Image credit: Marguerite Tonjes

Boosted jets and substructures for Higgs physics

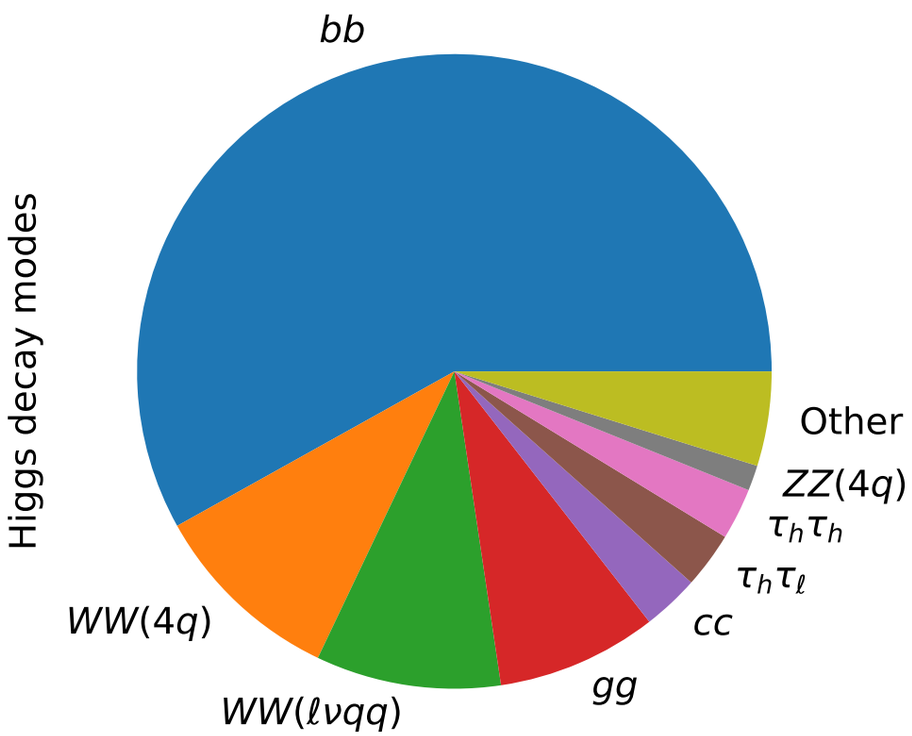
Nick Smith, on behalf of ATLAS+CMS

Higgs 2022

9 November, 2022

Outline

- History: boosted Higgs as a path to H(bb) observation?
- Developments in jet substructure
  - Mass grooming & flavor tagging
- Enabling a diverse set of Higgs and di-Higgs analyses
  - VH(bb)
  - ggH(bb)
  - VH(cc)
  - ggH(cc)
  - HH(bb bb)
  - Multi-boson resonances

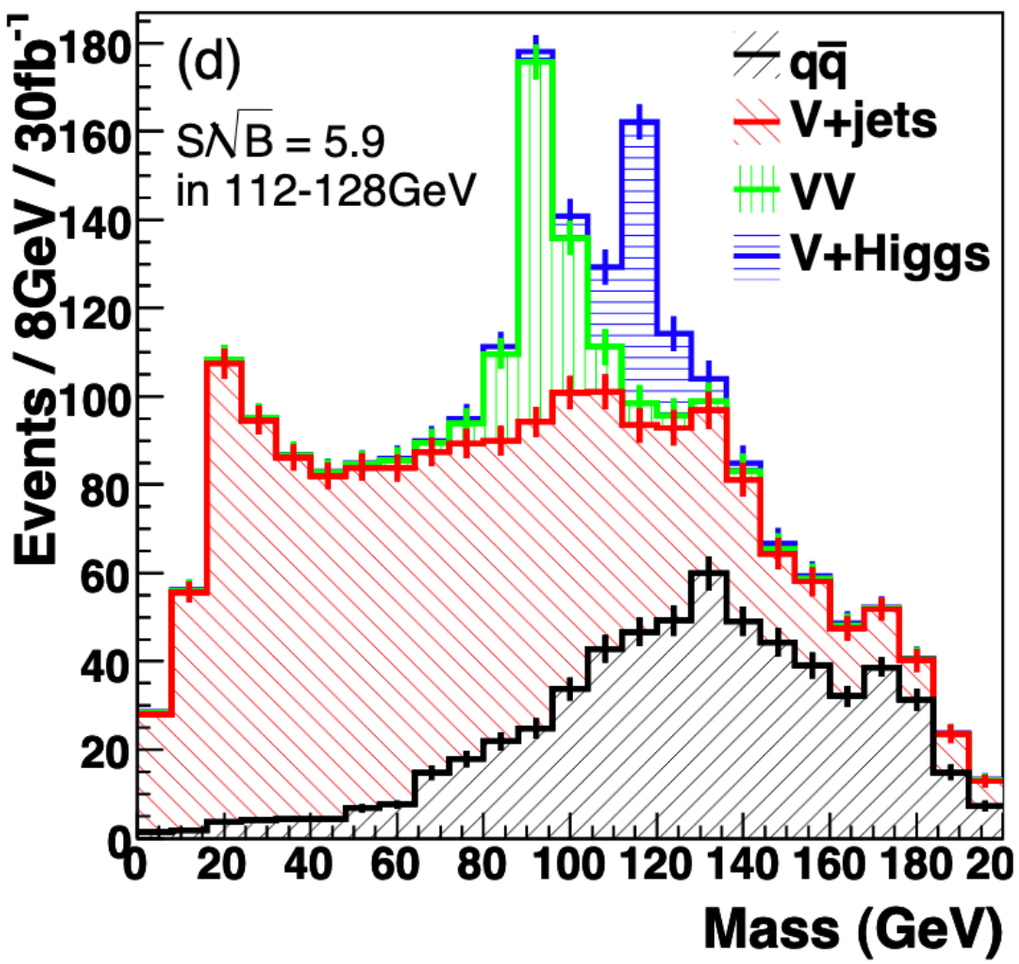


Popular for some time

- Inaugural BOOST 2009

UCL

Combined particle-level result



- Note excellent Z peak for calibration
- 5.9 σ; potentially very competitive
- bb branching information critical for extracting Higgs properties
  - *"Measuring the Higgs sector" Lafaye, Plehn, Rauch, D.Zerwas, Duhrssen, arXiv:0904.3866 [hep-ph]*
- Studies within ATLAS are promising and nearly public.

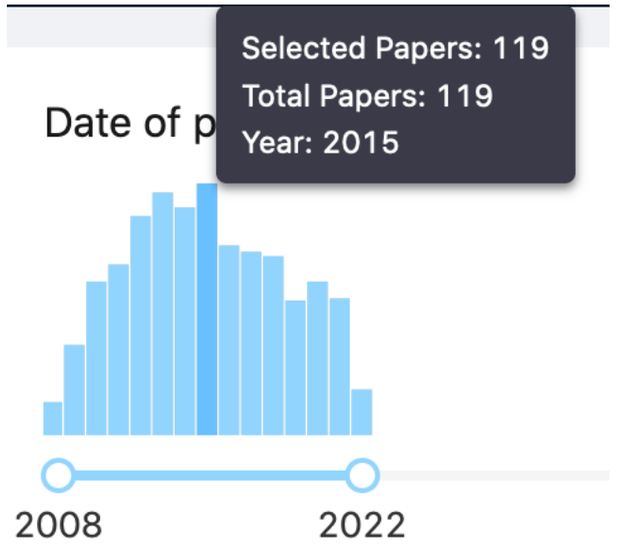
10/7/09 JMB, Boost 2009 Stanford 33

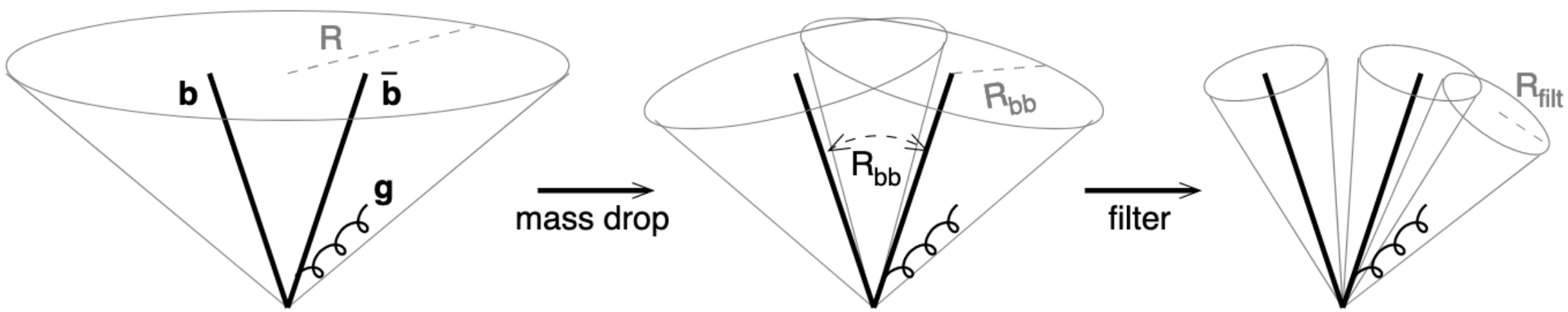
https://indico.cern.ch/event/56985/#3-why-subject-structure-works

The original plan

- Trigger on leptonic V decay
- Cluster into Cambridge-Aachen R=1.2 jet
  - Higgs $p_T > 200$ GeV
- Decluster until mass drop
  - If subjets are b-tagged, call it a Higgs
- 6 ➜ 3σ in ATLAS full-sim follow up (cds:1201444)
  - For 30/fb @ 14 TeV

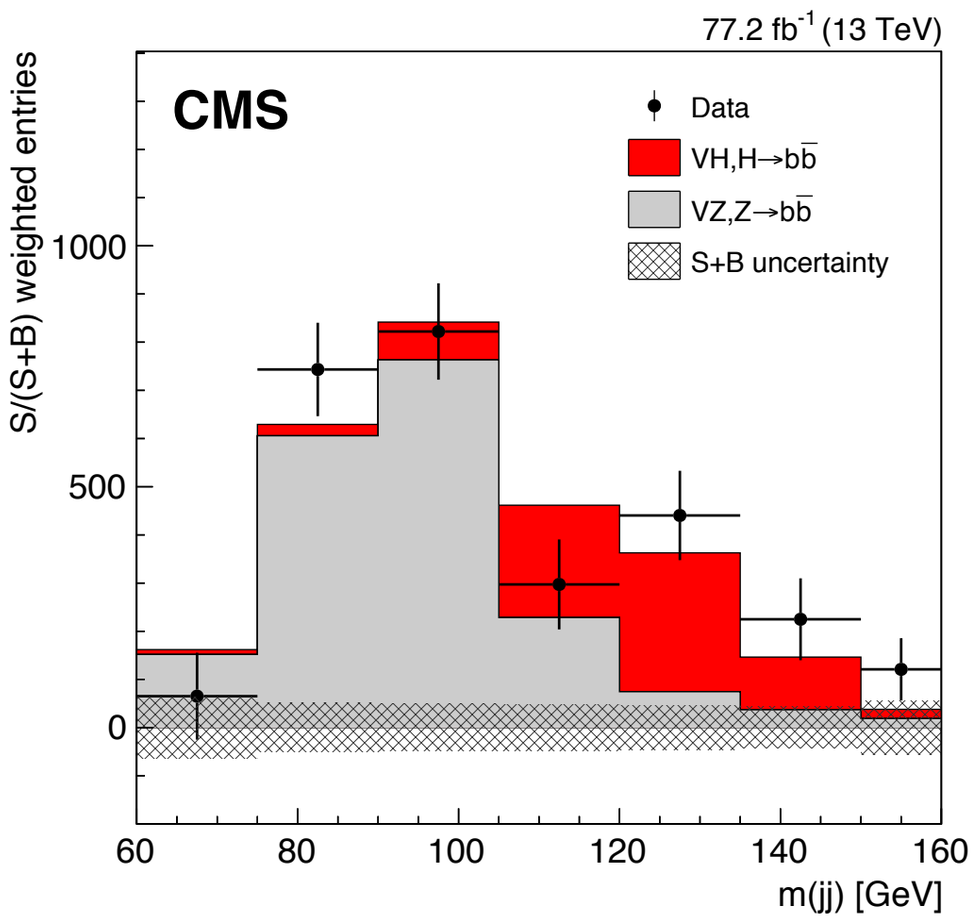
arxiv:0802.2470

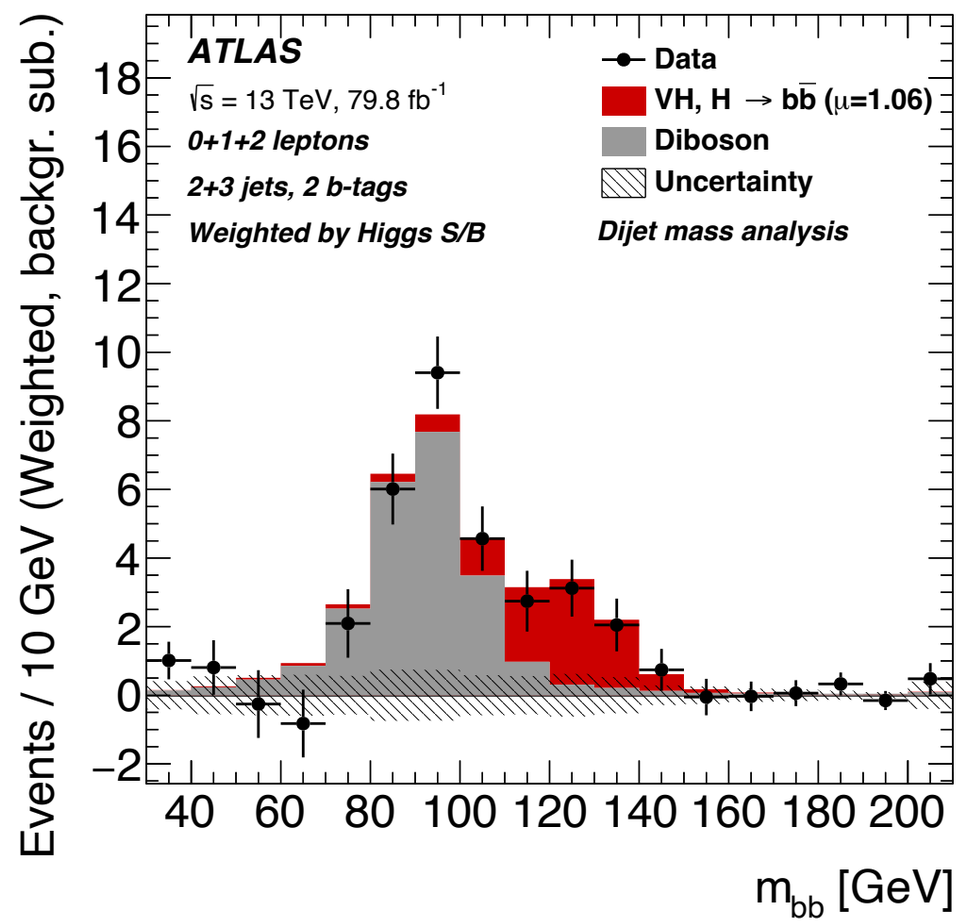




A decade later

- VH(bb) observation 2018
  - ATLAS arxiv:1808.08238 CMS arxiv:1808.08242
  - Neither used a declustering technique, rather Anti-k_T R=0.4 (AK4) dijet system
    - Though $p_T^V > 150$ GeV*
    - CMS: gluon radiation recovery



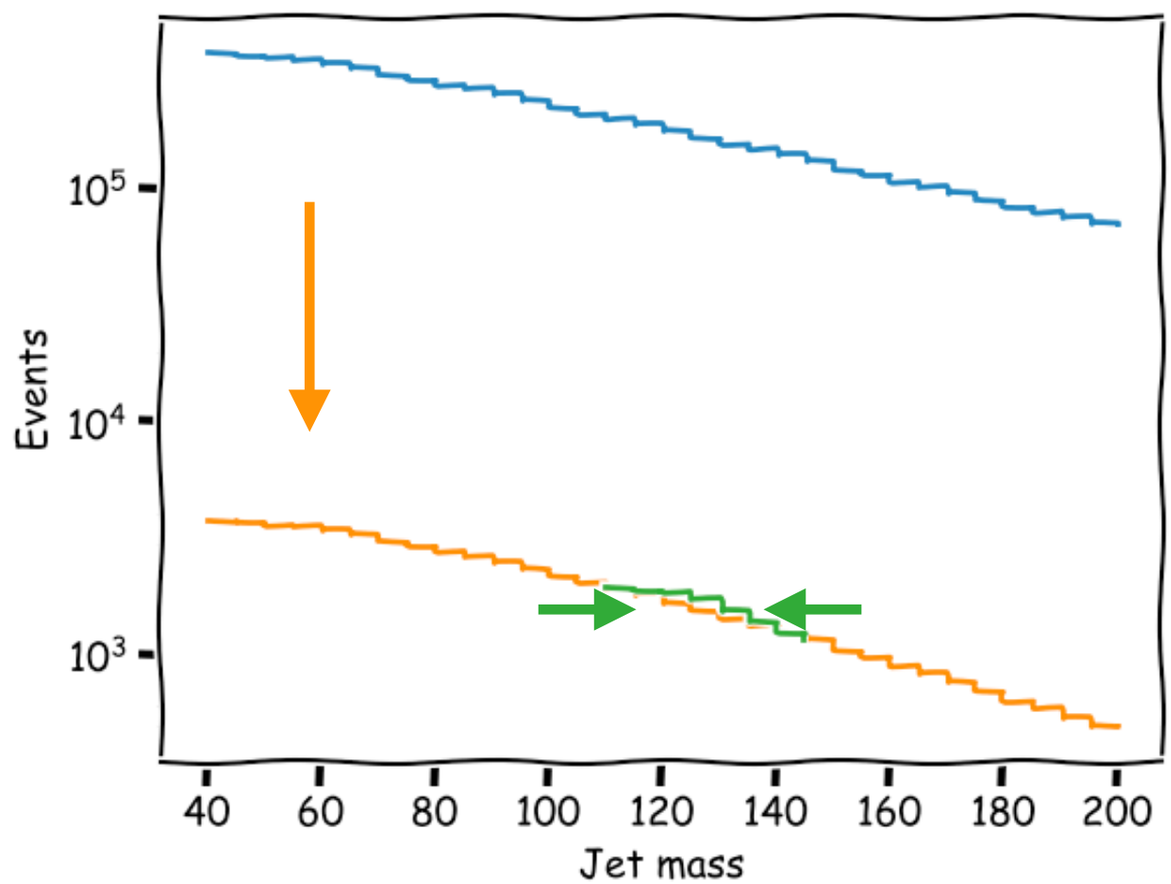


* Z(ll) channel has additional 50(75)-150 GeV bin

Meanwhile

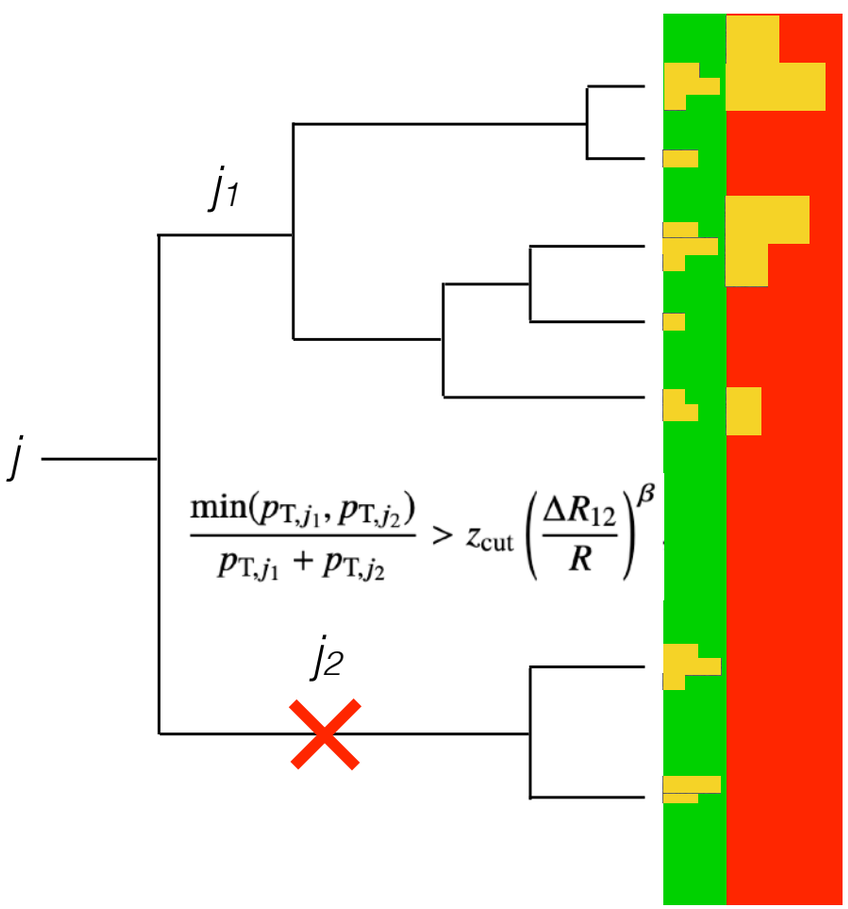
- 14 years of development in jet substructure
  - Latest usually found at BOOST
- Main products for our purposes:
  - Jet mass grooming
  - Jet flavor tagging

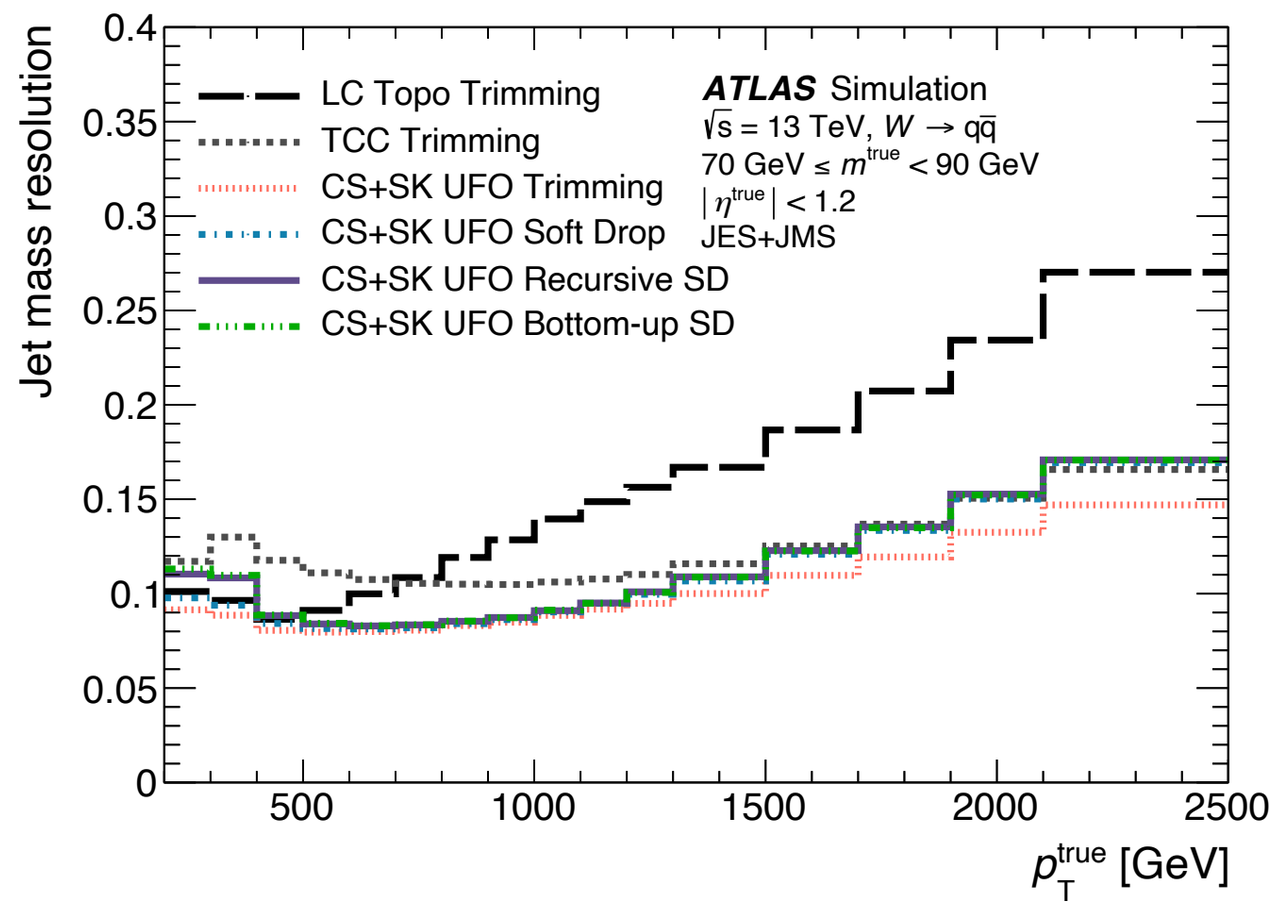
- SLAC (2009)
- Oxford (2010)
- Princeton (2011)
- Valencia (2012)
- Flagstaff(2013)
- London (2014)
- Chicago (2015)
- Zurich (2016)
- Buffalo (2017)
- Paris (2018)
- MIT (2019)
- Online (Hamburg) (2020)



Jet mass

- ATLAS & CMS converging on approach, mostly
  - Anti-k_T, radius 0.8-1.5
  - Pileup suppression: CS+SK vs. Puppi
  - Grooming: soft-drop β=0 (CMS) or 1 (ATLAS Run 3), z=0.1

