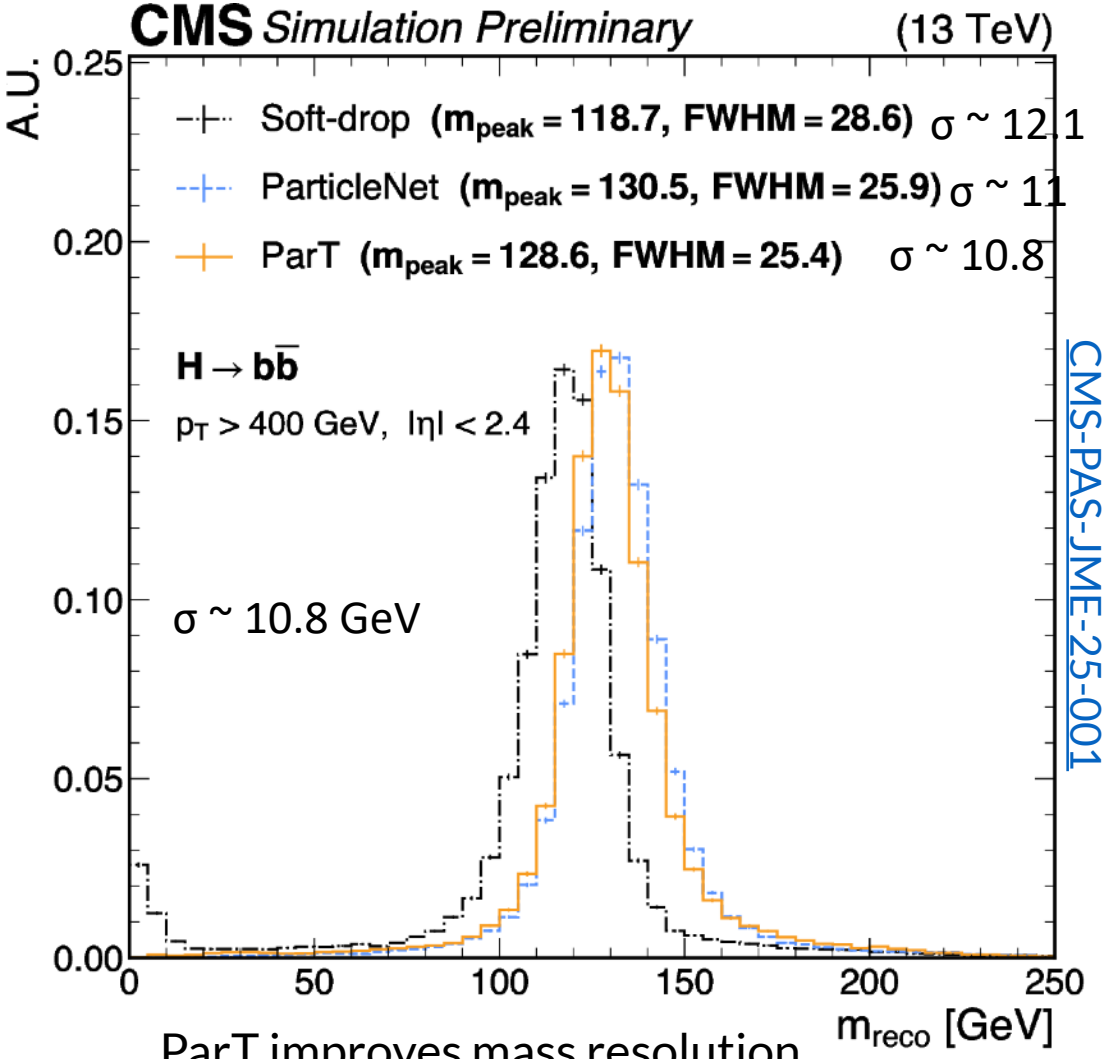
- Kinematic, charge, PID, and track-related features for each
- Outputs 37 class probabilities (cross-entropy loss) + **mass regression (log-cosh)**

