Analysis Methods in CMS

Zak Lawrence (Université catholique de Louvain)
on behalf of the CMS Collaboration

Higgs Hunting 2026
16^{th} -18^{th} September
Paris, France

CMS

fnrs
FREEDOM OF RESEARCH

UCLouvain

Introduction and Overview

Highlights of recent developments for CMS Higgs physics

- Jet Tagging:
  - Latest algorithms used for heavy flavour tagging and boosted jet tagging.
  - Developments toward multi task strategy, with performance of jet p_T regression
- Improved Triggers:
  - Application of flavour taggers to the HLT.
  - Dedicated parking streams optimised for HH topologies.
- Enhanced signal extraction and background estimation with machine learning techniques:
  - Several example cases looking at some of the latest Run 3 results in the search for Higgs pair production.

Jet tagging in CMS

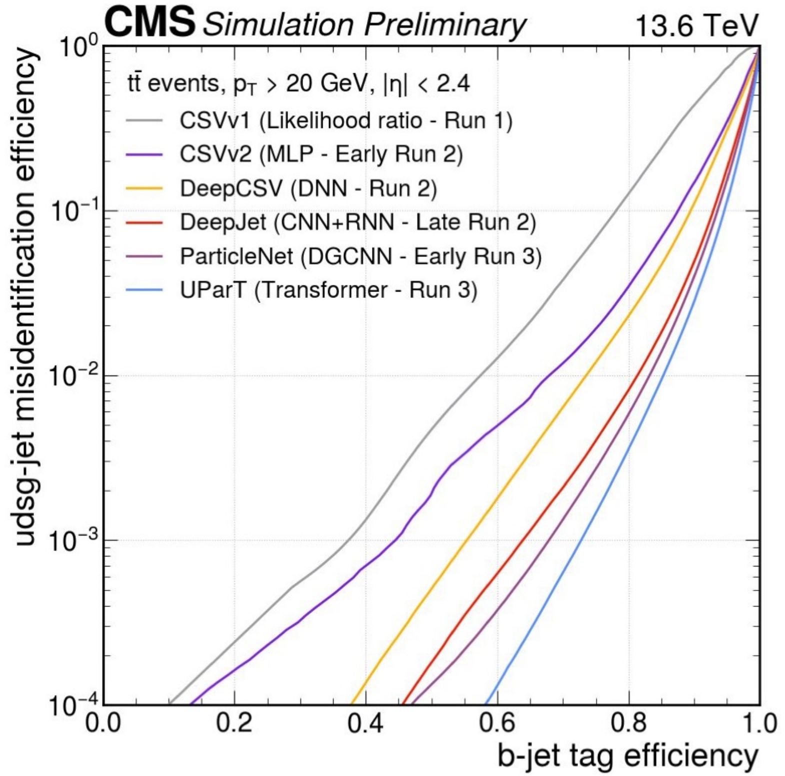
Improvements in jet tagging performance leverage developments in model architecture:

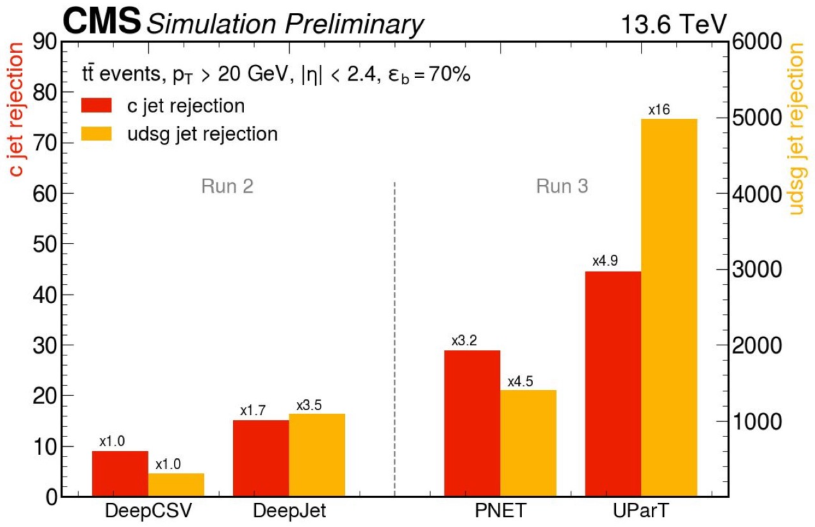
Particle Net: a Dynamic Graph Convolutional Network (DGCN) which represents jets as an unorder set of particles, particle cloud.