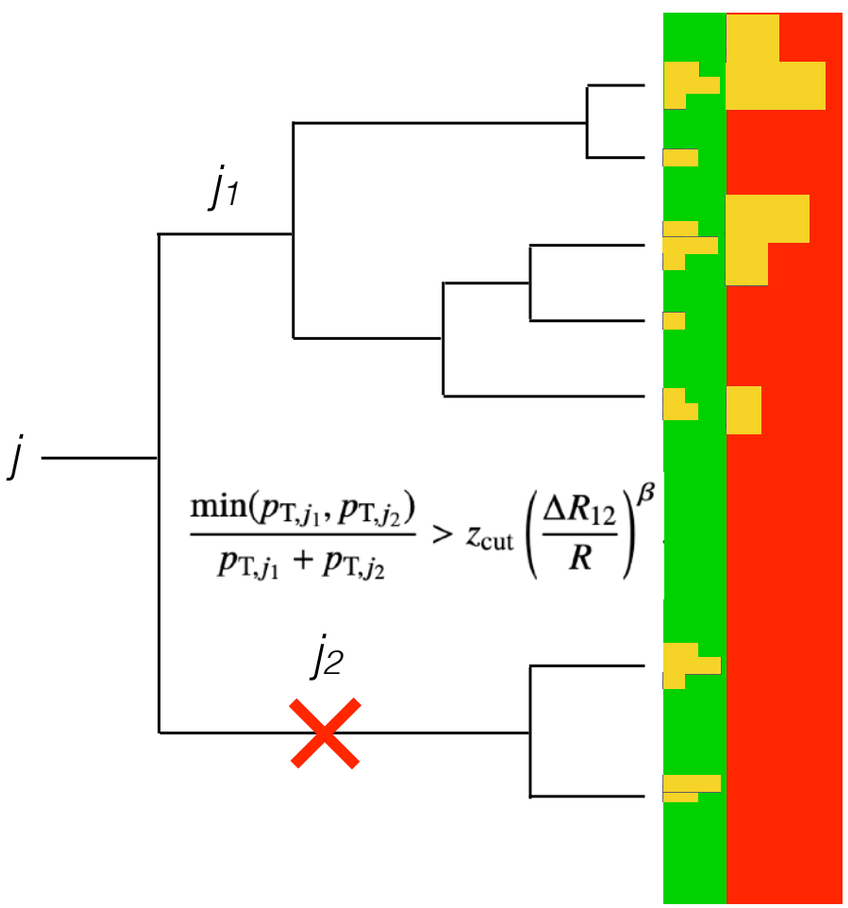


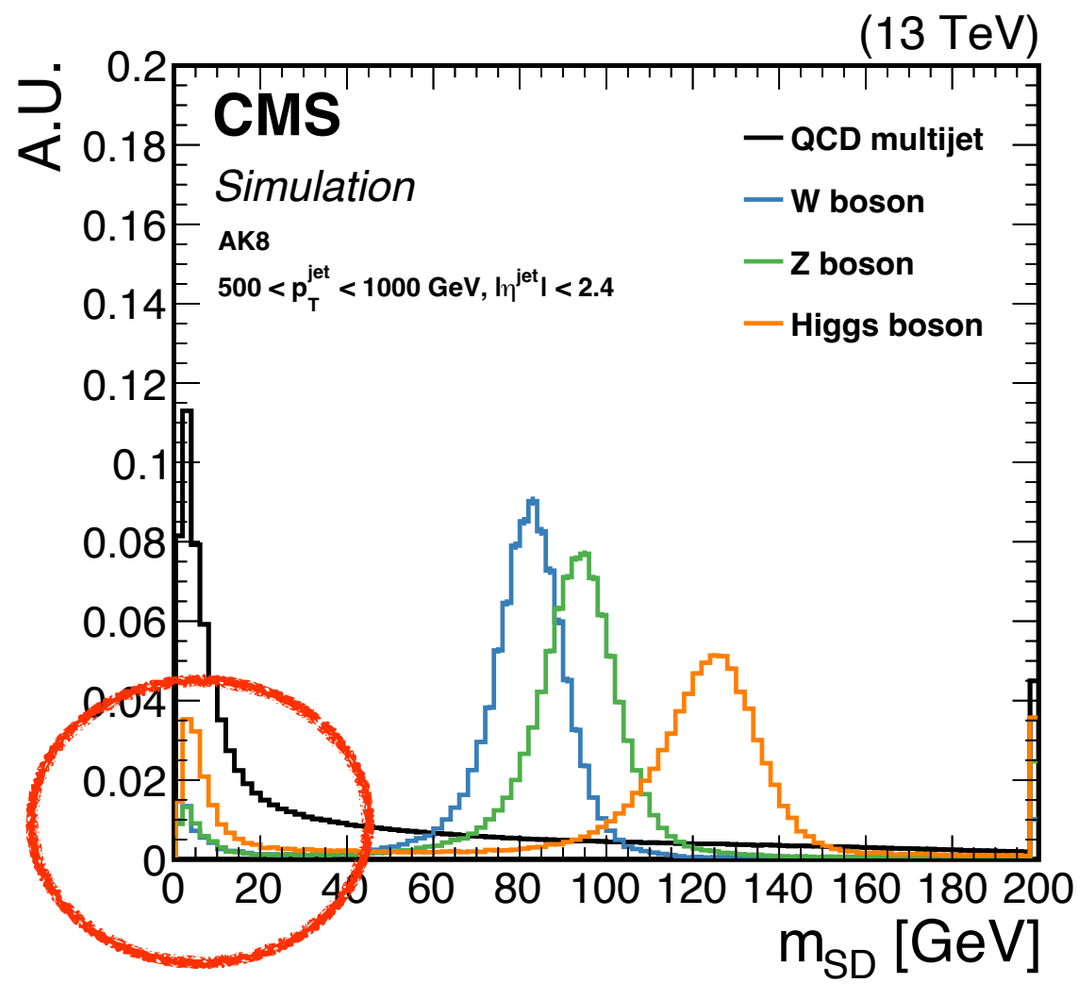


arxiv:2009.04986

Jet mass

- ATLAS & CMS converging on approach, mostly
  - Anti-k_T, radius 0.8-1.5
  - Pileup suppression: CS+SK vs. Puppi
  - Grooming: soft-drop β=0 (CMS) or 1 (ATLAS Run 3), z=0.1
    - Can send mass to zero



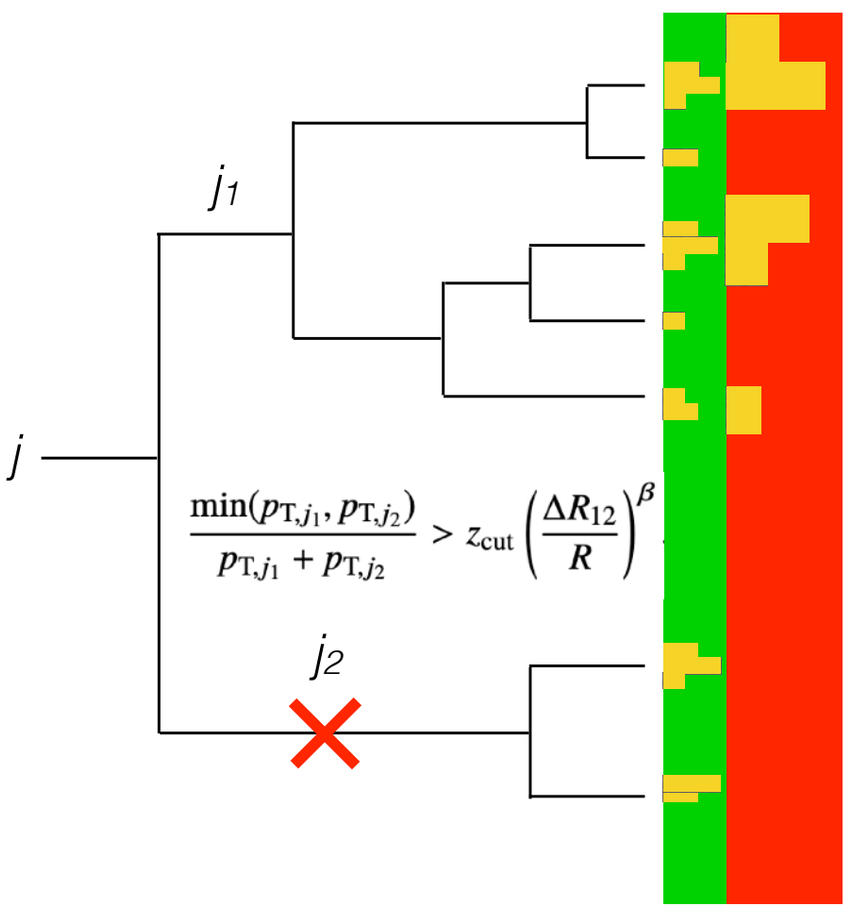


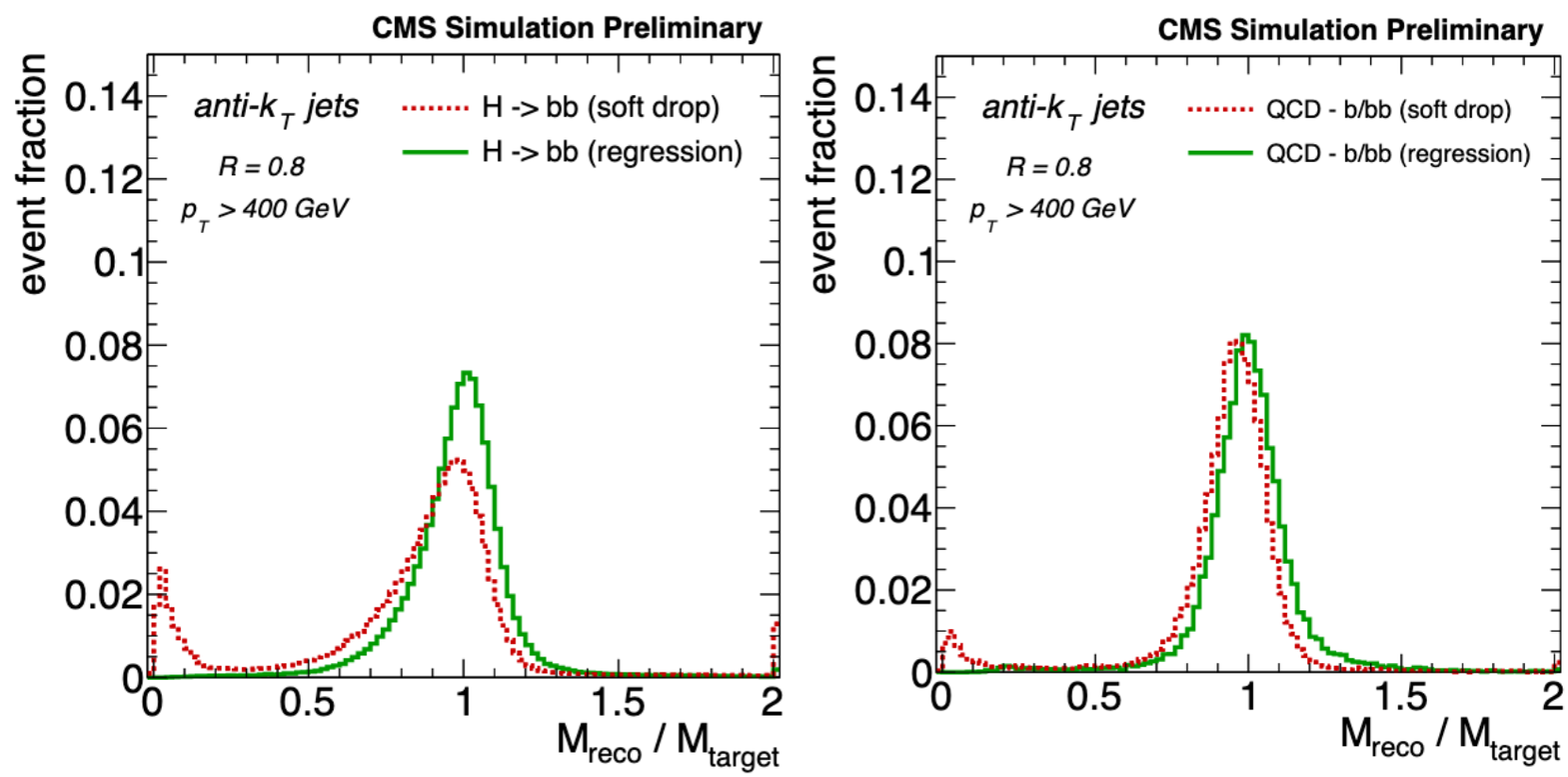
arxiv:2004.08262

Jet mass

- ATLAS & CMS converging on approach, mostly
  - Anti-k_T, radius 0.8-1.5
  - Pileup suppression: CS+SK vs. Puppi
  - Grooming: soft-drop β=0 (CMS) or 1 (ATLAS Run 3), z=0.1
    - Can send mass to zero, recover with ParticleNet regression

CMS-DP-2021-017



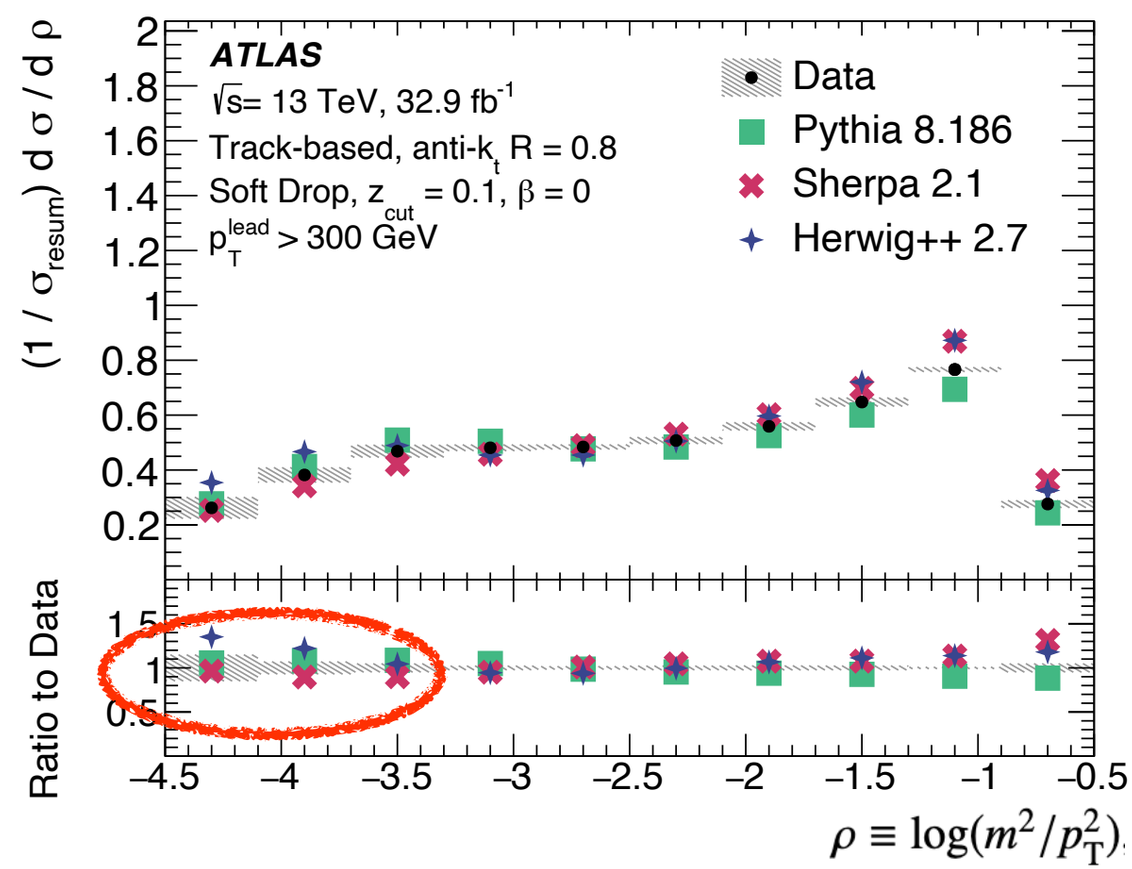


$$M_{\text{target}} = \begin{cases} M_{\text{SD}}^{\text{gen}} & \text{if jet is QCD} \\ m_{\text{H}} \in [15, 250] & \text{otherwise} \end{cases}$$

Jet mass

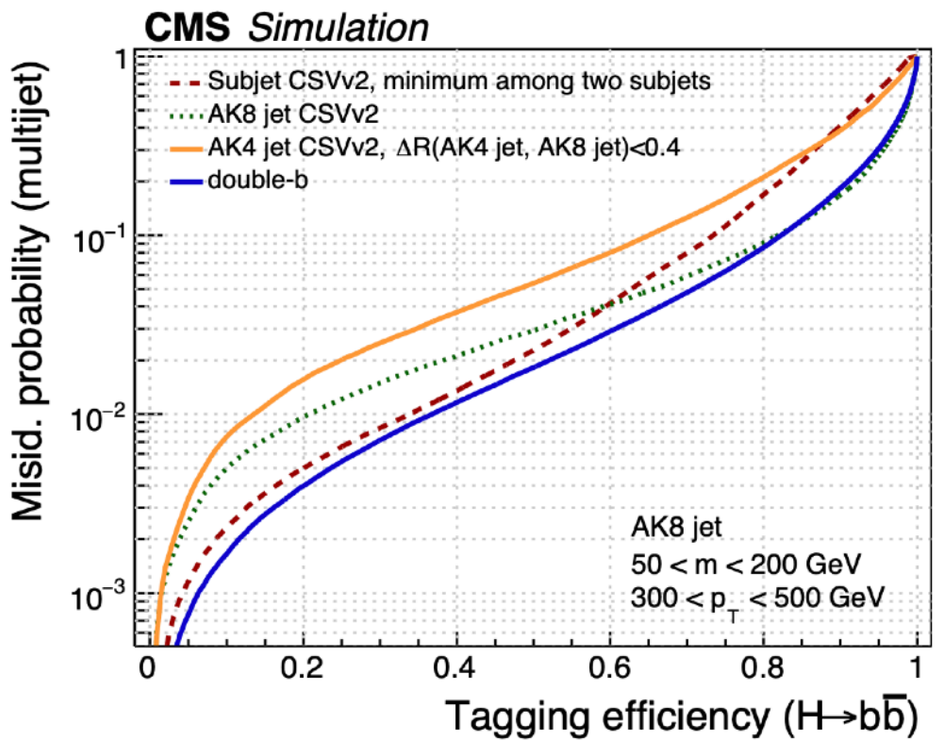
- ATLAS & CMS converging on approach, mostly
  - Anti-k_T, radius 0.8-1.5
  - Pileup suppression: CS+SK vs. Puppi
  - Grooming: soft-drop β=0 (CMS) or 1 (ATLAS Run 3), z=0.1
    - Can send mass to zero, recover with ParticleNet regression
    - Well-modeled observable? Yes except at very high boost

- ATLAS arxiv:1912.09837
  - Soft-drop (1402.2657) observables
  - "All of the parton shower … calculations provide an excellent description of the data in **most** regions of phase space."

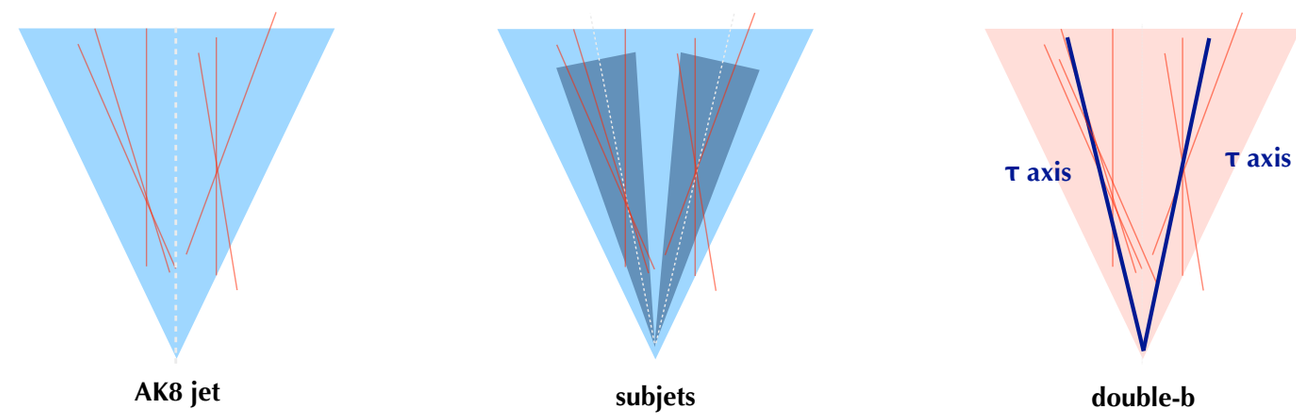


Jet flavor

- Whole large-R jet tagger wins over sub-jet tag
  - Tracks, secondary vertices along subjet axis

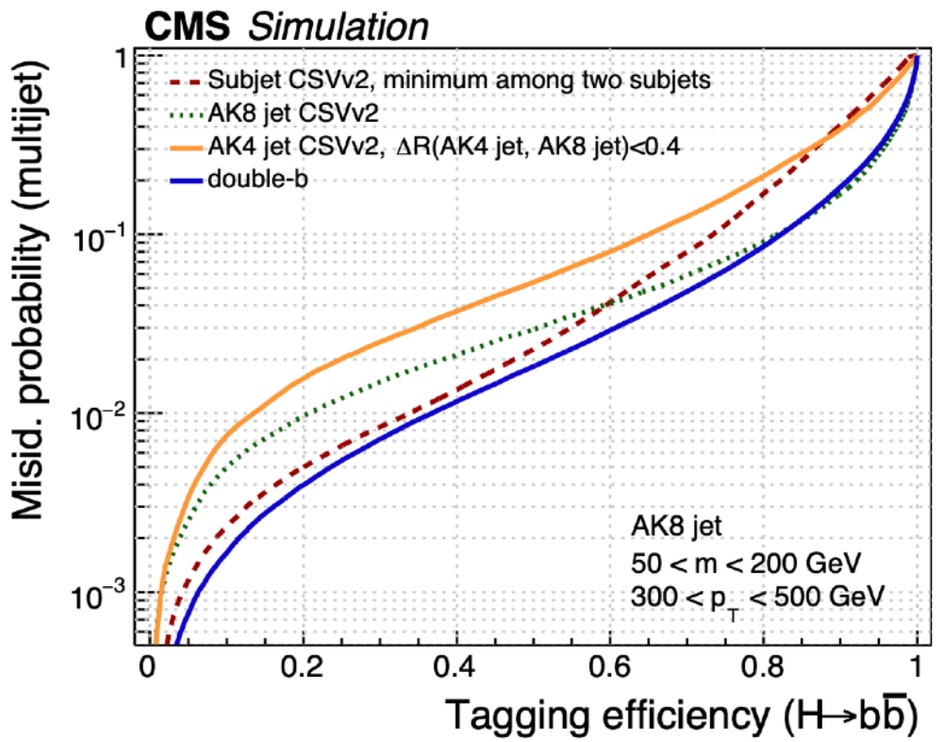


arxiv:1712.07158

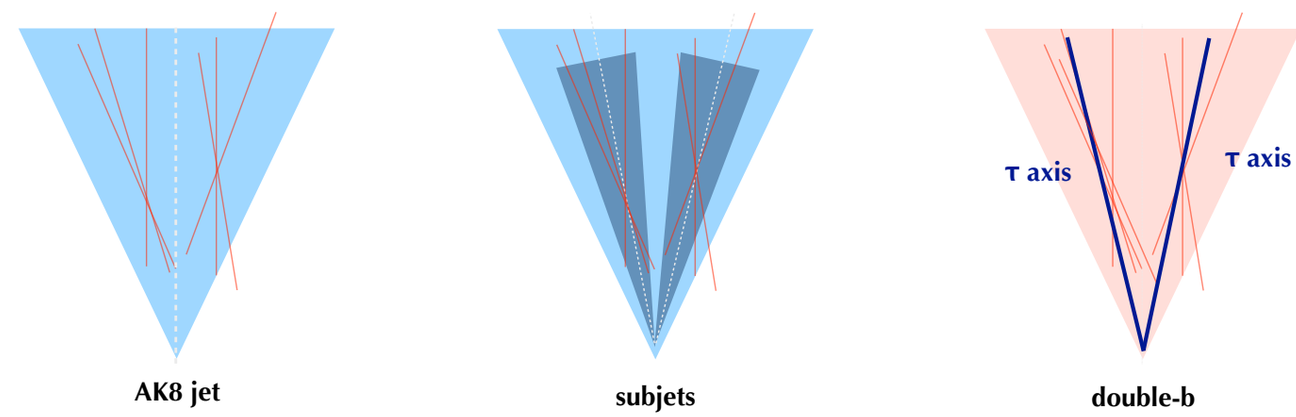


Jet flavor

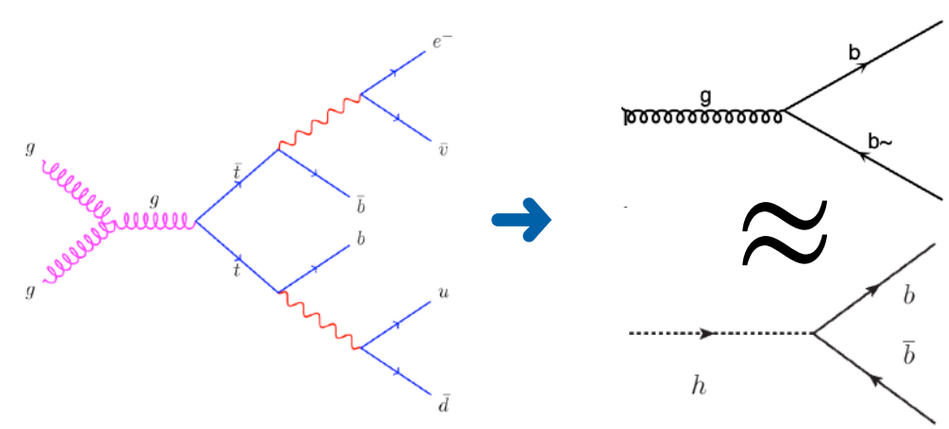
- Whole large-R jet tagger wins over sub-jet tag
  - Tracks, secondary vertices along subjet axis



arxiv:1712.07158

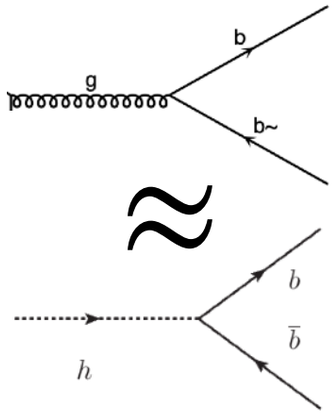


- Cannot calibrate using top tag & probe
  - Calibration through gluon splitting proxy jet
    - Tag via soft muons from weak B hadron decay