With mass regression included in the training

* PAB: particle attention blocks

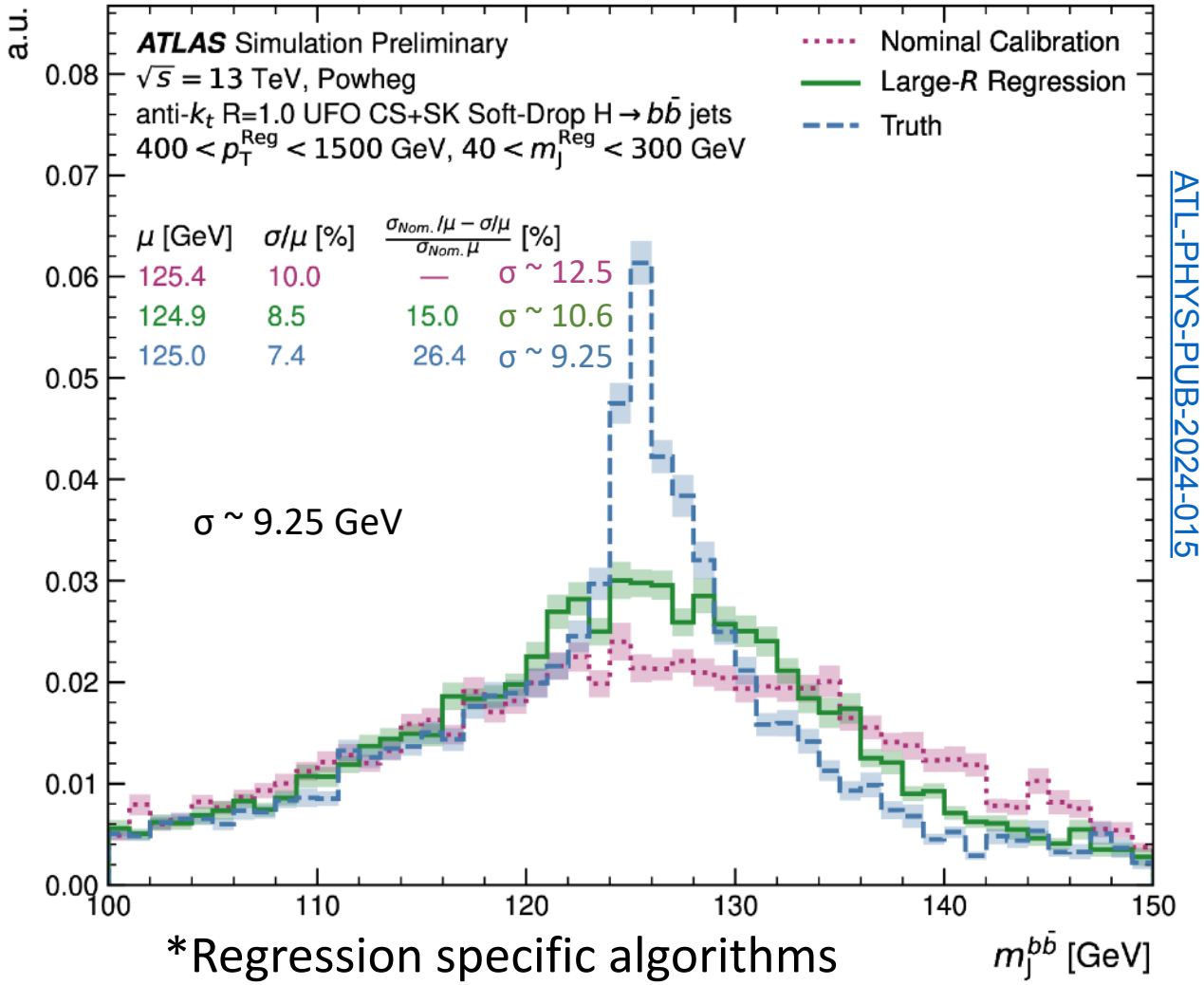
* CAB: class attention blocks

Comparison for resolution improvements



CMS-PAS-JME-25-001

ParT improves mass resolution



ATL-PHYS-PUB-2024-015

*Regression specific algorithms

Out perform parT in mass resolution