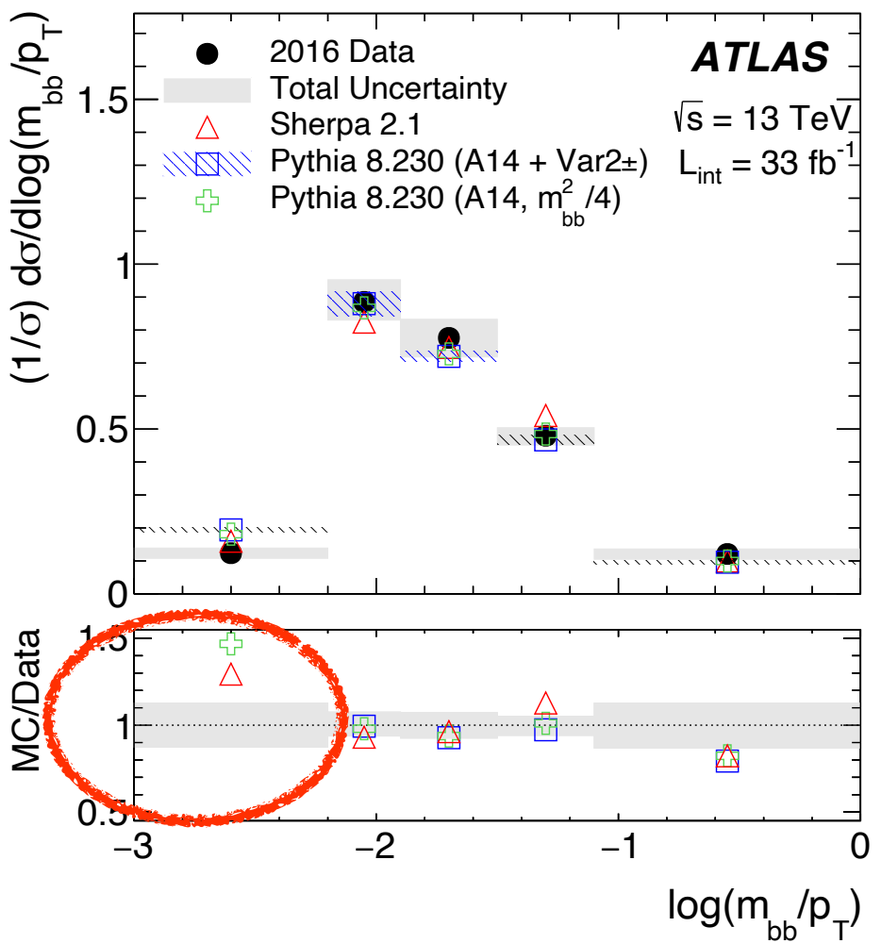


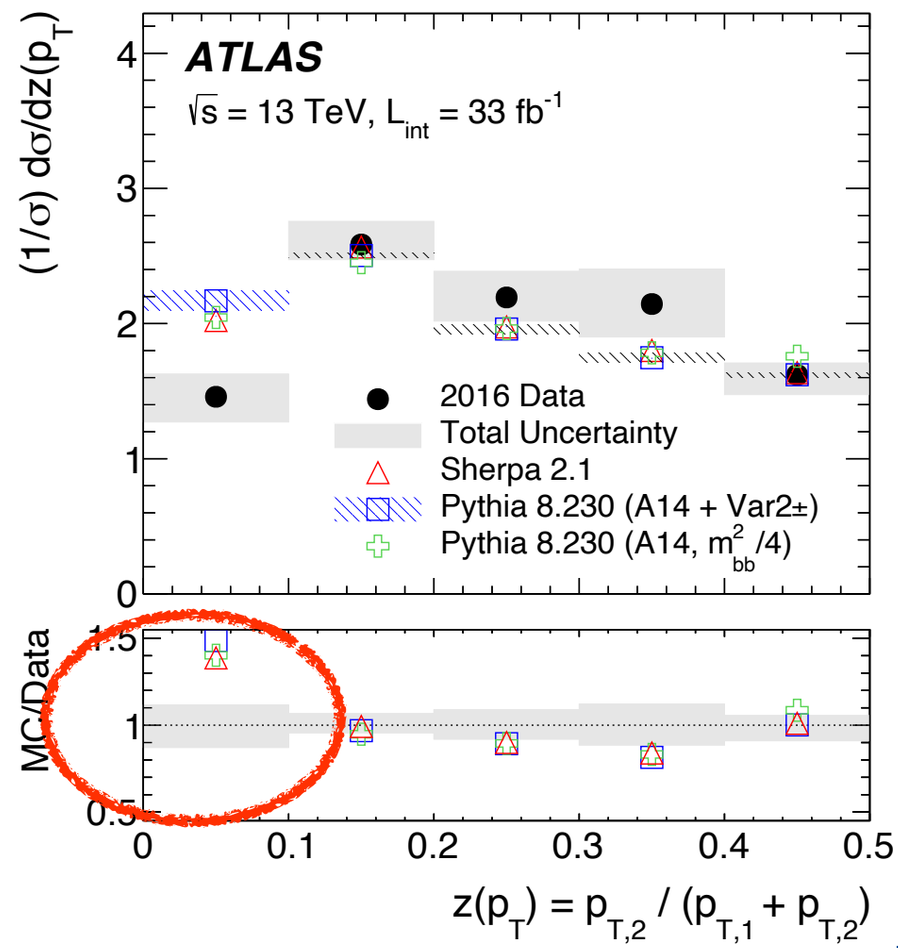
Jet flavor

- Gluon splitting modeling ok except for
  - Low ρ
    - High boost, important for EFT
  - High momentum asymmetry
    - Soft-drop: z > 0.1
- Common theme: use Z as validation (later)



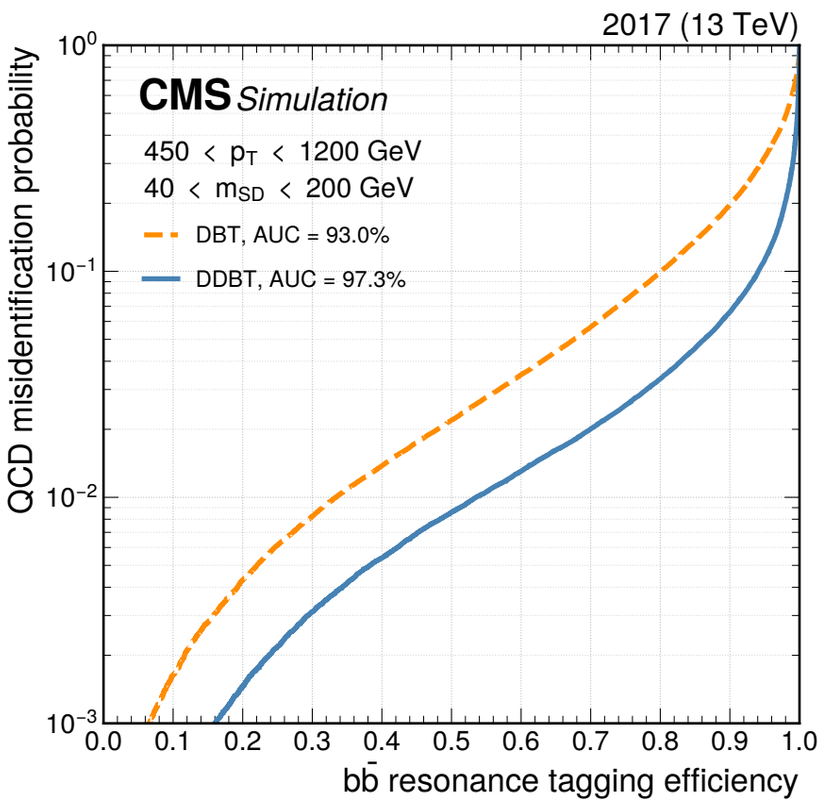
arxiv:1812.09283





Jet flavor

- Big gains with more complex DNN

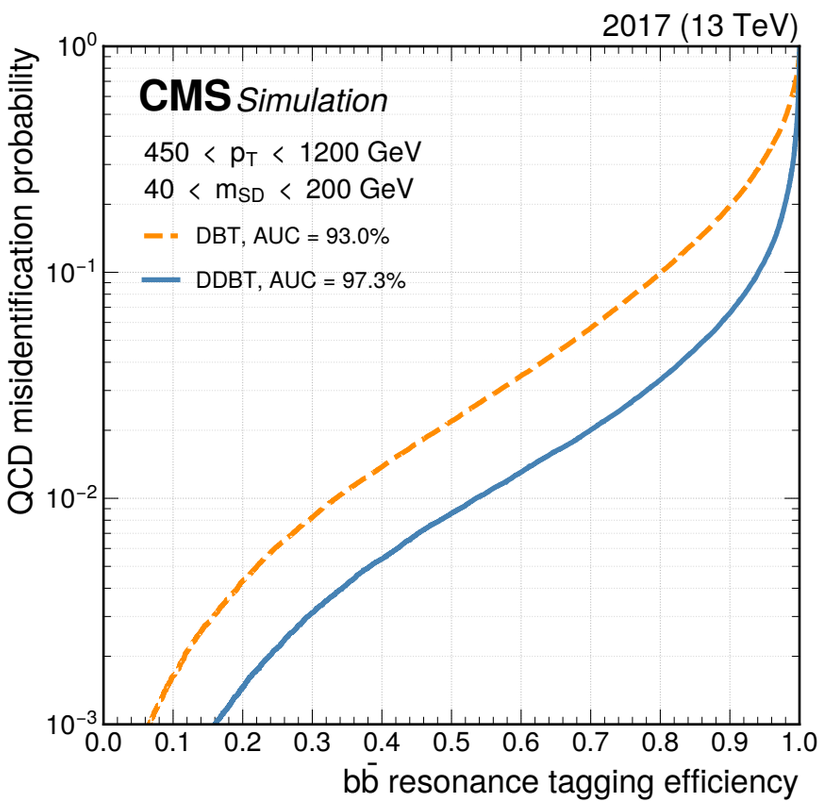


arXiv:2006.13251

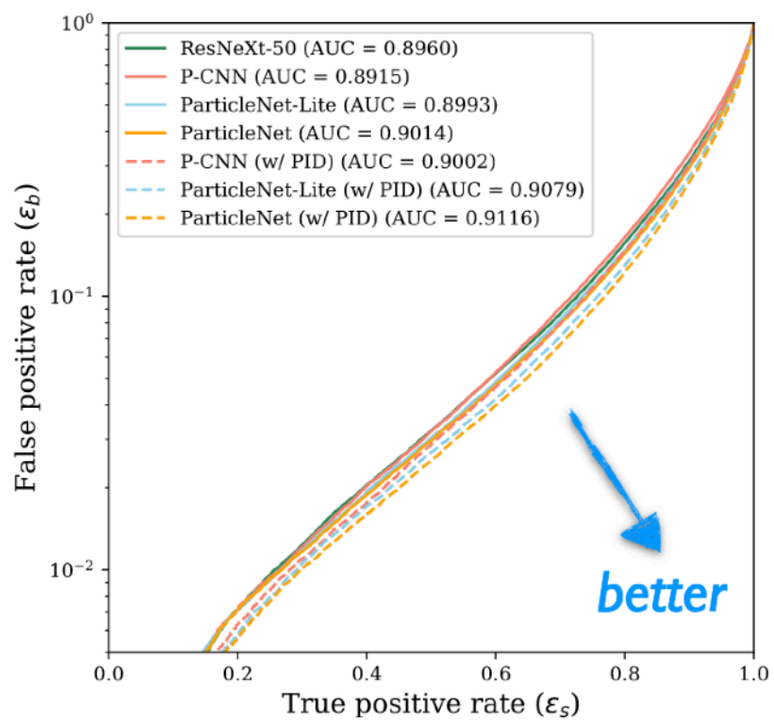
Latest flavor tagging: Nilotpal Kakati's talk

Jet flavor

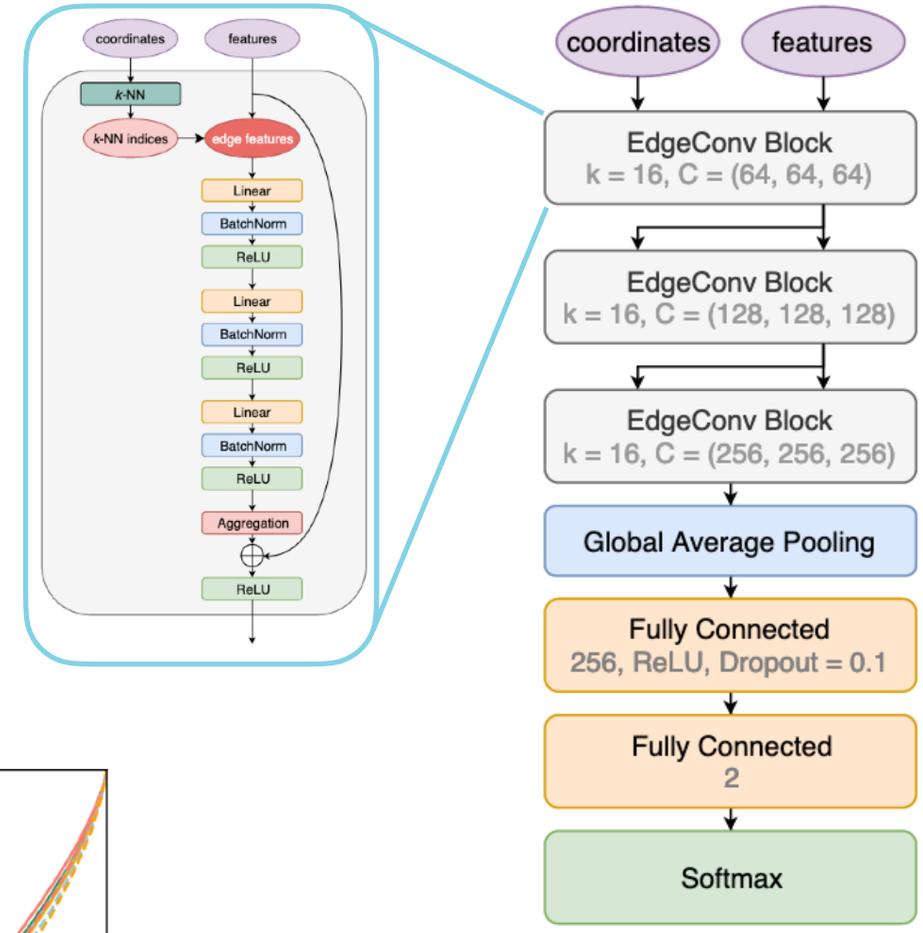
- Big gains with more complex DNN
  - SOTA graph networks



arXiv:2006.13251



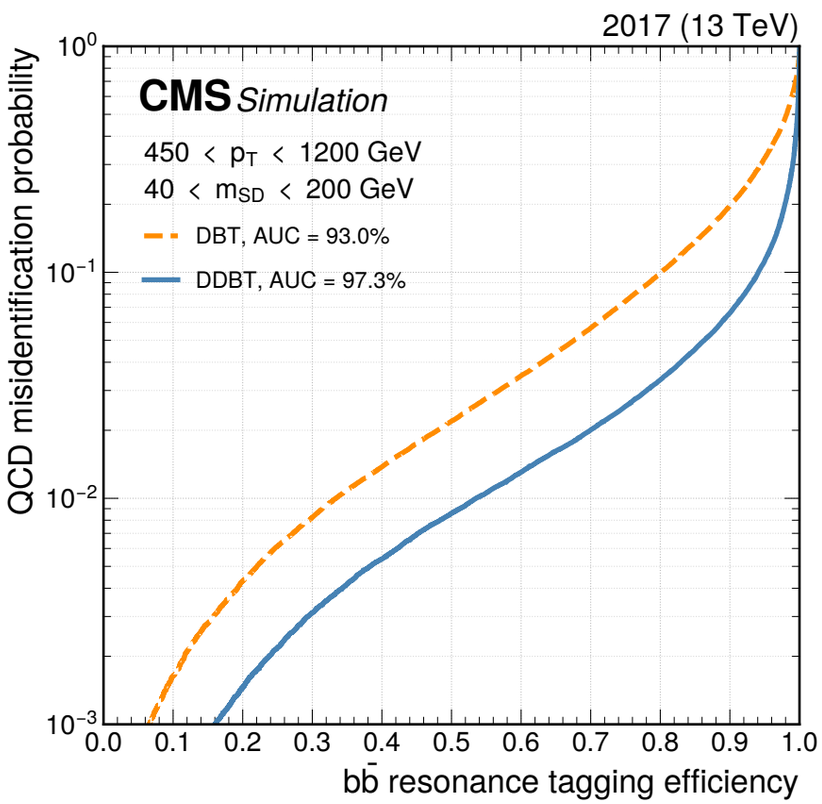
arxiv:1902.08570



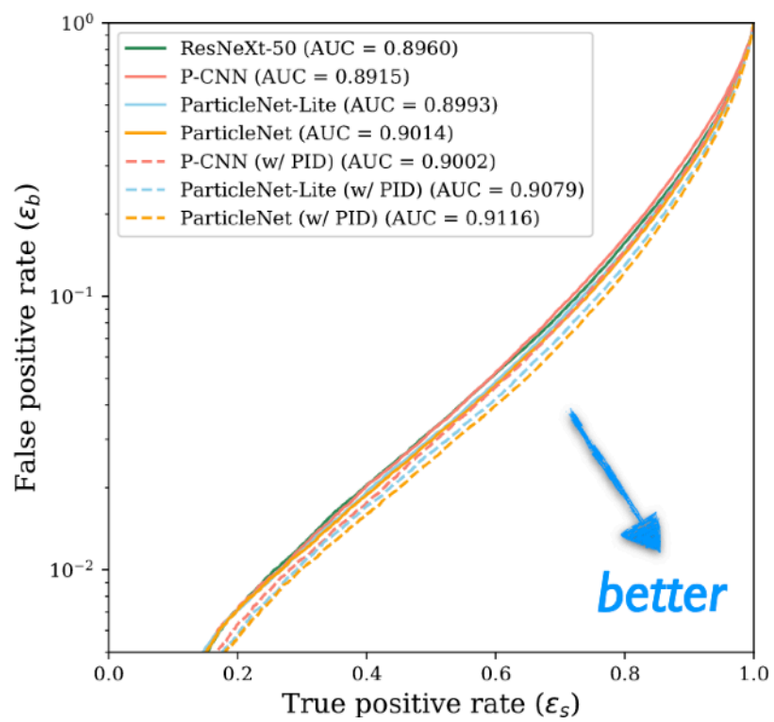
Latest flavor tagging: Nilotpal Kakati's talk

Jet flavor

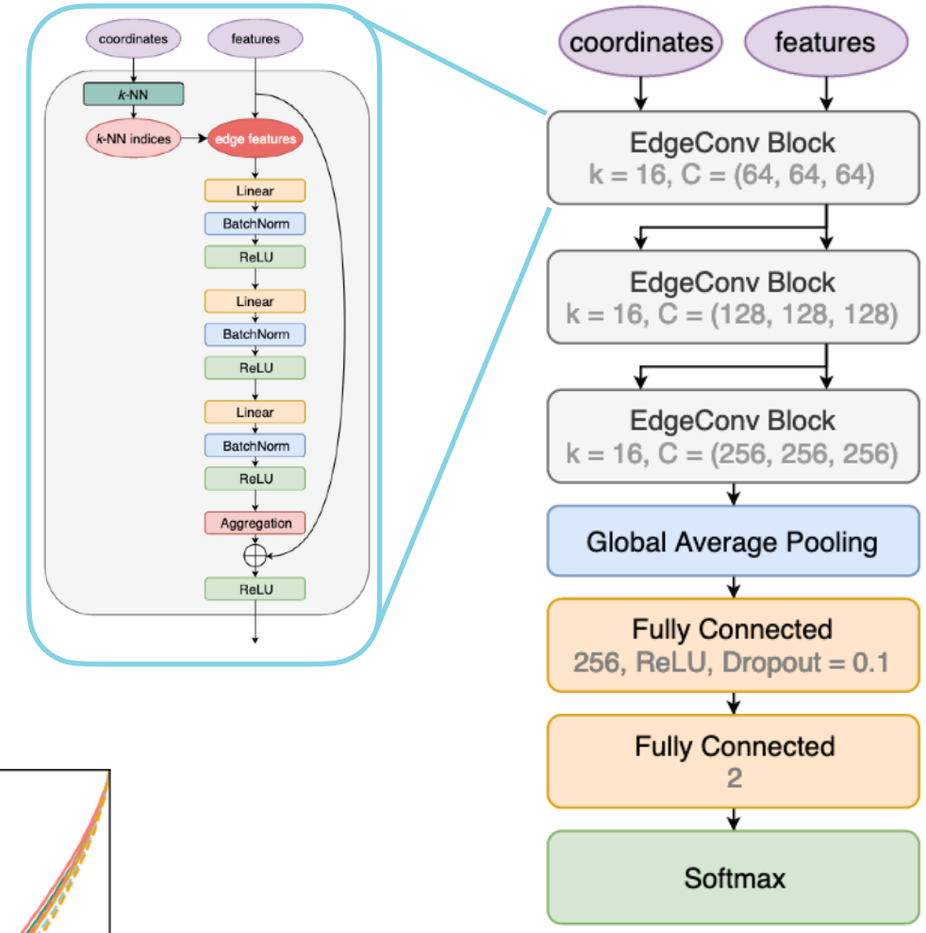
- Big gains with more complex DNN
  - SOTA graph networks

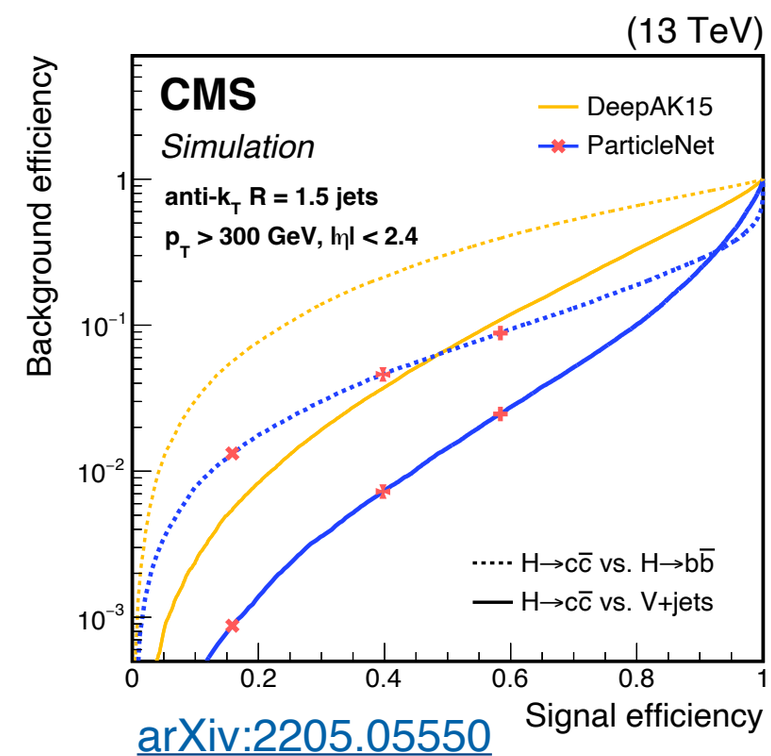


arXiv:2006.13251



arxiv:1902.08570





arXiv:2205.05550

Latest flavor tagging: Nilotpal Kakati's talk

Jet flavor

- Big gains with more complex DNN
  - SOTA ~~graph networks~~ transformers

arxiv:2202.03772

	All classes		$H \to b\bar{b}$	$H \to c\bar{c}$
	Accuracy	AUC	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$
PFN	0.772	0.9714	2924	841
P-CNN	0.809	0.9789	4890	1276
ParticleNet	0.844	0.9849	7634	2475
ParT	**0.861**	**0.9877**	**10638**	**4149**
ParT (plain)	0.849	0.9859	9569	2911

Model complexity

	Accuracy	# params	FLOPs
PFN	0.772	86.1 k	4.62 M
P-CNN	0.809	354 k	15.5 M
ParticleNet	0.844	370 k	540 M
ParT	**0.861**	2.14 M	340 M
ParT (plain)	0.849	2.13 M	260 M

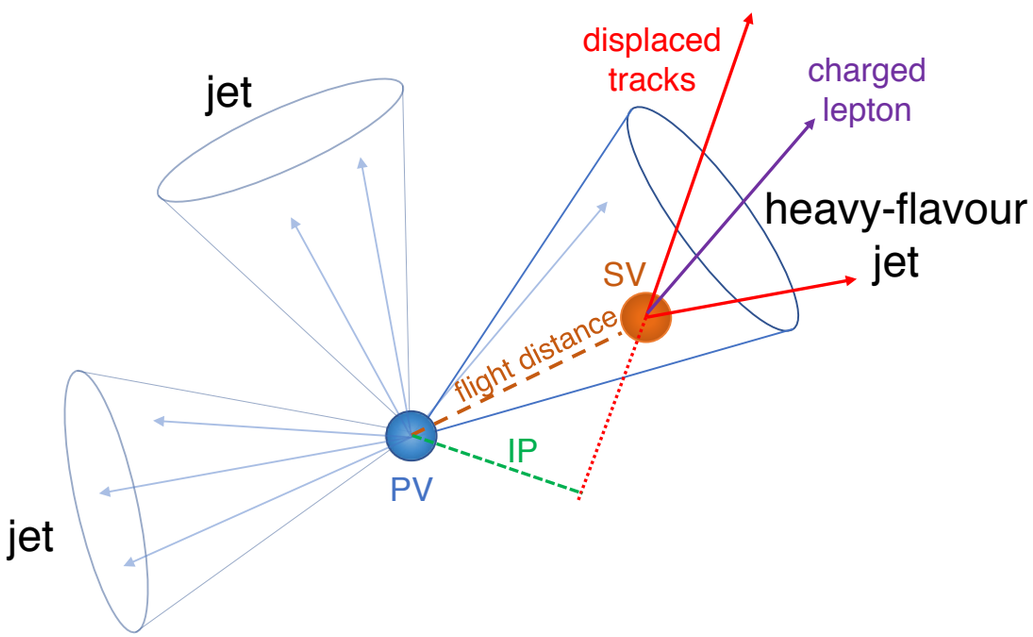
- Computation cost still under control.

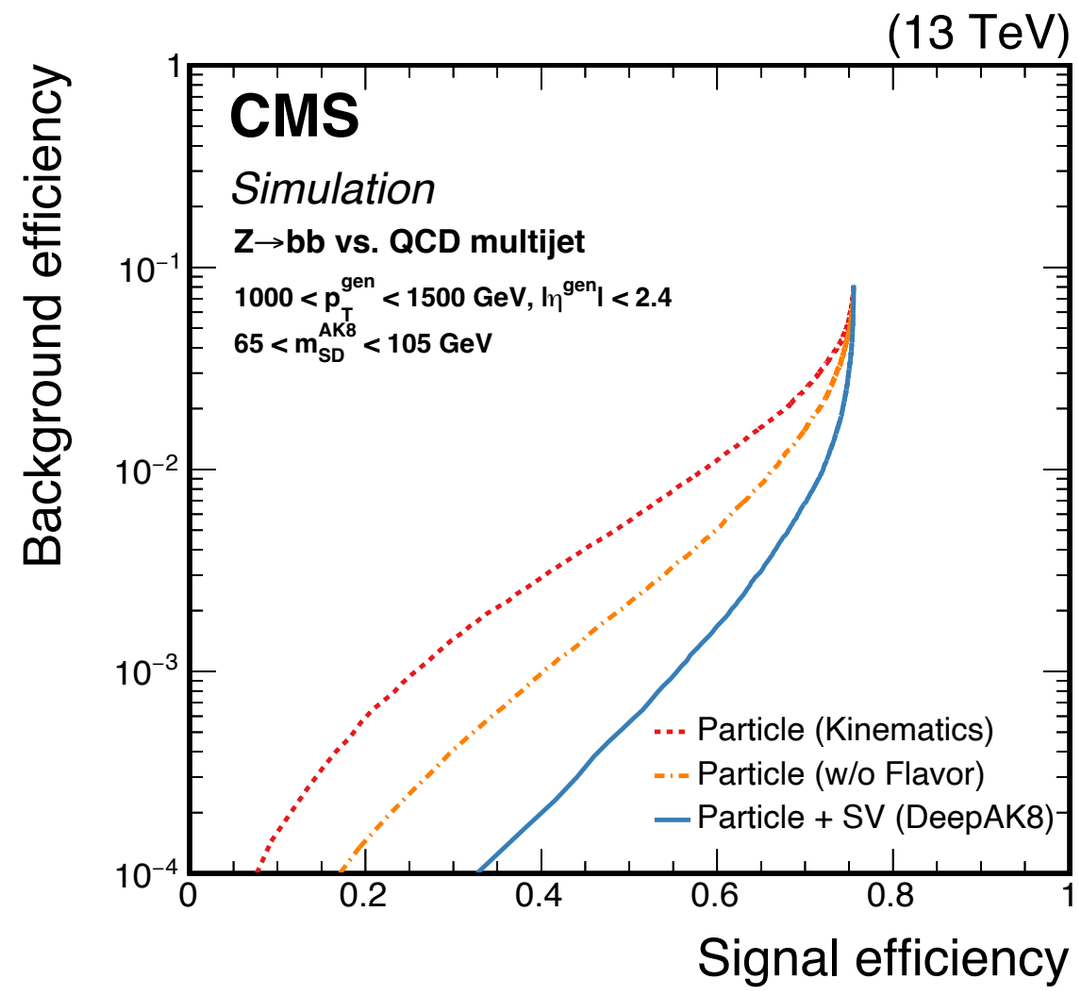
Latest flavor tagging: Nilotpal Kakati's talk

Fermilab

Jet flavor

- Per-particle information beyond kinematics crucial for performance
  - Particle-flow ID (μ / e / charged had. / γ / neutral had.), charge
  - Secondary vertices

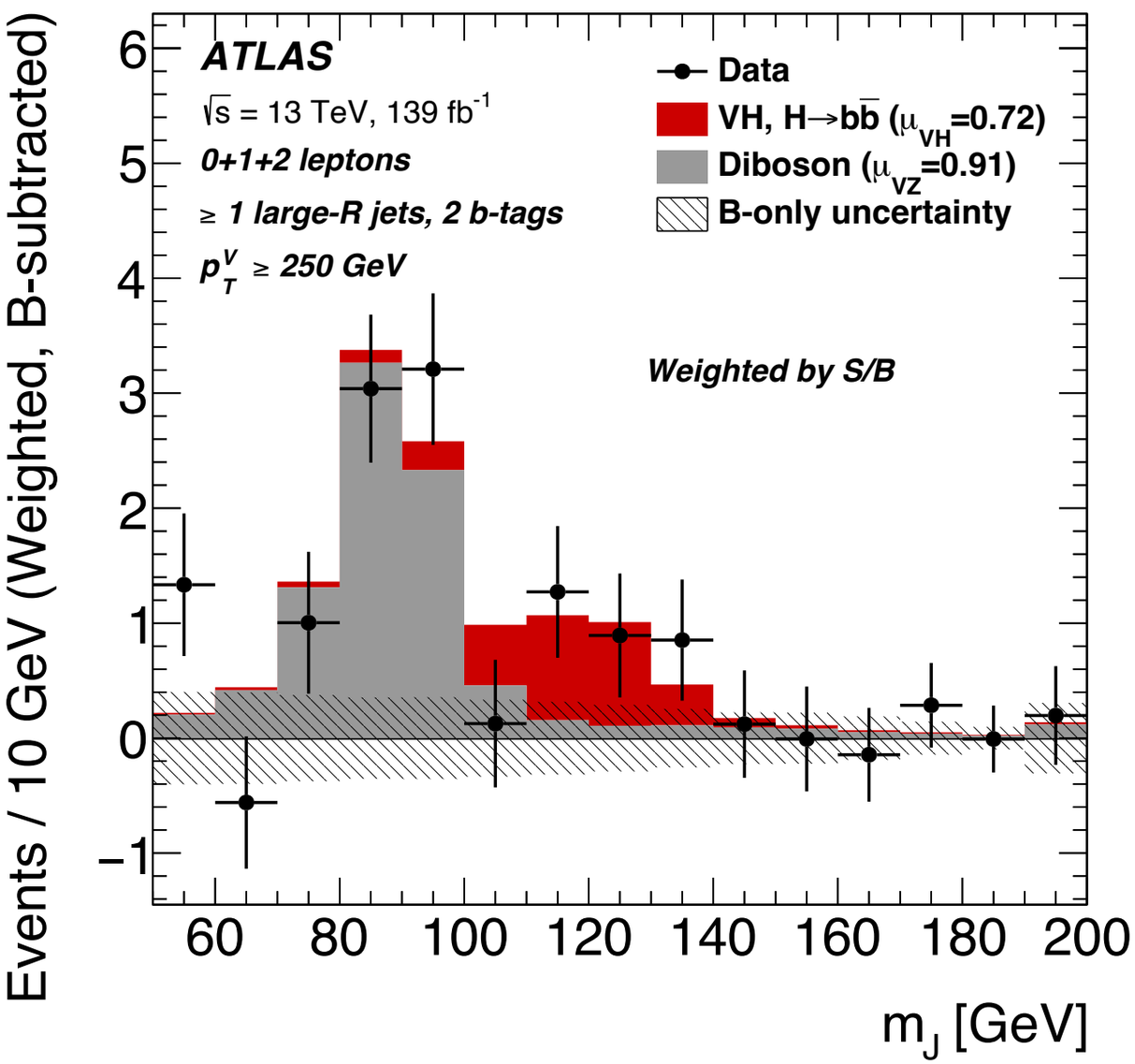


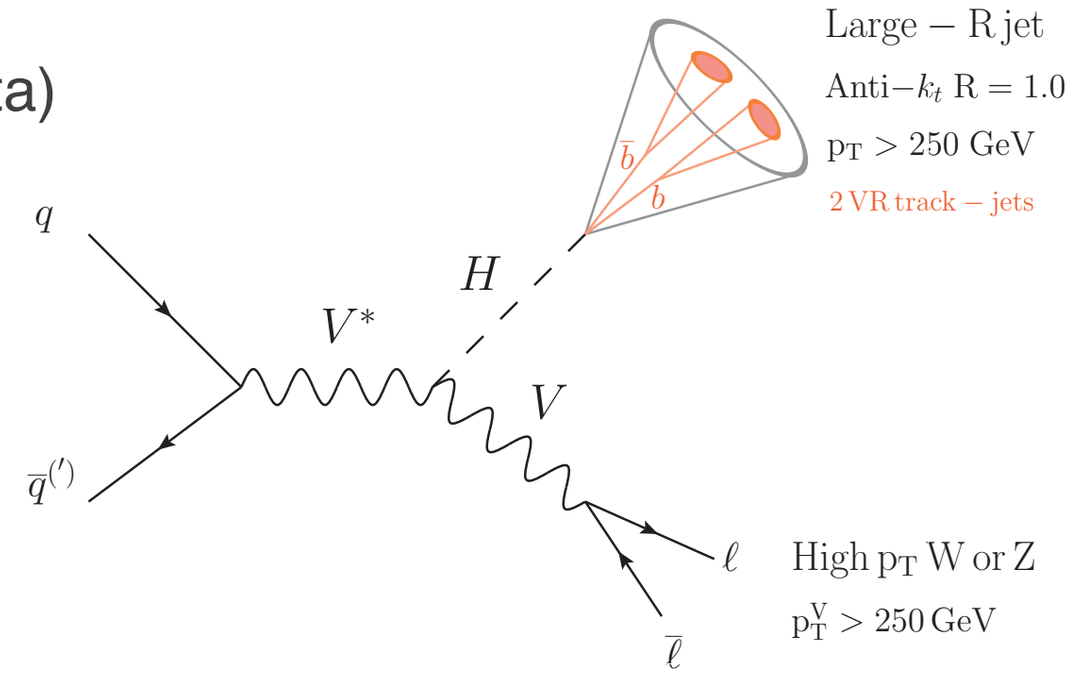


arxiv:2004.08262

Enabling boosted VH(bb)

- ATLAS boosted VH(bb) result
  - Following the 2008 proposal in spirit (with 5x data)
  - 2.1 (2.7 exp.) σ excess





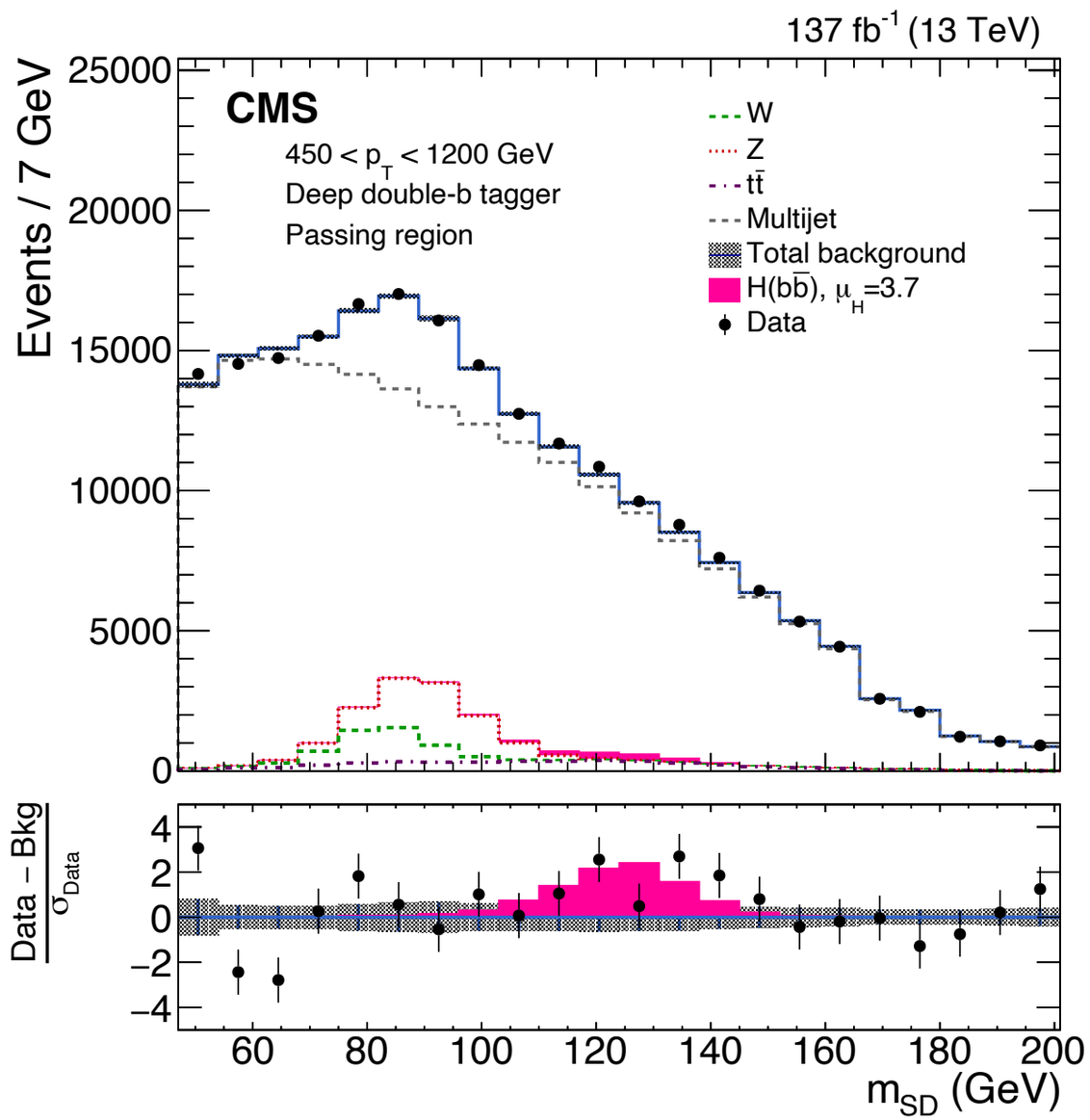
$$\mu_{VH}^{bb} = 0.72^{+0.39}_{-0.36} = 0.72^{+0.29}_{-0.28}(\text{stat.})^{+0.26}_{-0.22}(\text{syst.})$$

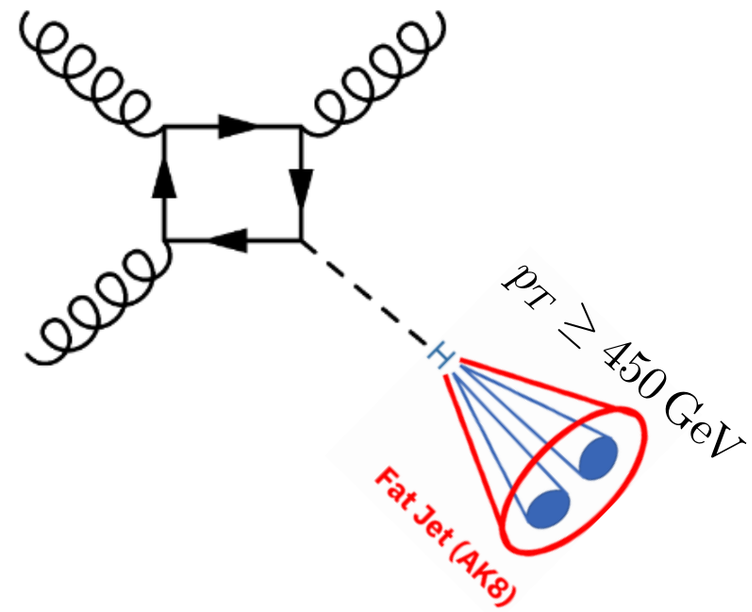
$$\mu_{VZ}^{bb} = 0.91^{+0.29}_{-0.23} = 0.91 \pm 0.15(\text{stat.})^{+0.24}_{-0.17}(\text{syst.})$$

arxiv:2008.02508

Enabling boosted ggH(bb)

- CMS boosted ggH result
  - Anti-k_T, R=0.8, PUPPI softdrop mass
  - 2.5 (0.7 exp.) σ excess over background





$$\mu_H = 3.7 \pm 1.2\,(\text{stat})^{+0.8}_{-0.7}\,(\text{syst})^{+0.8}_{-0.5}\,(\text{theo})$$

$$\mu_Z = 1.01 \pm 0.05\,(\text{stat})^{+0.20}_{-0.15}\,(\text{syst})^{+0.13}_{-0.09}\,(\text{theo})$$

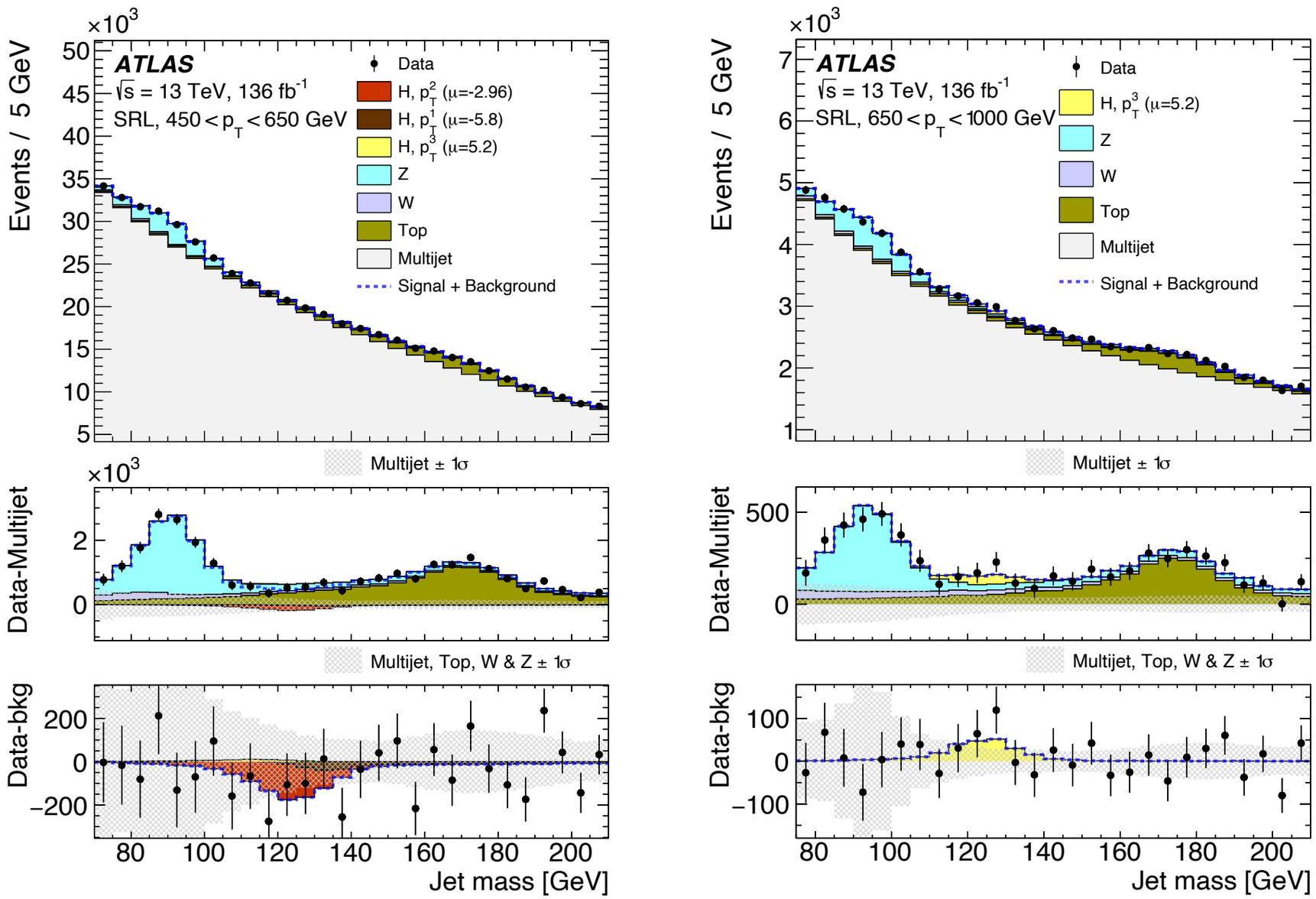
arxiv:2006.13251

High-p_T Z(bb): tagging uncertainty larger than theory

Enabling boosted ggH(bb)

- ATLAS boosted ggH result
  - Anti-k_T, R=1, trimmed mass
  - μ_H = 0.8±3.2

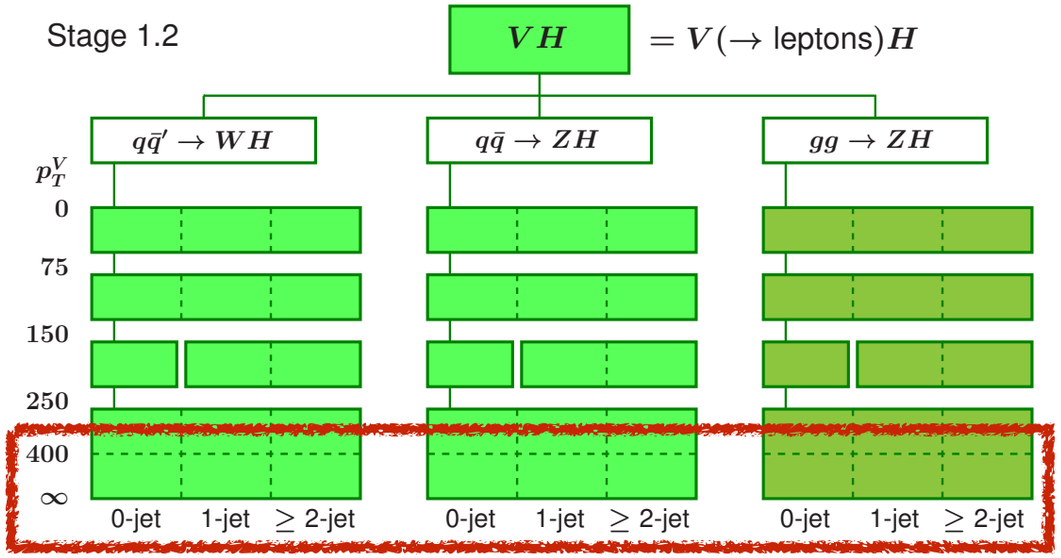
arxiv:2111.08340

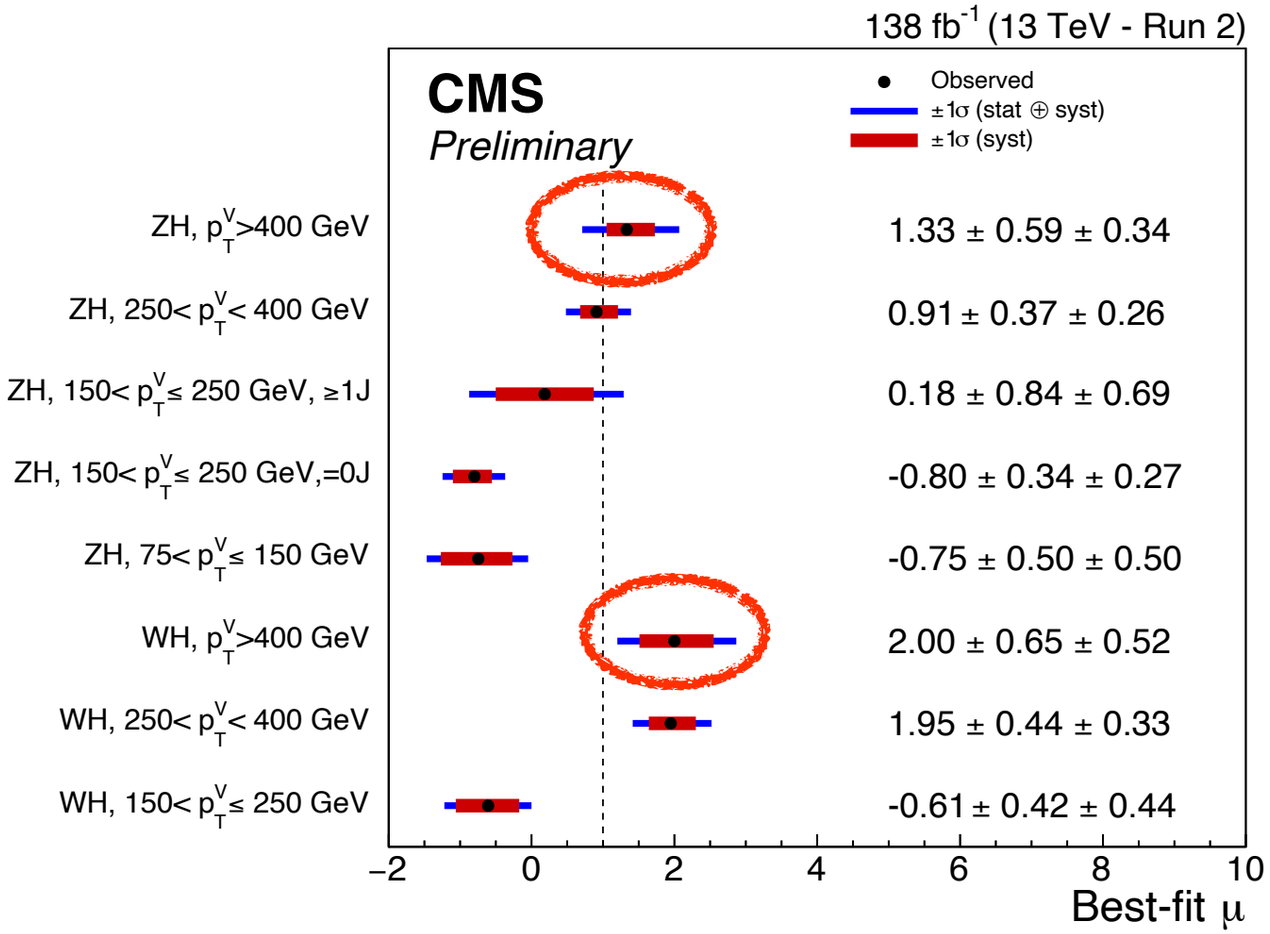


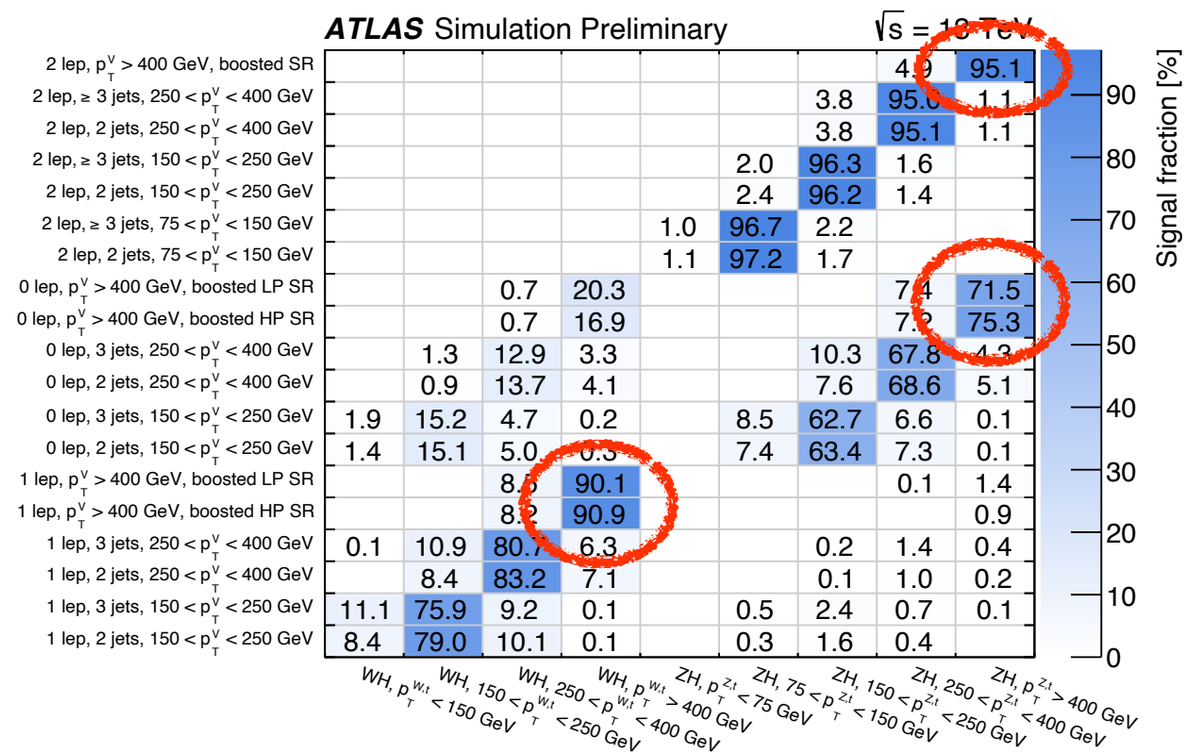
Further details on CMS VH(bb): Nicholas Haubrich's talk

Improving STXS tails

- ATLAS, CMS: full Run-2 VH(bb) STXS
  - Both include boosted categories
    - Merged jet: CMS: AK8, ATLAS: AK10
  - Strong constraint at $p_T(V) > 400$





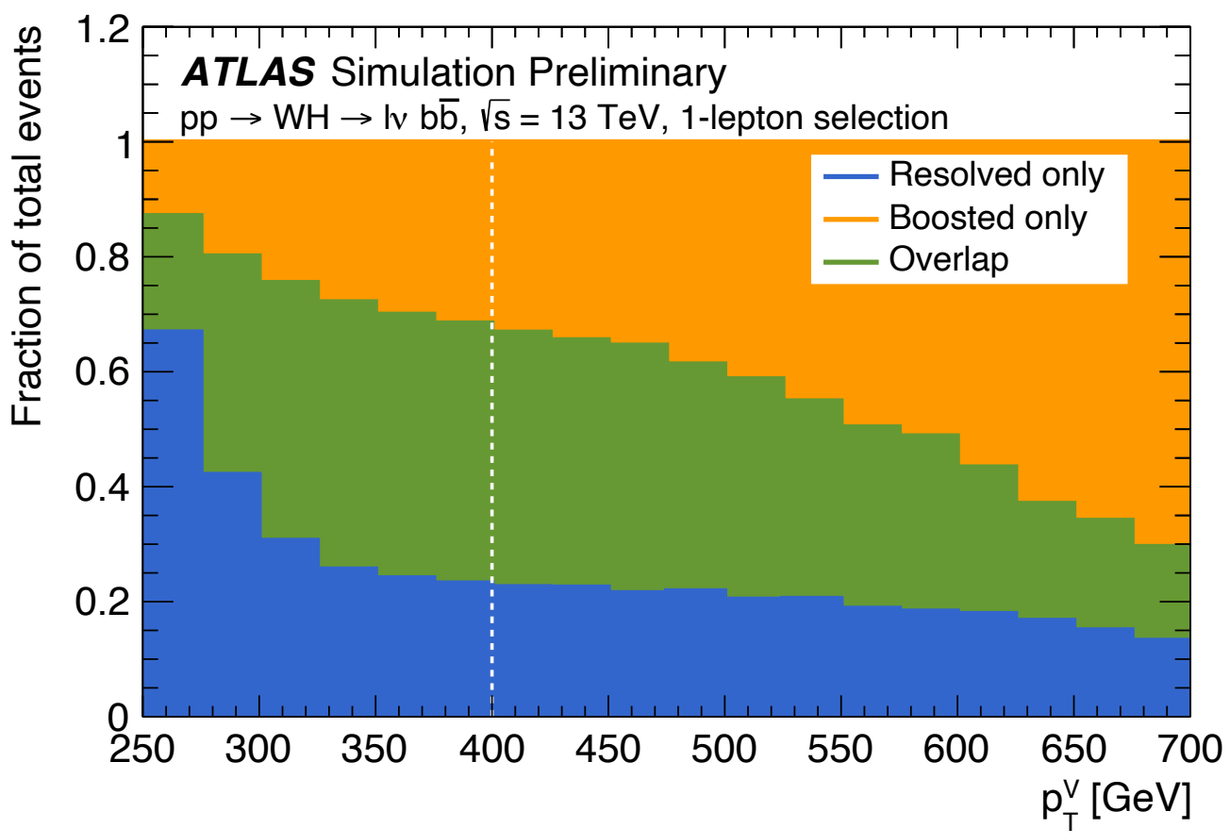


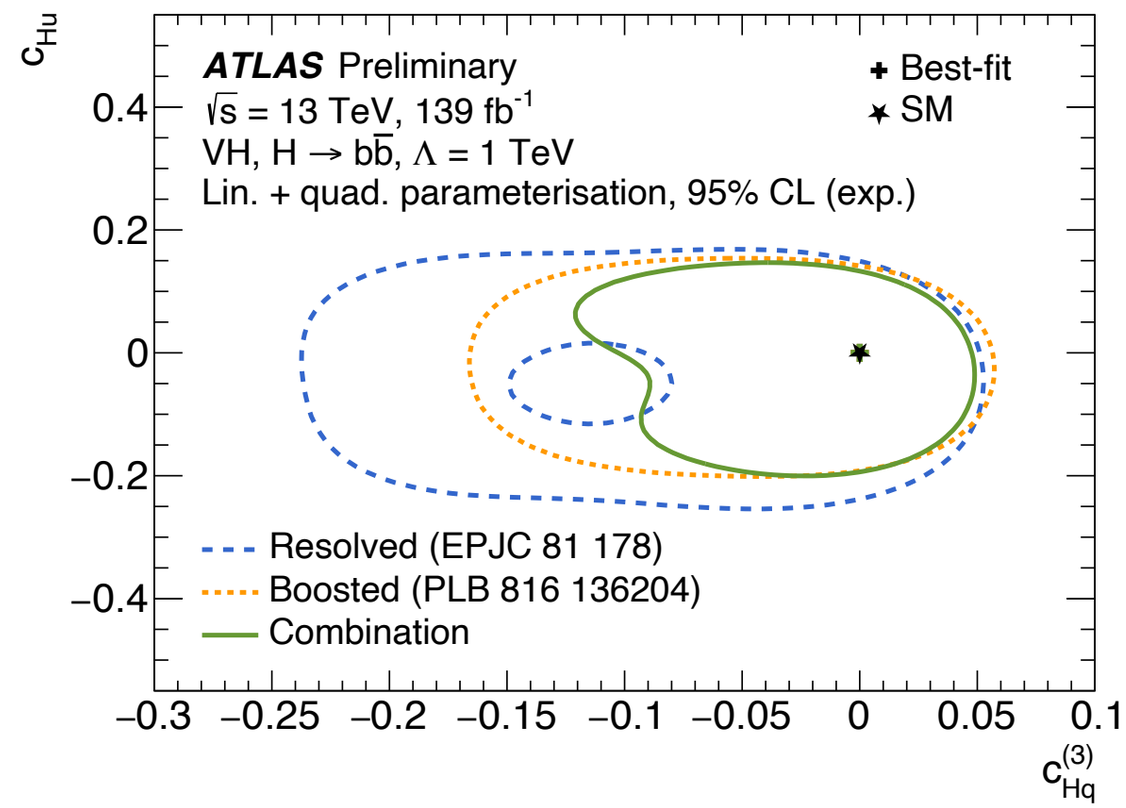
HIG-20-001

ATLAS-CONF-2021-051

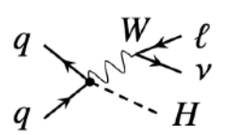
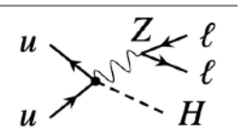
Improving EFT constraints

- ATLAS result: SMEFT interpretation
  - Boosted/resolved crossover ~ 400 GeV
- Combination resolves degeneracies
  - Complementary kinematic regimes





ATLAS-CONF-2021-051

$c_{Hq}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}^I_\mu H)(\bar{q}_p \tau^I \gamma^\mu q_r)$	
c_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u}_p \gamma^\mu u_r)$	

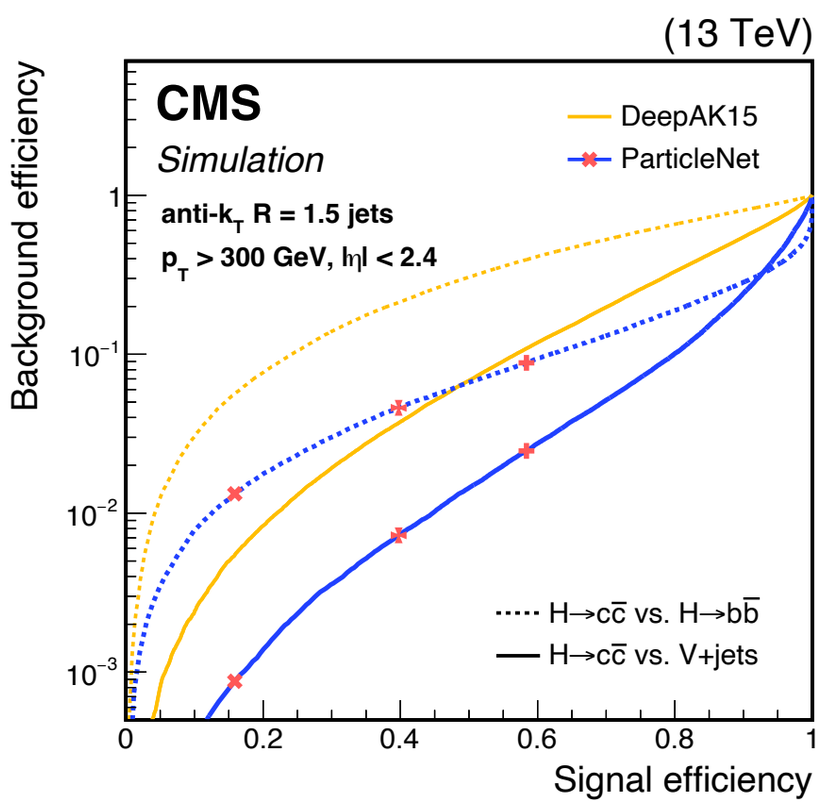
Representative diagrams courtesy ATLAS-CONF-2020-053

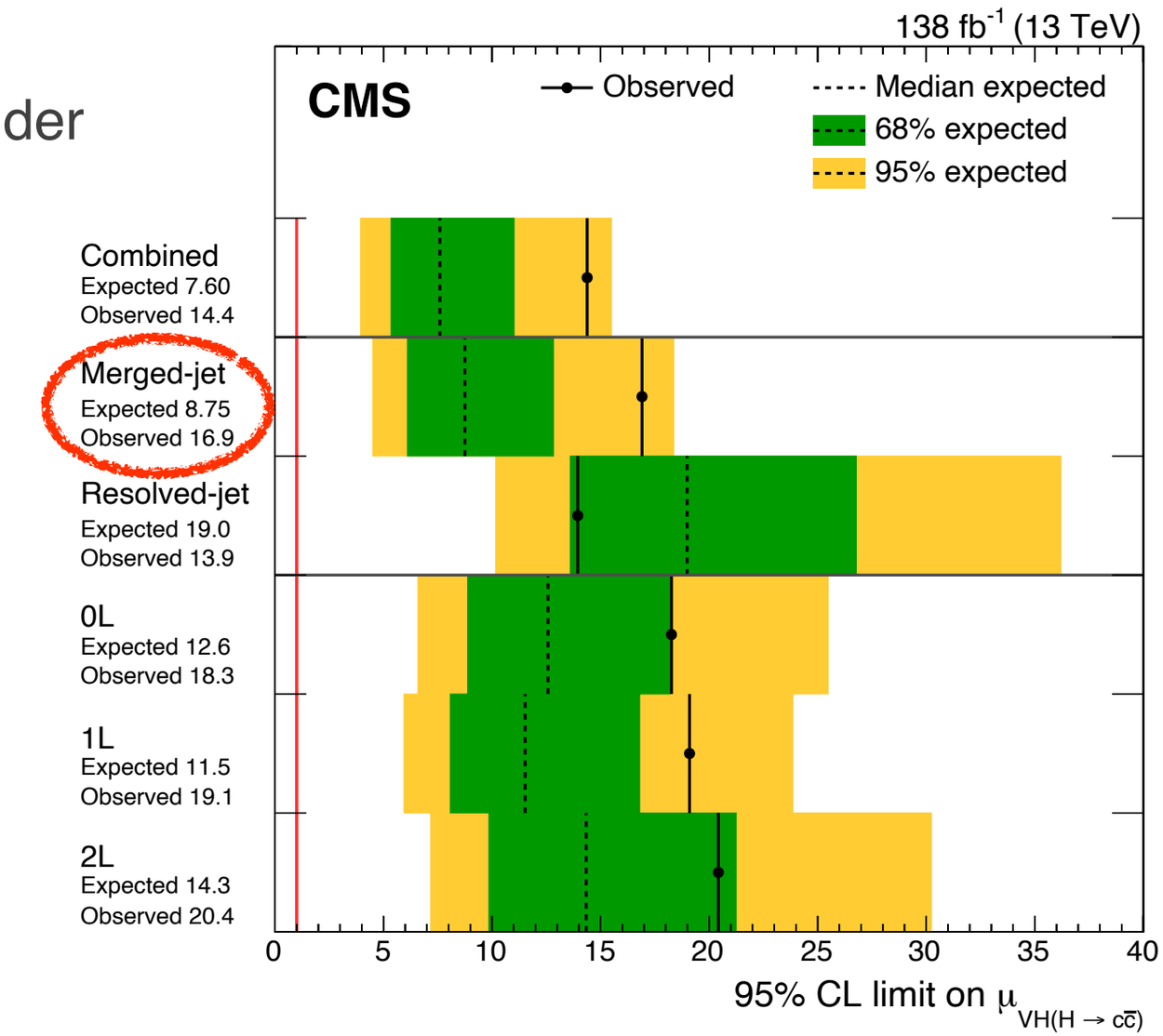
Further details on this result: Marko Stamenkovic's talk

Enabling boosted VH(cc)

- CMS full Run II VH(cc)
  - Combines resolved and merged-jet channels
    - Merged: one AK15 ParticleNet cc-tagged jet
- Validated using VZ(cc)
  - Excess significance 5.7 (5.9) σ
  - First observation of Z(cc) at a hadron collider
- VH(cc) upper limit: 14x SM

arxiv:2205.05550



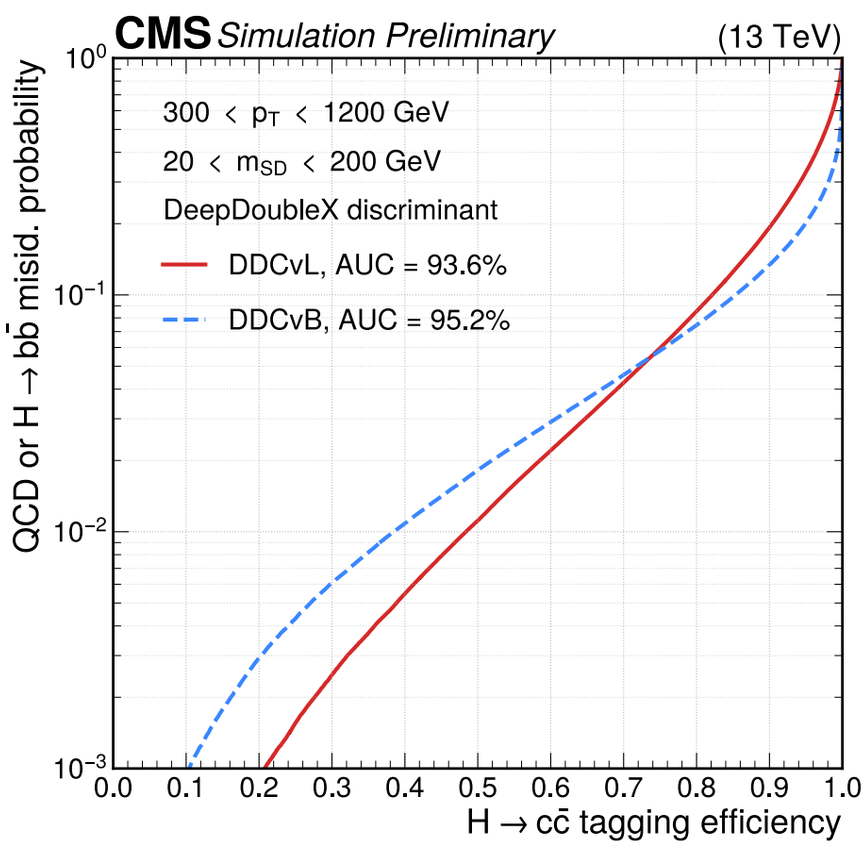


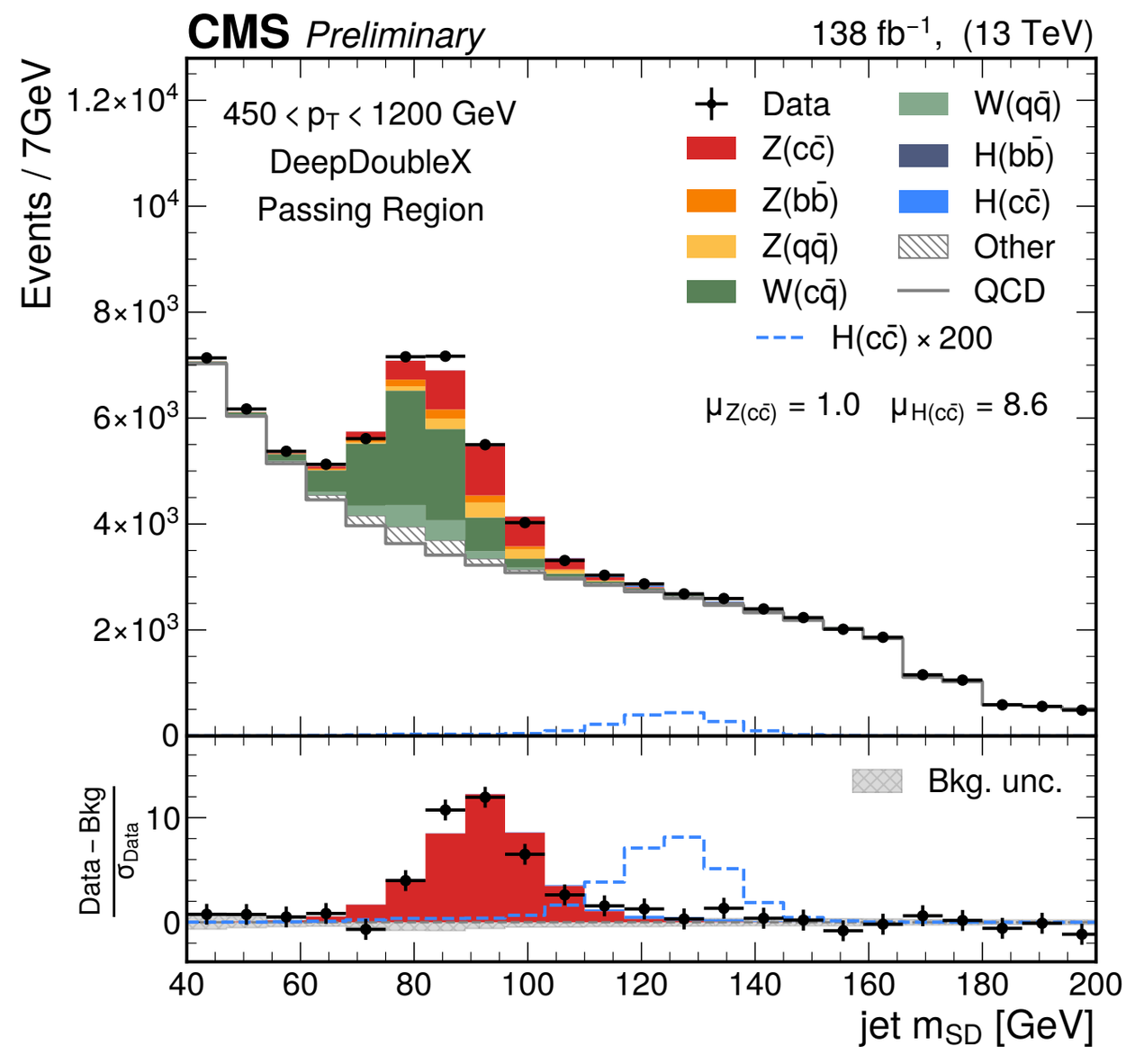
Further details on this result: Marko Stamenkovic's talk

Enabling boosted ggH(cc)

- CMS full Run II ggH(cc)
  - Similar technique to ggH(bb): one AK8 DeepDoubleX cc-tagged jet
  - QCD estimation via CR + sideband constraint ("differential alphabet")
- Validated using Z(cc)
  - Observation of Z(cc)+jets
- ggH(cc) upper limit: 45x SM

HIG-21-012



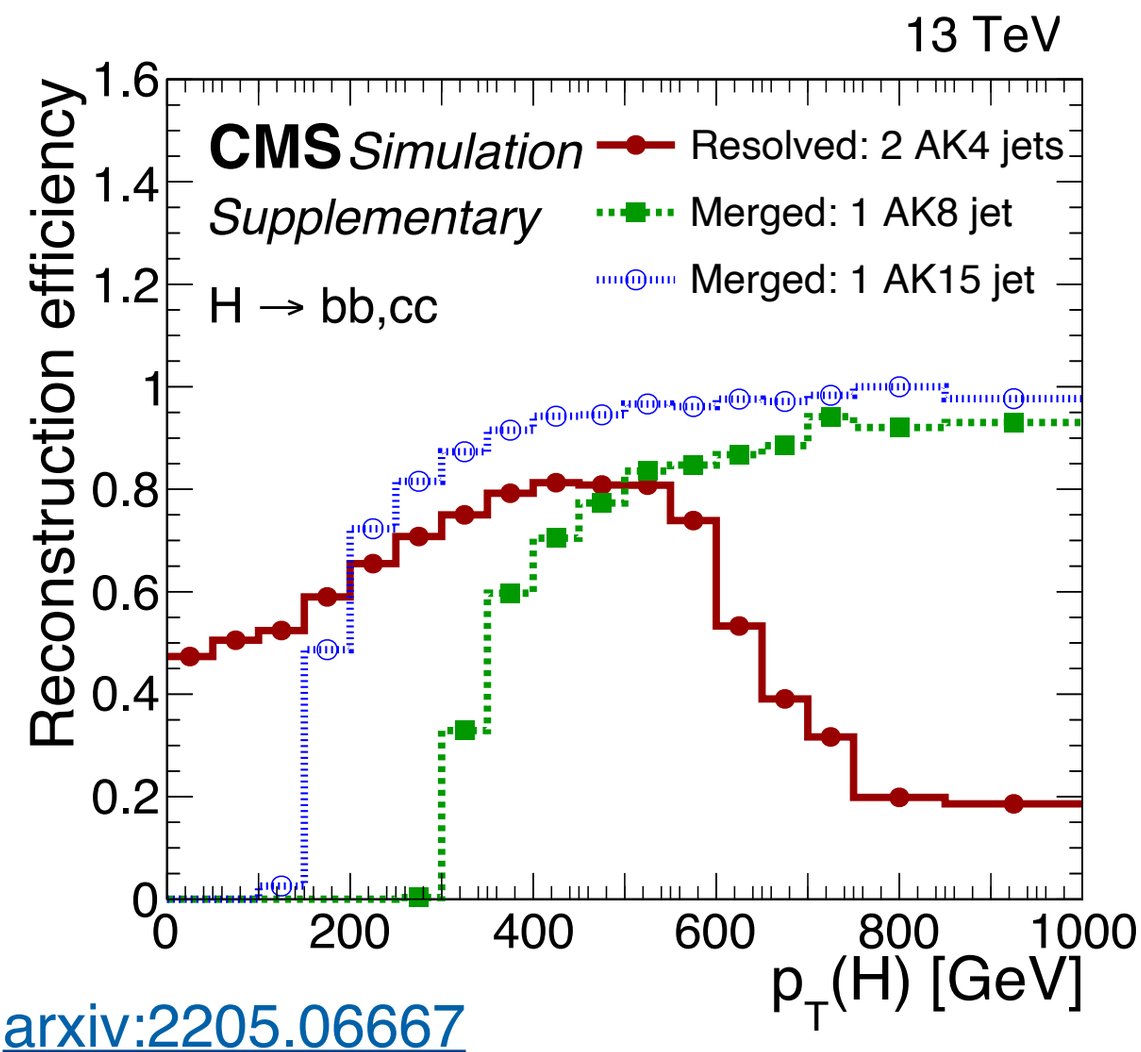


Towards HH

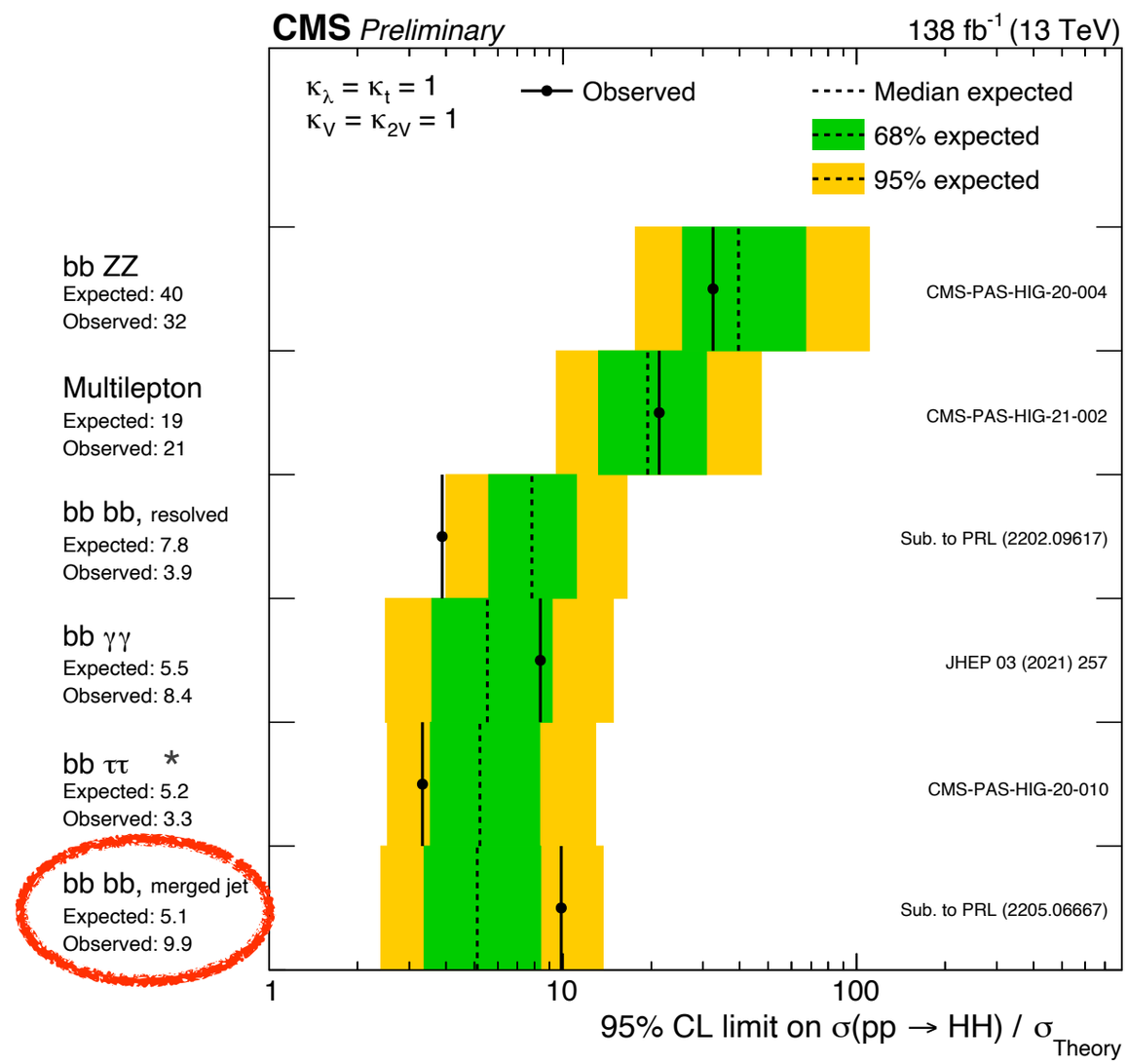
- CMS: merged jet channel for HH(4b)
  - R=1.5 jet improves low-p_T acceptance
  - Leading performance for cross section limit

Further details on CMS combination: Jin Wang's talk

* bb$\tau\tau$ also has a boosted region, details in Jaime Holgado's talk



arxiv:2205.06667

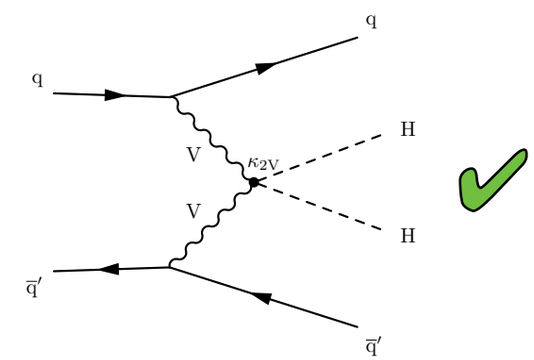


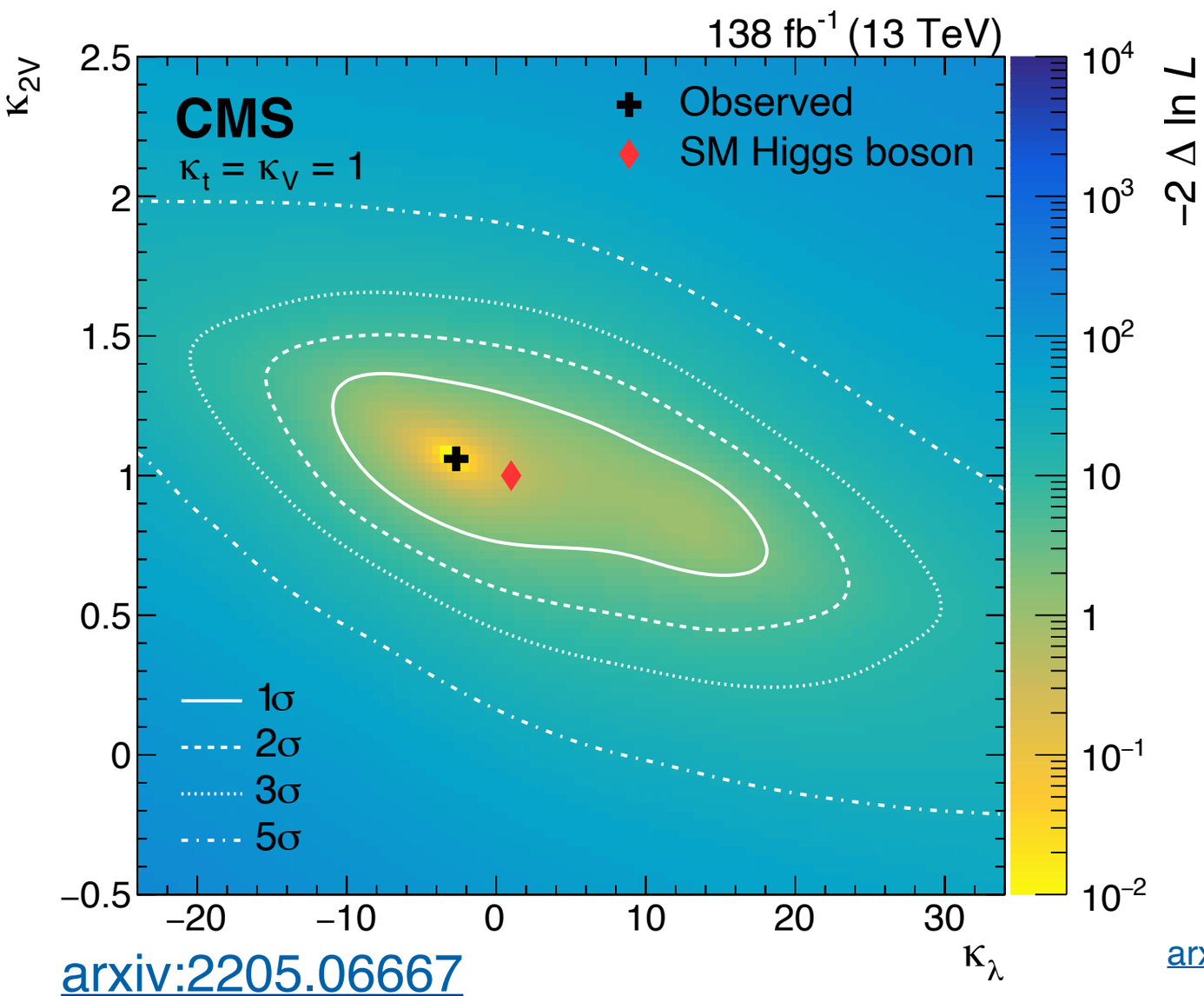
arxiv:2207.00043, CMSPublic twiki: SummaryResultsHIG

Towards HH

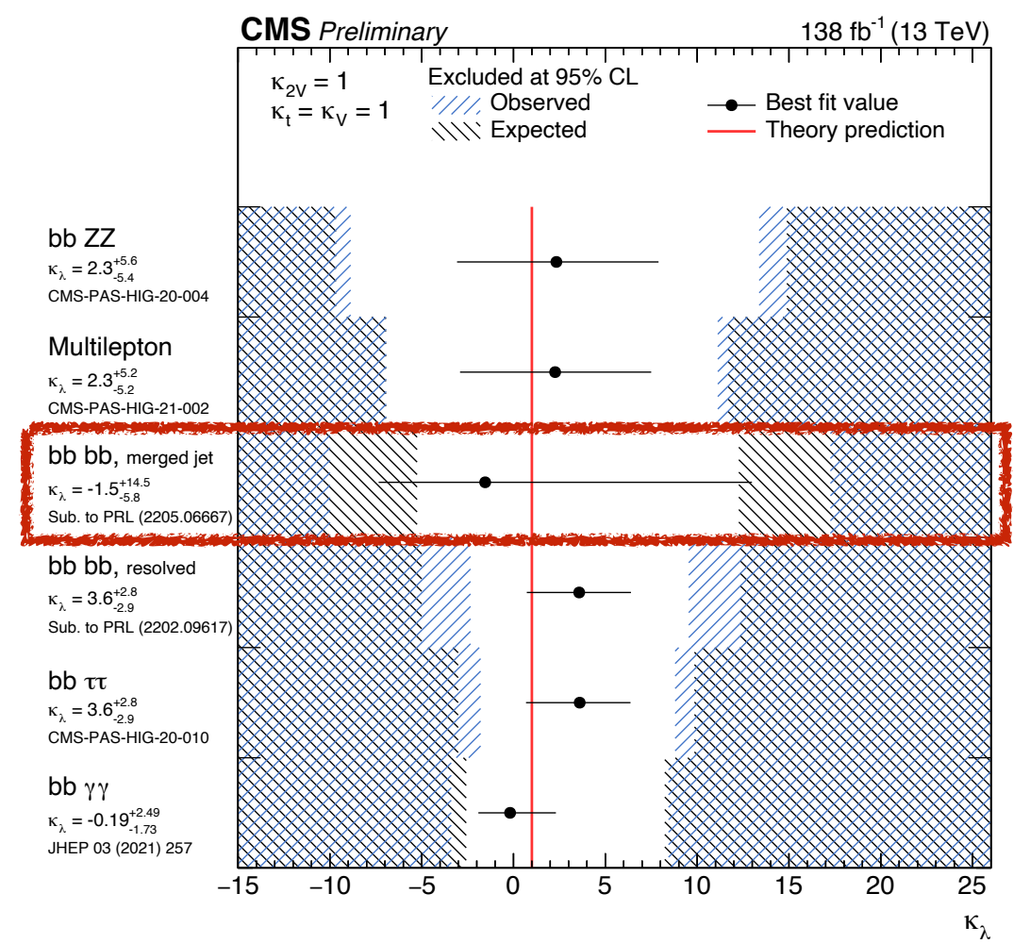
Further details: Nicholas Haubrich's talk

- CMS: merged jet channel for HH(4b)
  - Quartic coupling $\kappa_{2V} = 0$ excluded at 6.3 σ
  - For κ_λ, slightly less sensitive





arxiv:2205.06667

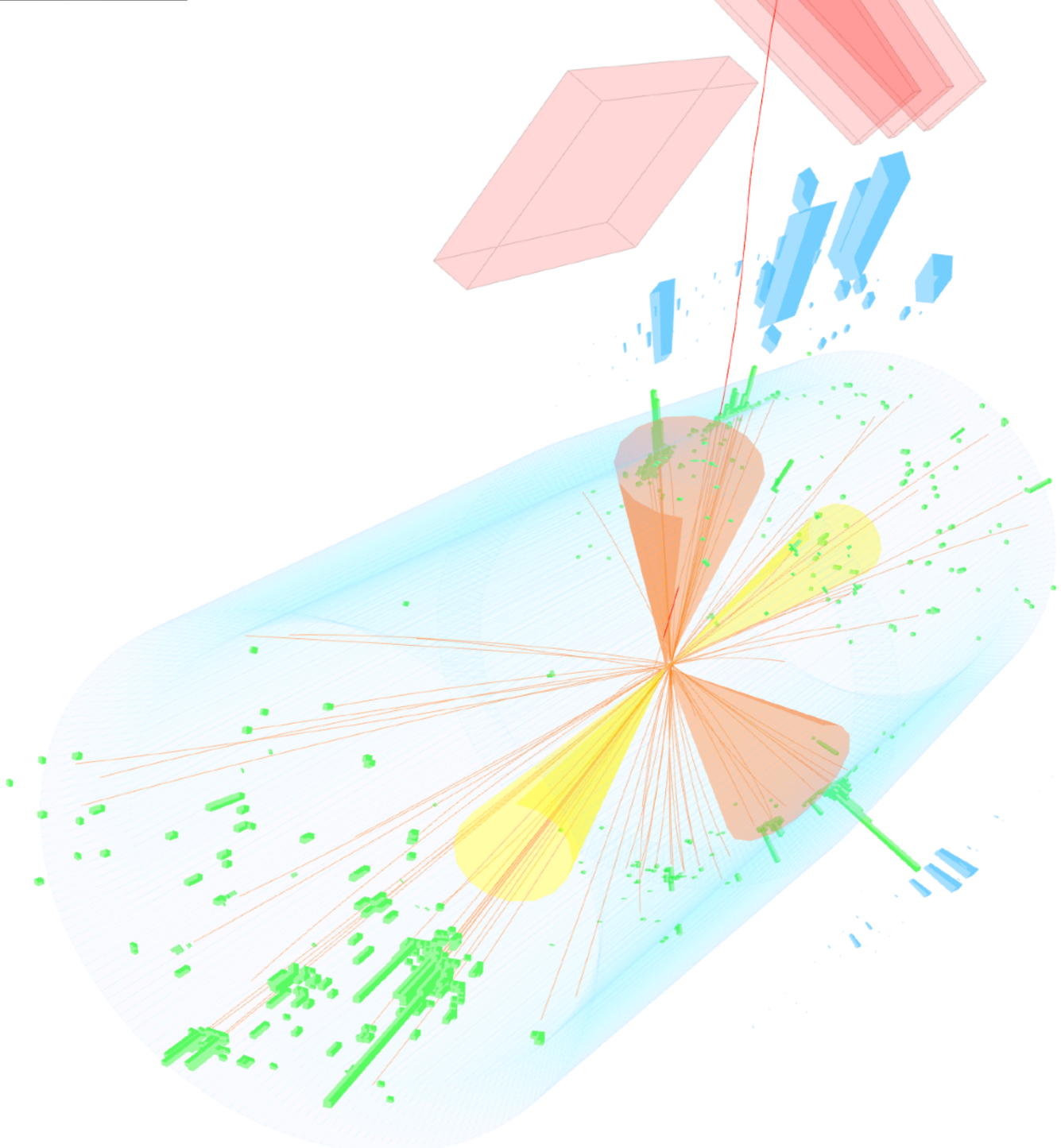


arxiv:2207.00043, CMSPublic twiki: SummaryResultsHIG

Fermilab

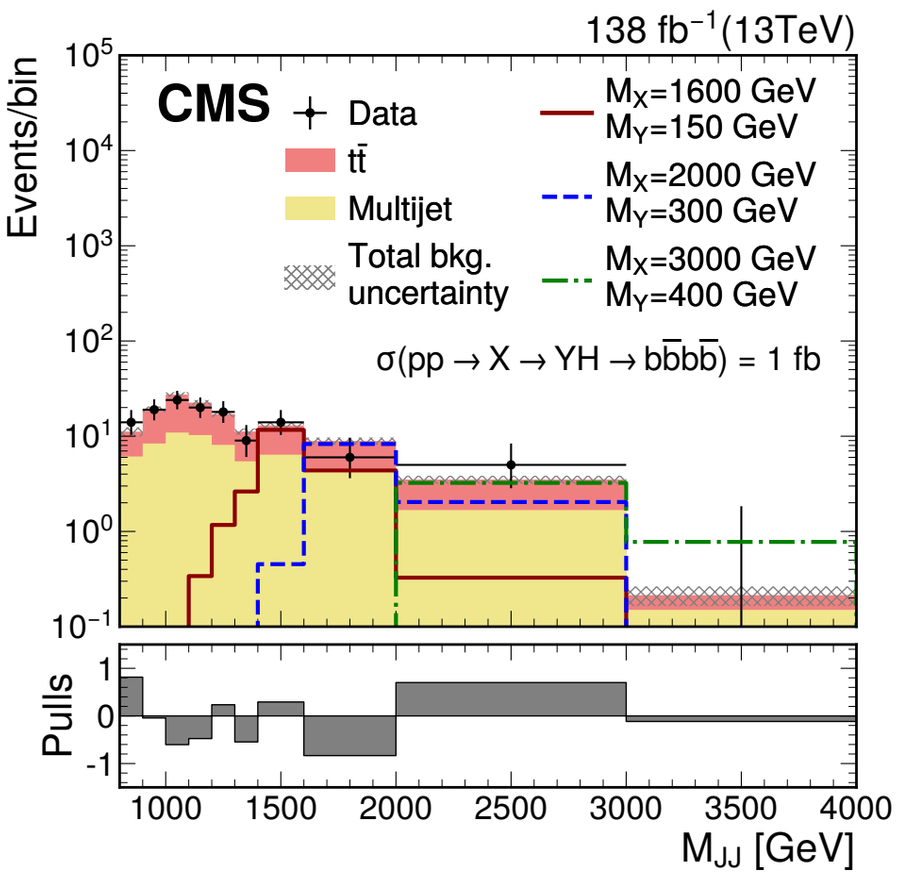
VBF HH

CMS Experiment at the LHC, CERN
Data recorded: 2017-Aug-07 19:13:22.727552 GMT
Run / Event / LS: 300633 / 525384863 / 347

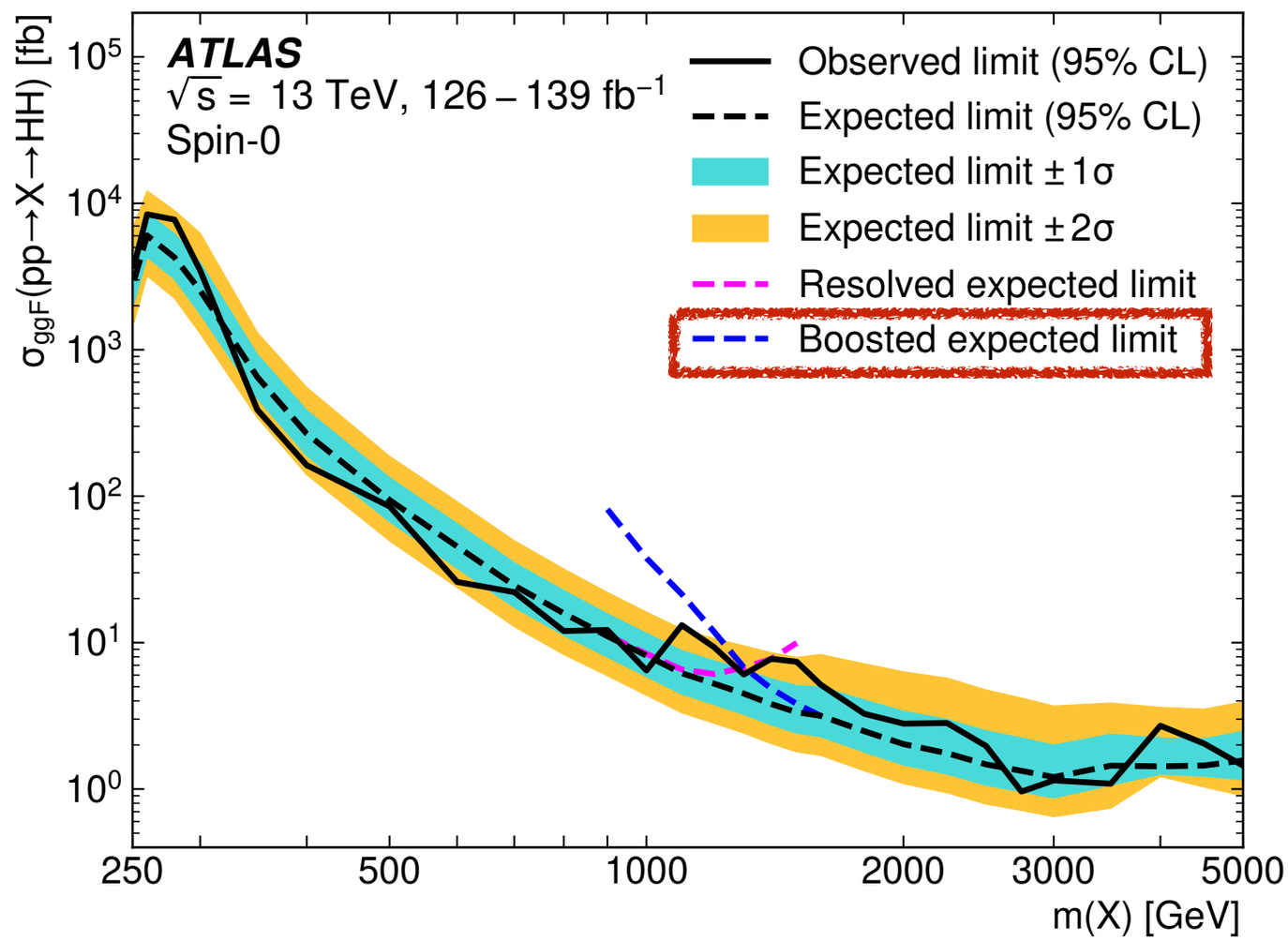


Resonant HH

- Hadronic Higgs at very high boost ➜ HH resonance search
  - And YH resonance in NMSSM model



arxiv:2204.12413

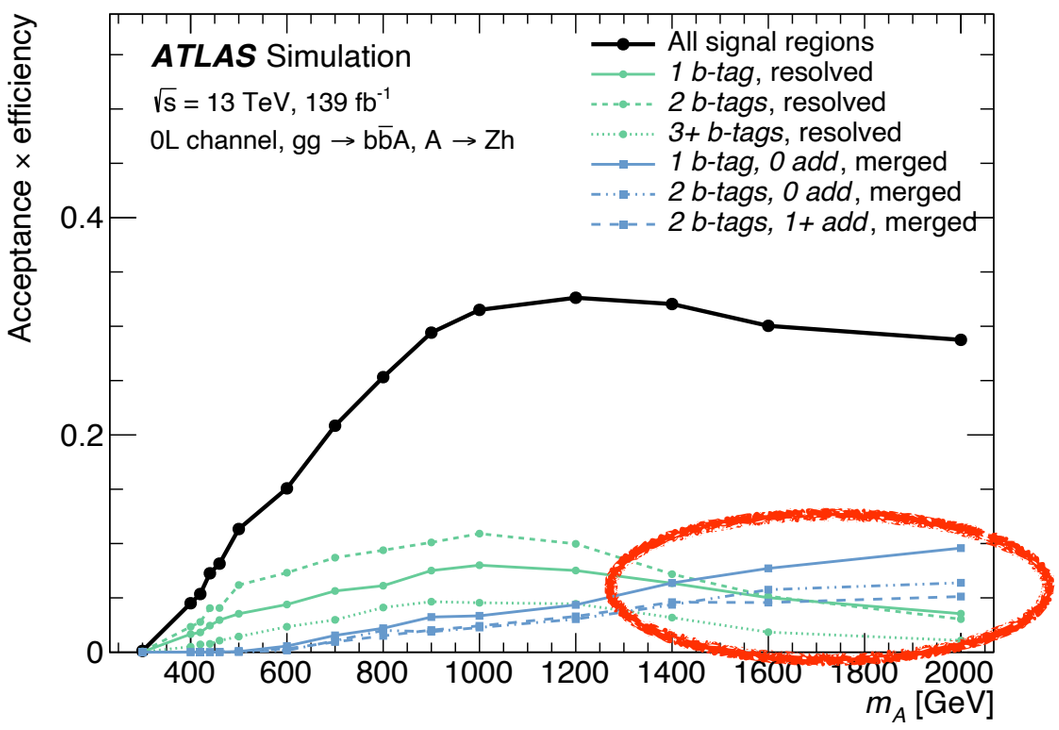


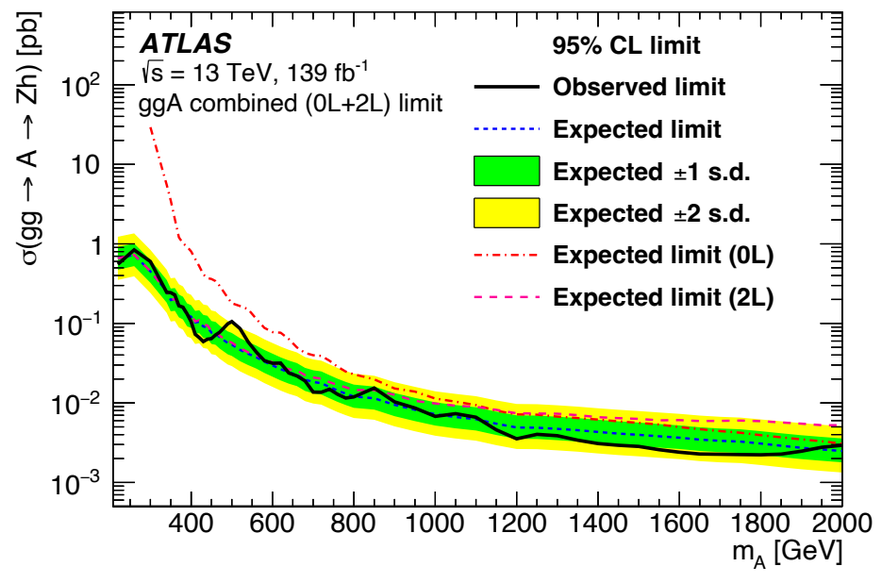
arxiv:2202.07288

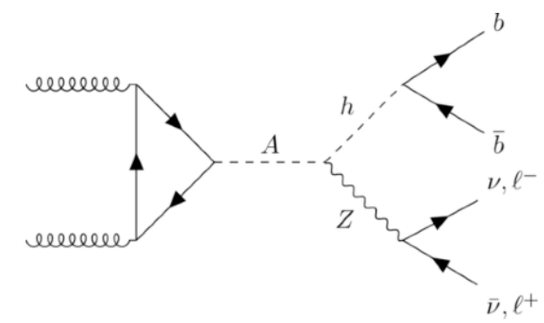
Resonant multiboson

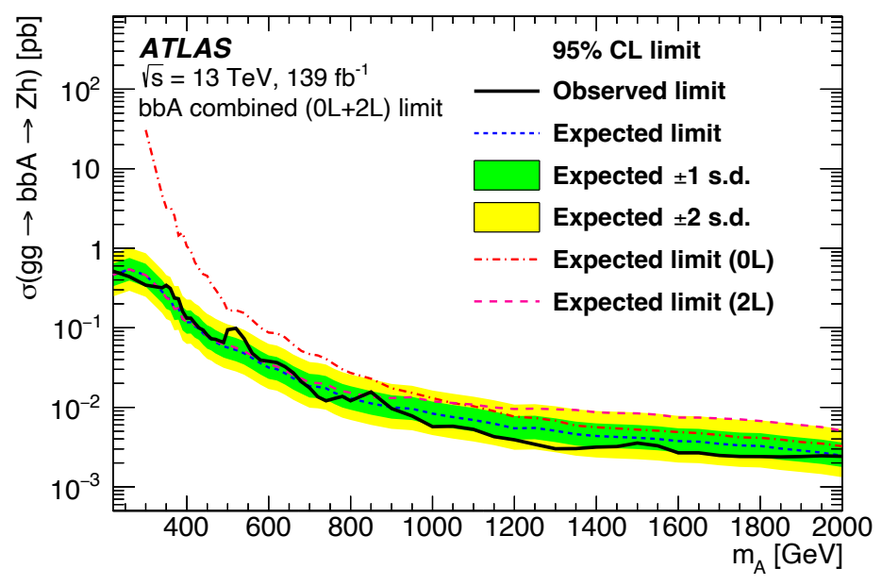
Also relevant: CMS all-hadronic VH resonant: arxiv:2210.00043

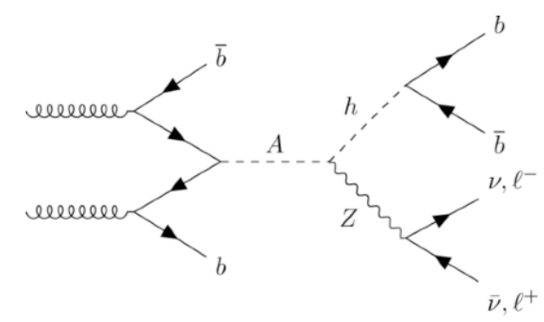
- ATLAS: MSSM A(Zh) search
  - Merged-jet category extends m_A reach







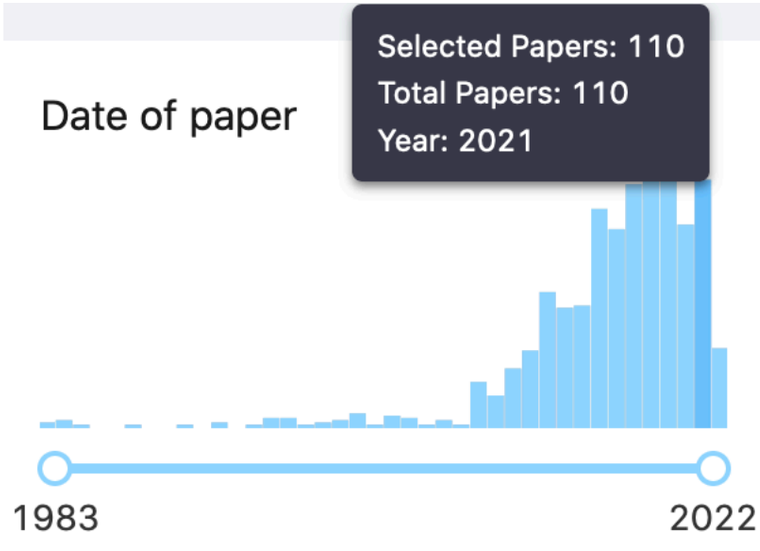




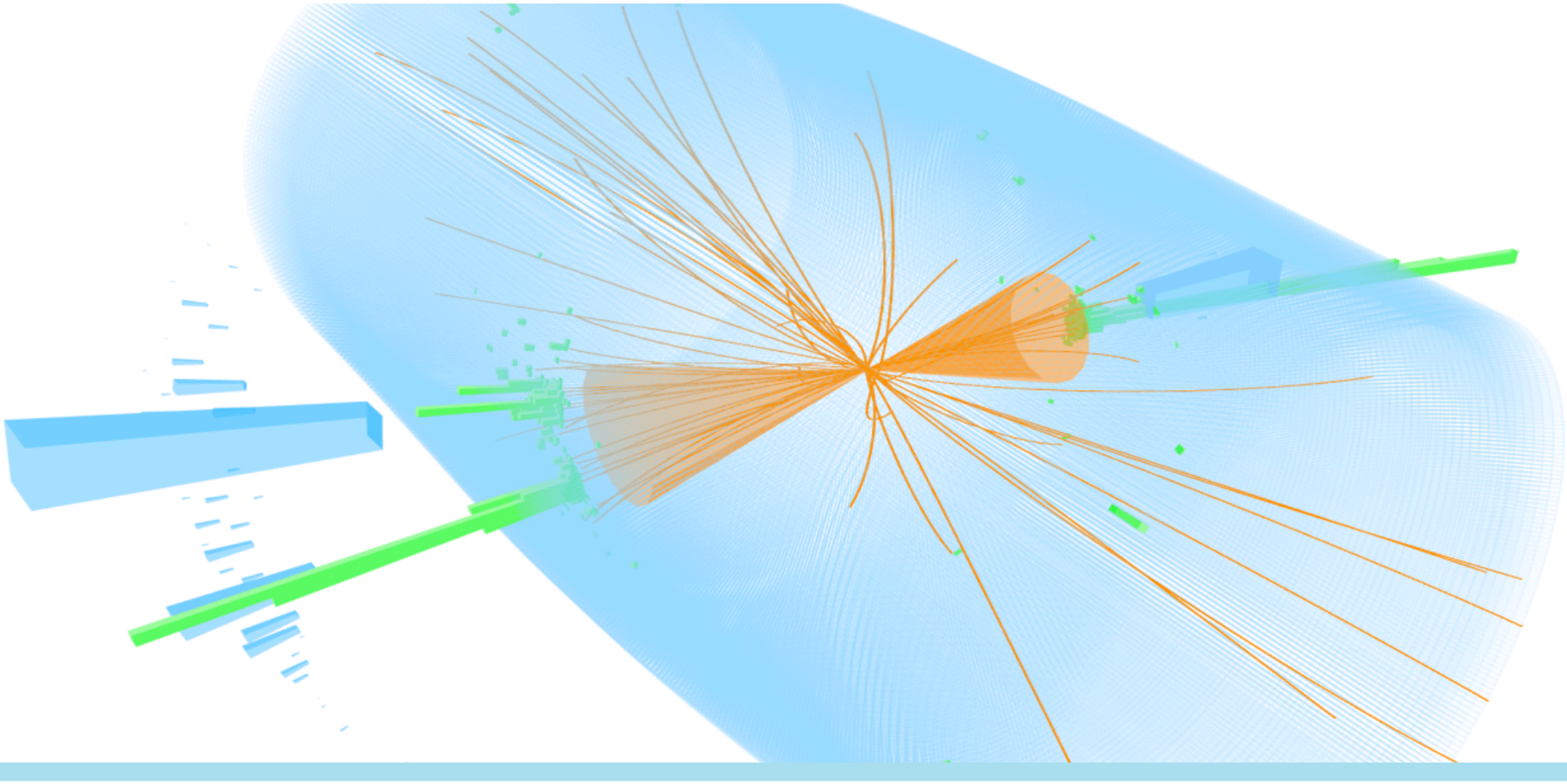
arxiv:2207.00230

Conclusion

- Jet substructure methods are delivering results
- Boosted channels add value to LHC Higgs program
  - Across a diverse set of analyses
- Outlook: more final states!
  - Boosted H($\tau\tau$), H(WW) on the horizon
  - H(gluons) ?



"jet substructure"

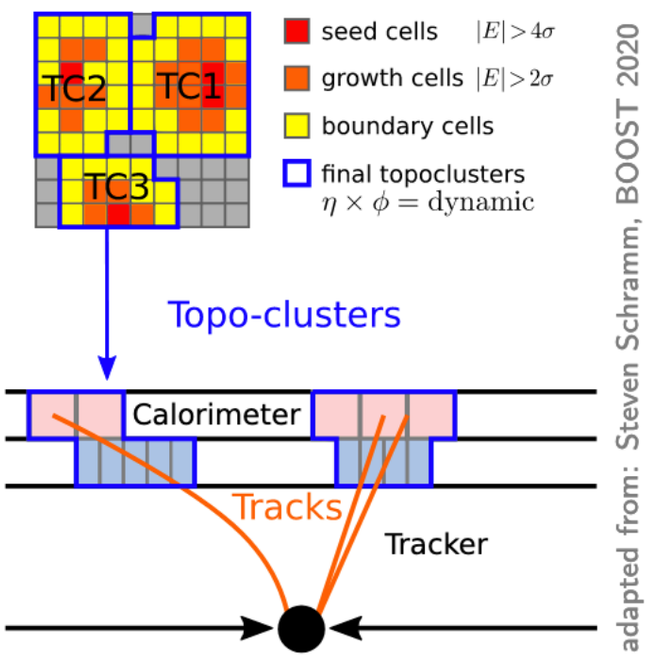


Backup

UFO jets

Calorimeter only:

- **LCTopo**: Topological calorimeter clusters



Combined with tracking:

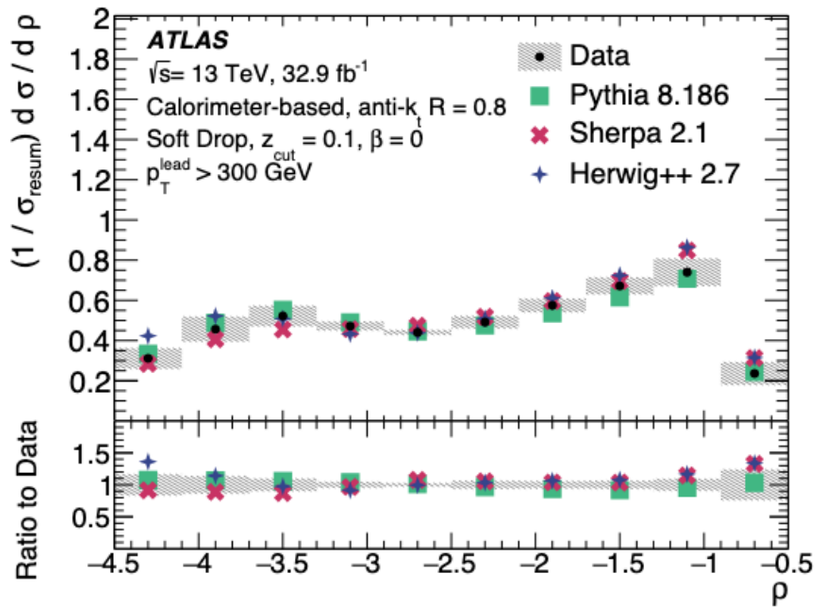
- **PFlow:** Particle Flow Objects
  - Low p_T: Use track 4-vector for charged particles, subtract energy from cluster 4-vectors
  - High p_T: Use cluster 4-vectors, ignore tracks
- **TCC**: Track Calo Clusters
  - Low p_T: Use cluster 4-vectors, ignore tracks
  - High p_T: Split clusters using tracks, get energy from clusters but angles from tracks

Combining PFlow and TCC:

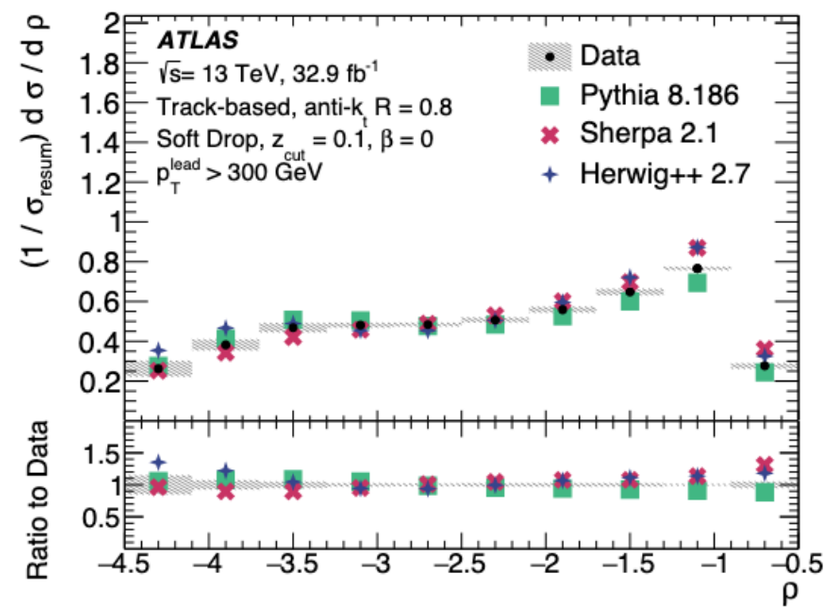
- **UFO** combines **TCC** and **PFlow** to achieve optimal performance over a broad kinematic (p_T) range

T. Fitschen

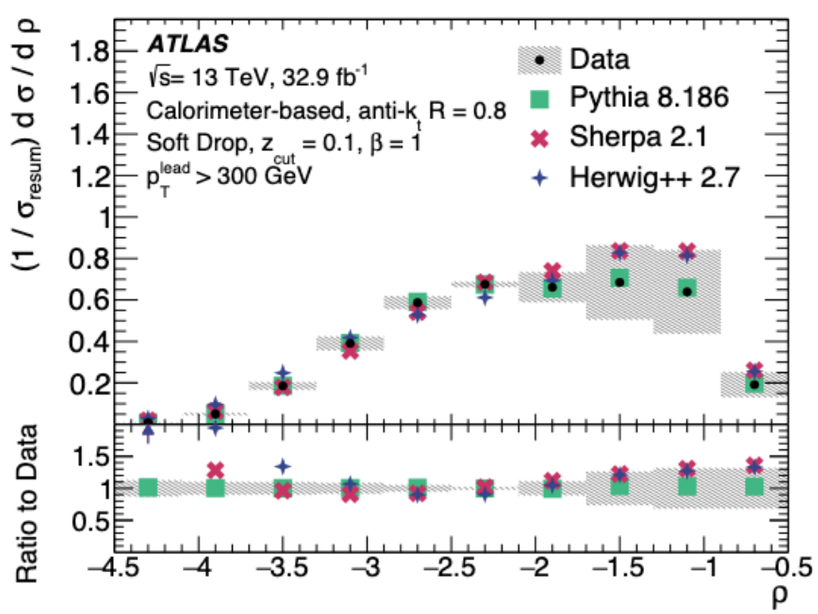
Softdrop observables (β=0,1)



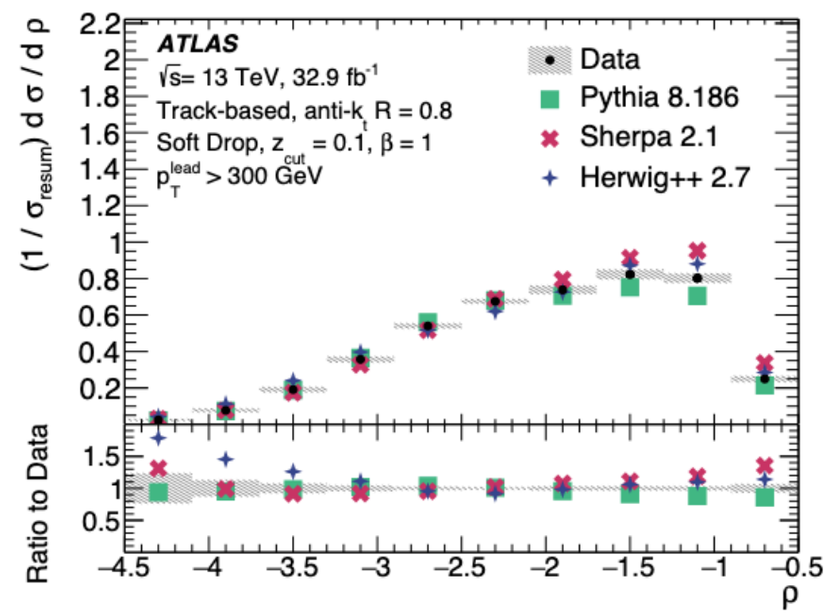
(a) $\beta = 0$, calorimeter-based



(b) $\beta = 0$, track-based



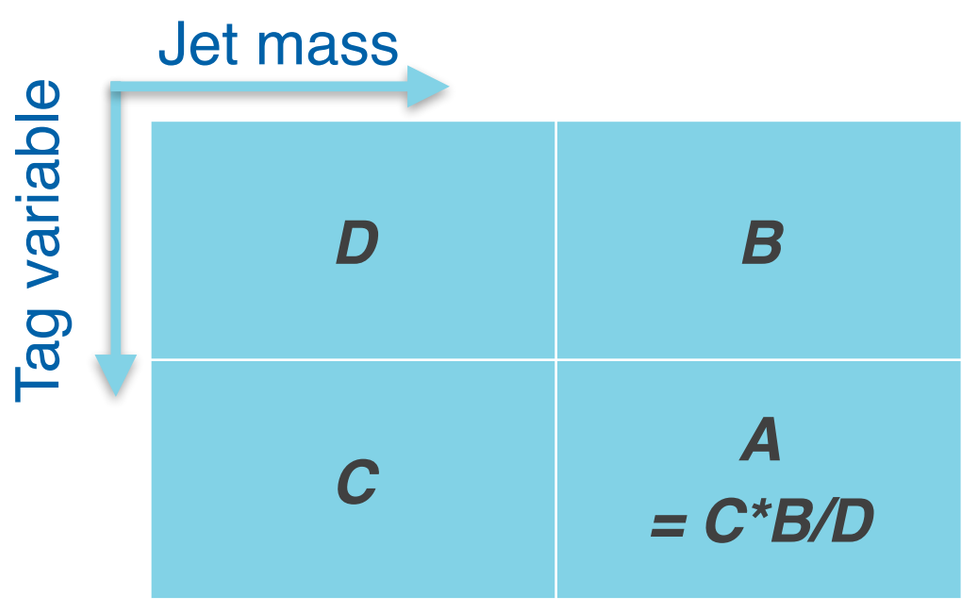
(c) $\beta = 1$, calorimeter-based

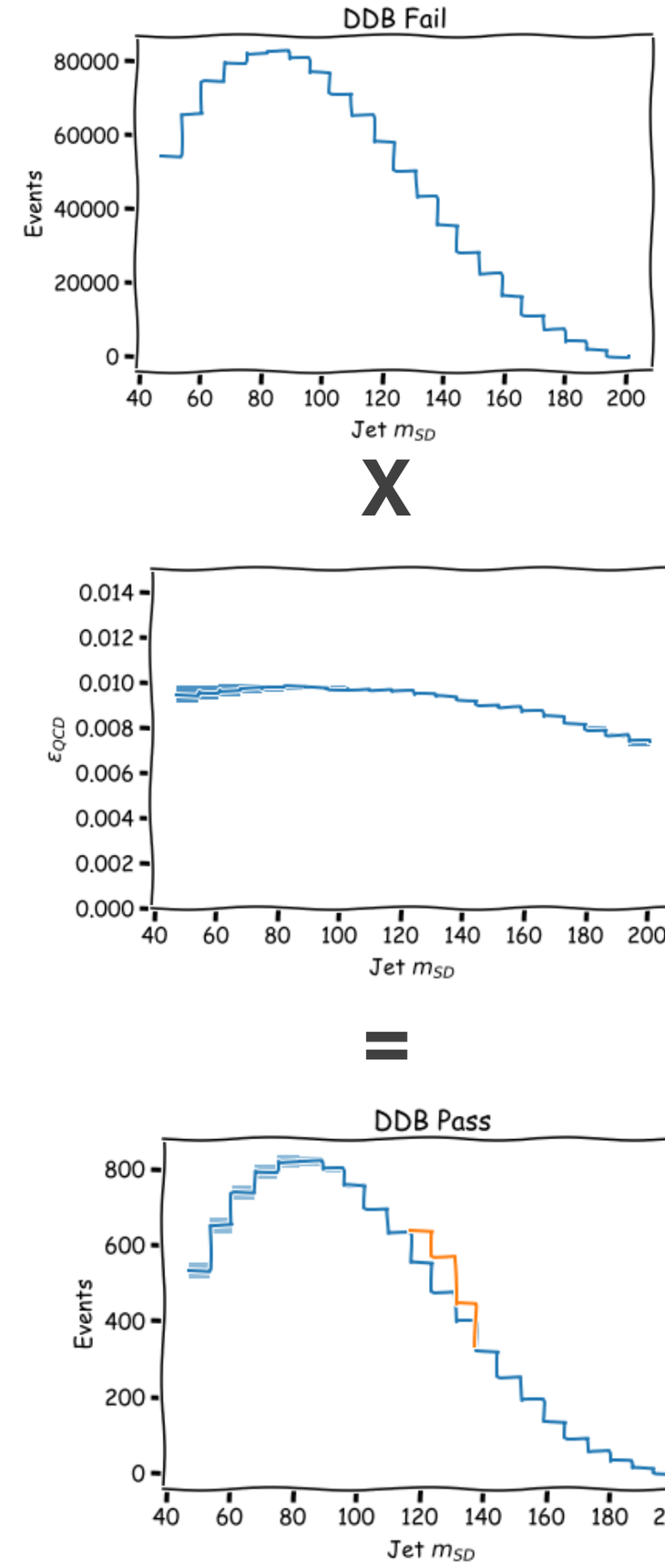


(d) $\beta = 1$, track-based

Mass decorrelation

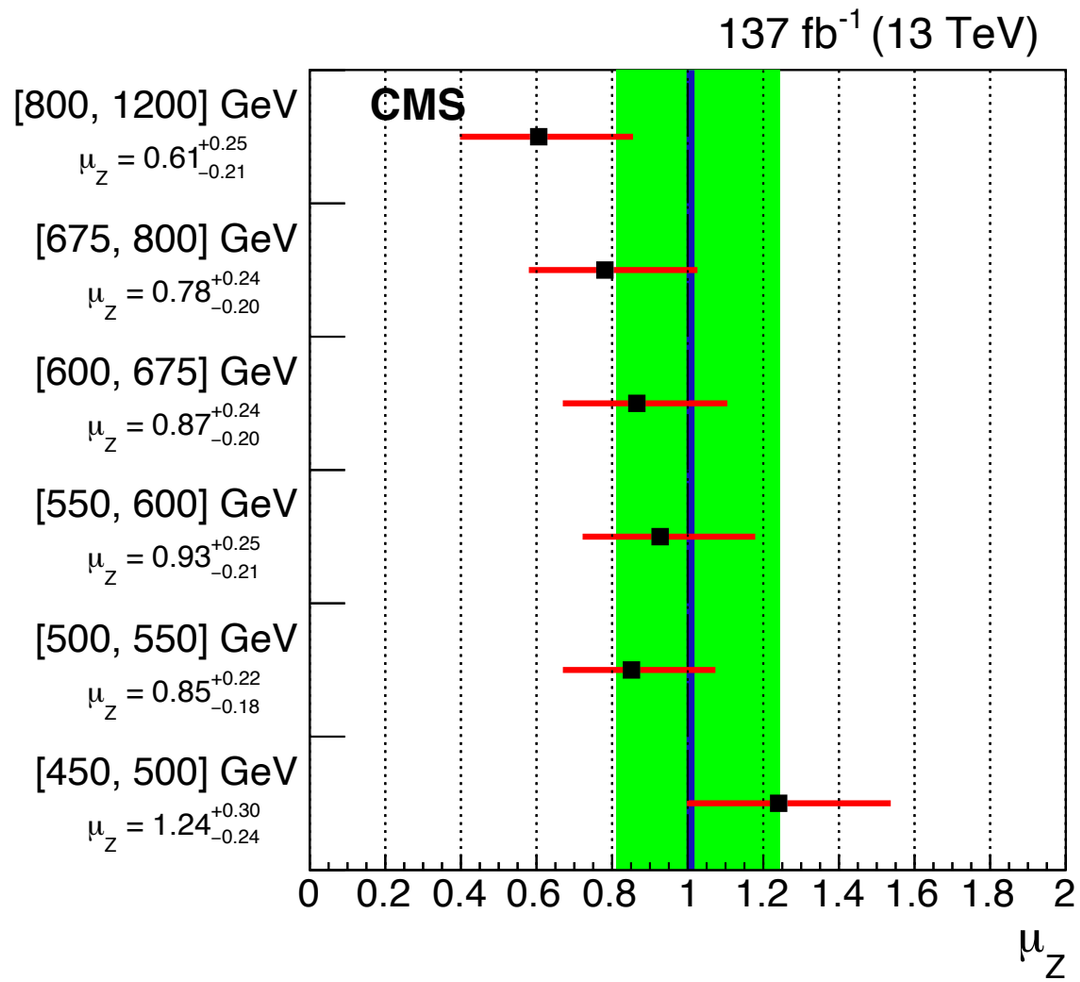
- Always one “impossible” background
  - Too expensive and unreliable to model from sim
  - For boosted H(bb), QCD. For VH(bb), V+HF
- Use data in mass sideband with “alphabet method”
- Required assumption:
  - Minimal correlation between tagger and mass in background
- Many ways to accomplish decorrelation in design
  - DDT 1603.00027, DisCo 2001.05310, adversarial training…
  - Engineer synthetic signals at many mass points



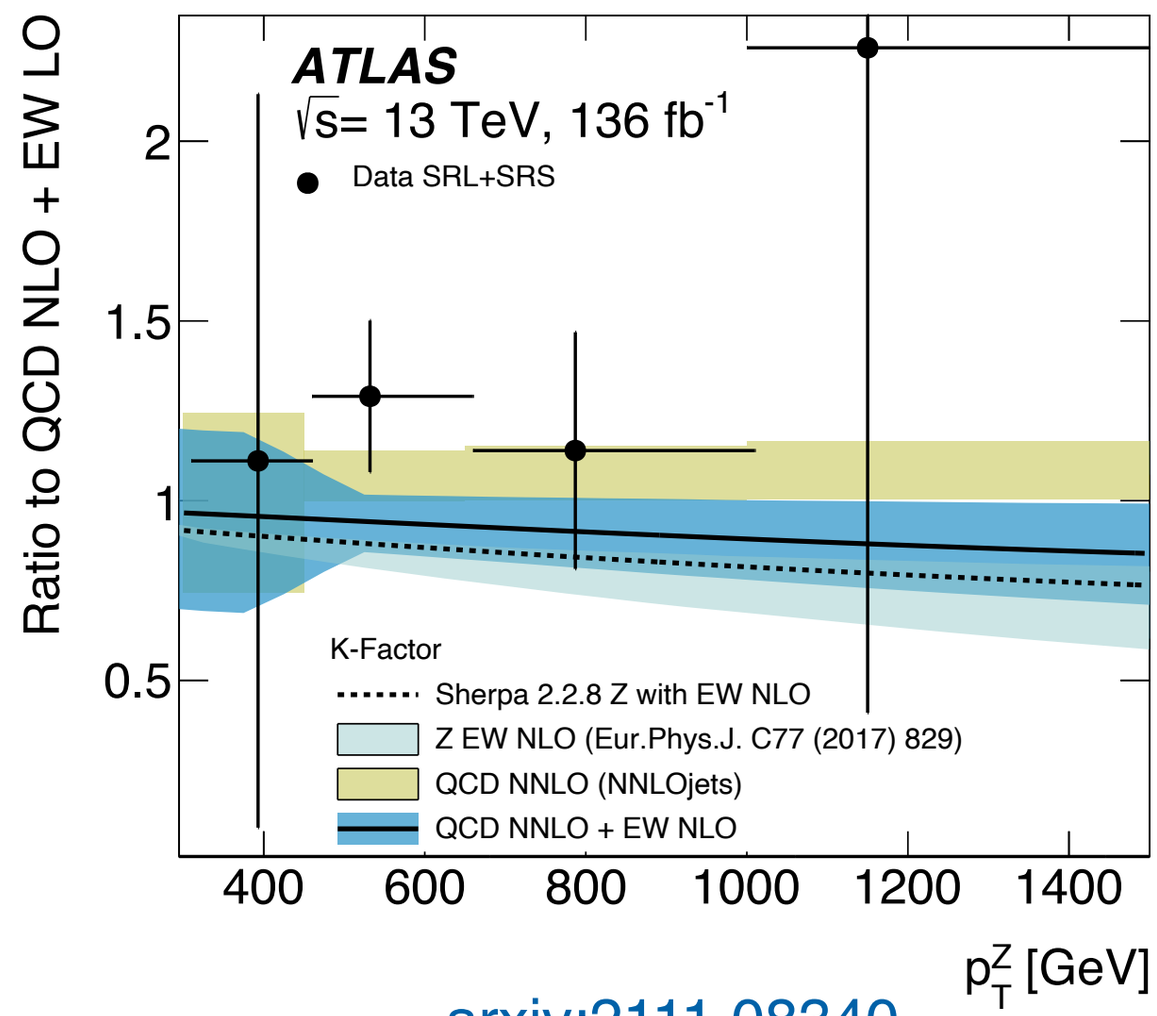


Boosted Z(bb)

CMS w.r.t. NLO (Z+1,2j FxFx) QCD + NLO EW



arxiv:2006.13251

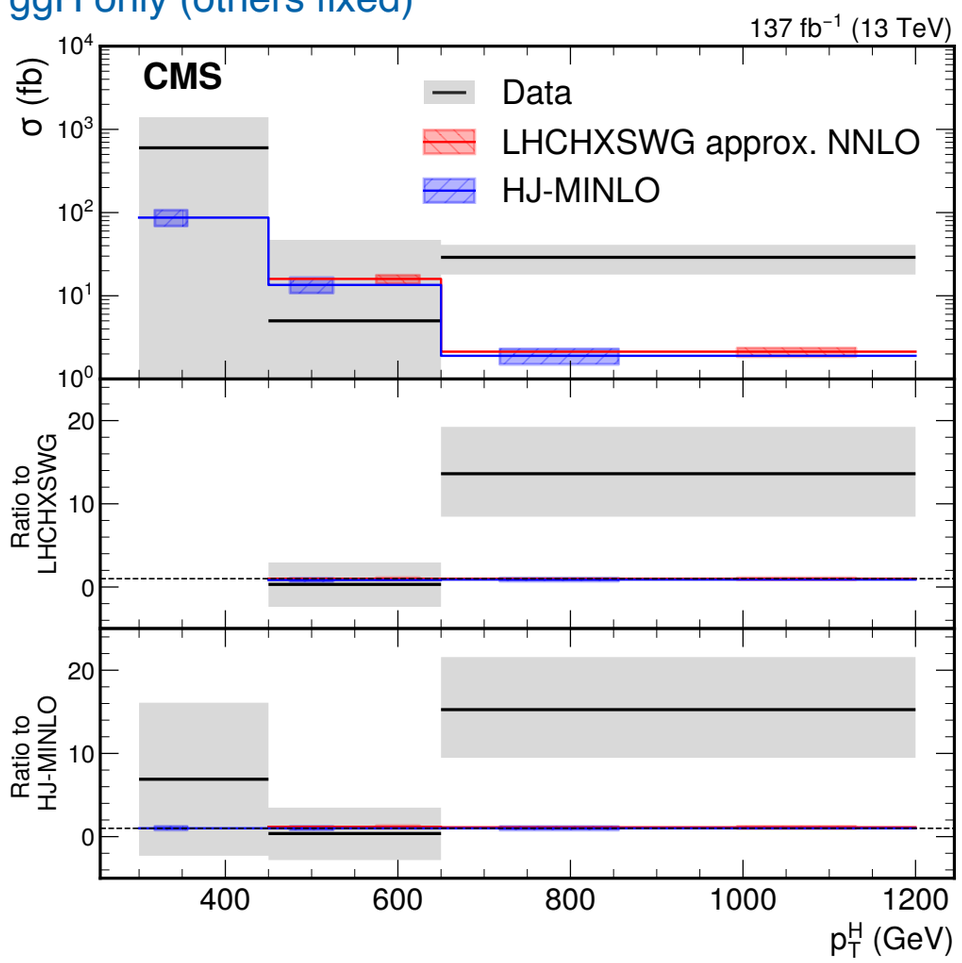


arxiv:2111.08340

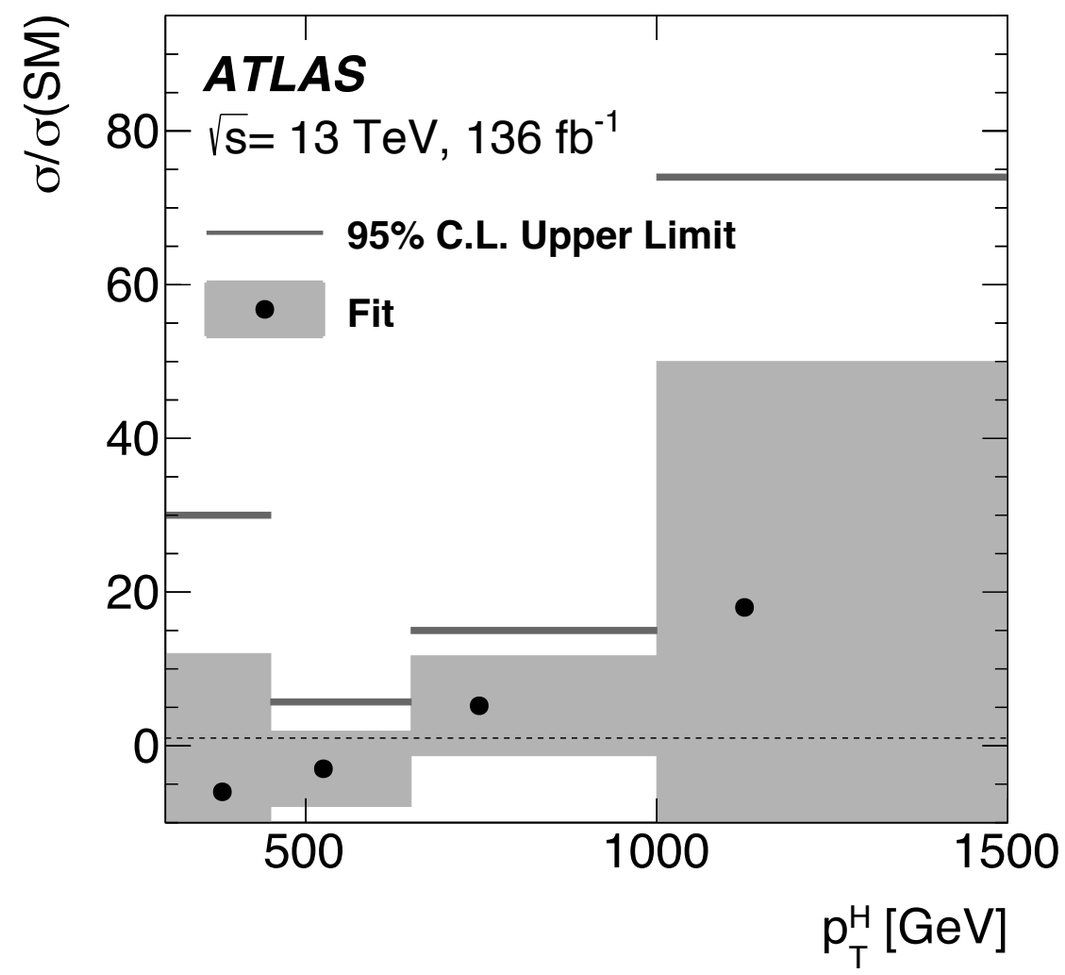
Fermilab

Boosted ggH(bb) unfolded

ggH only (others fixed)



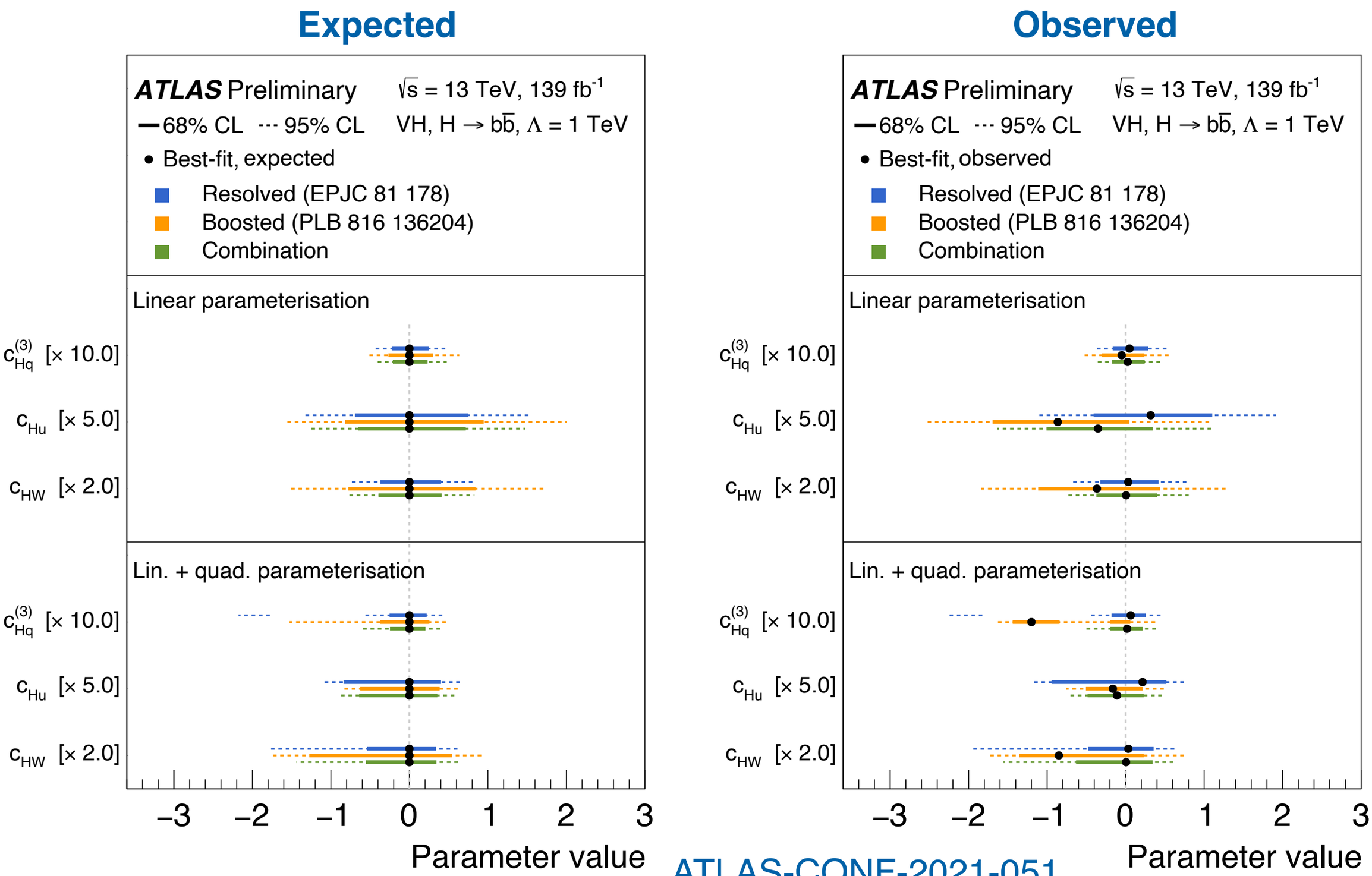
SM = HJ-MiNLO (ggH) + others



arxiv:2006.13251

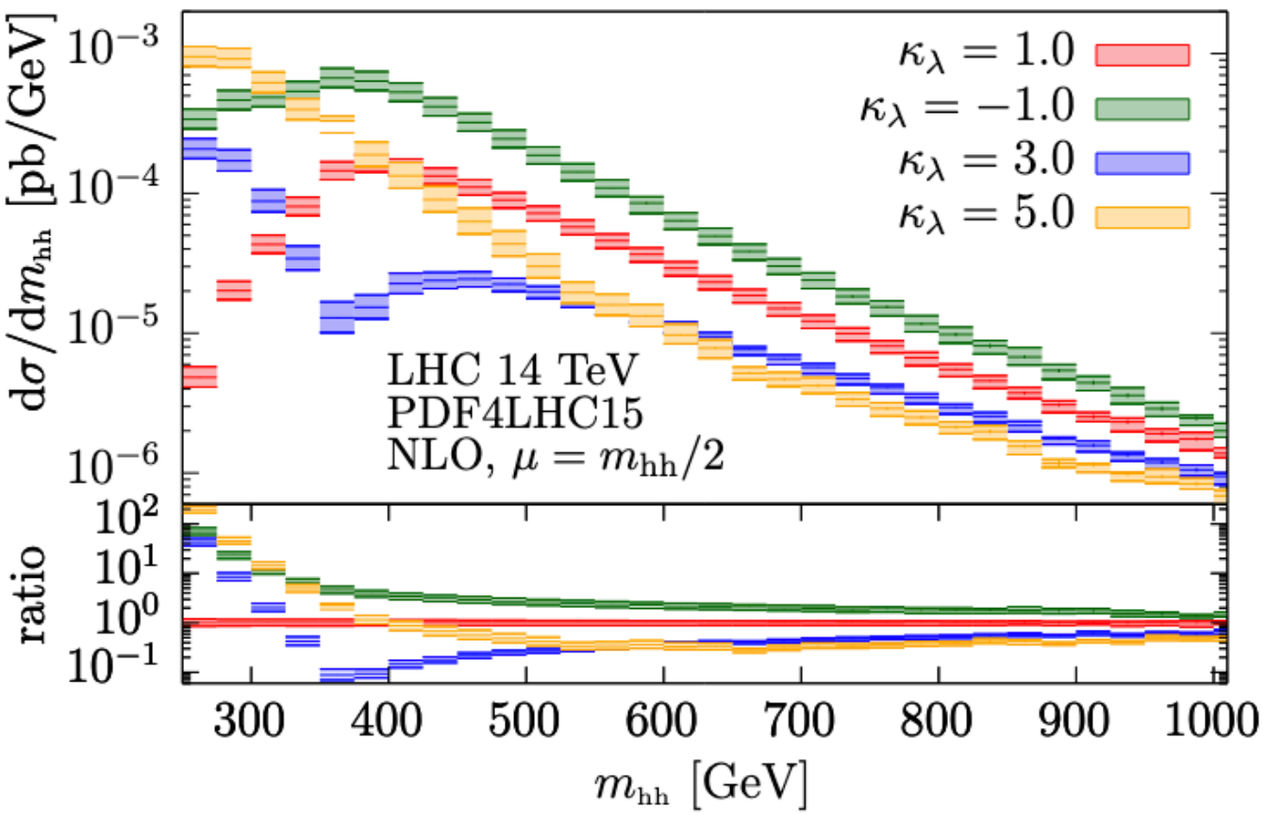
arxiv:2111.08340

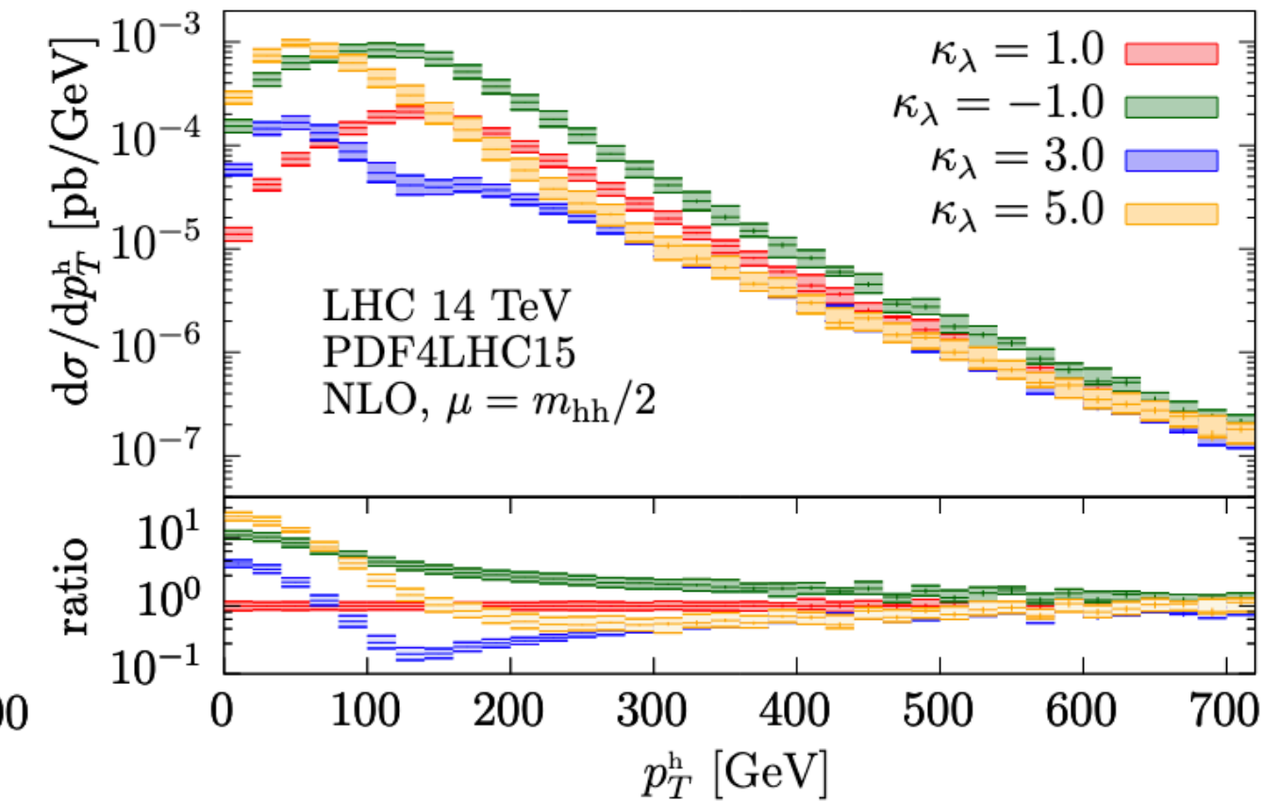
ATLAS resolved+boosted VH(bb)



ATLAS-CONF-2021-051

ggHH differential XS





arxiv:1903.08137