Particle Transformer: Adapts transformer attention mechanism with particle interaction features, giving particle attention block.

Tagging for both AK4 and AK8 jets are shifting towards a more unified approach to tagging:

- Multitask: models perform jet identification, mass or p_T regression, jet resolution regression.
- Multiclass: models provide outputs for a wider range of jet classes.

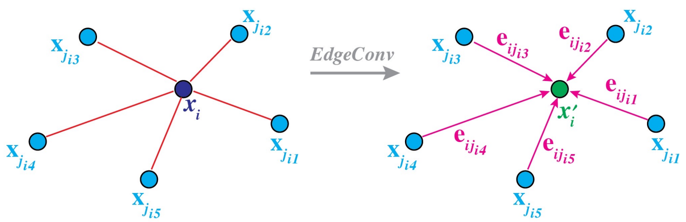


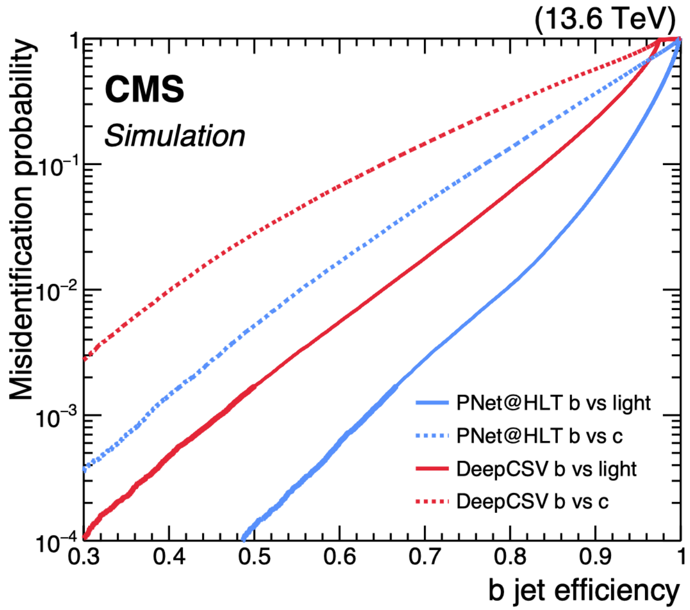


A unified approach for jet tagging in Run 3 (CMS DP -2024/066)

[Jet Tagging via Particle Clouds (H Qu, L Gouskos 2020)]

Particle Net

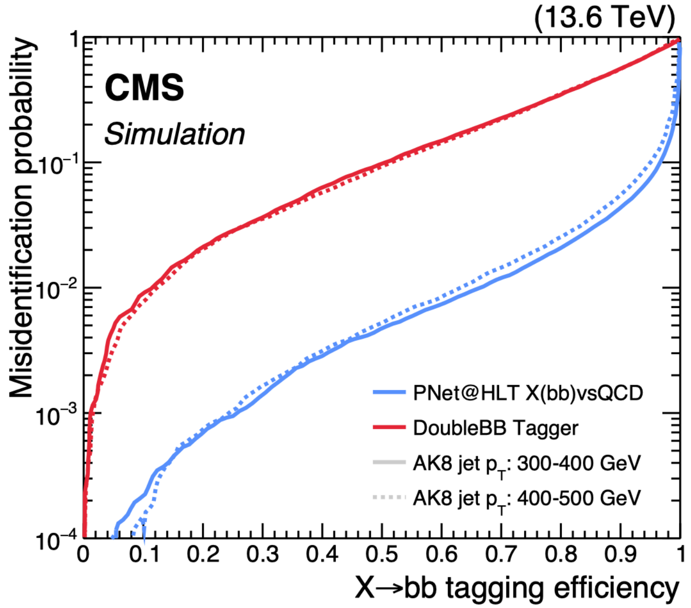




Developed at the end of Run 2, two independent trainings for AK4 and AK8 jet tagging.

Particle cloud representation of jets provides:

- More efficient at incorporating low level jet features than ordered collections.
- Explicitly respects permutation symmetry.



Model is implemented into the HLT system of CMS for both AK4 and AK8 jets.

- The AK4 jet PNET@HLT algorithm outperforms previous method and increases the background rejection by around a factor of 5.
- The AK8 PNET@HLT $H \rightarrow b\bar{b}$ algorithm is the first tagger capable of achieving the performance to design a highly efficient trigger for merged $H \rightarrow b\bar{b}$ decays.

CMS DP -2023/021
CMS DP -2023/050
CMS-HIG-24-010 (2026)

Impacts of improved trigger

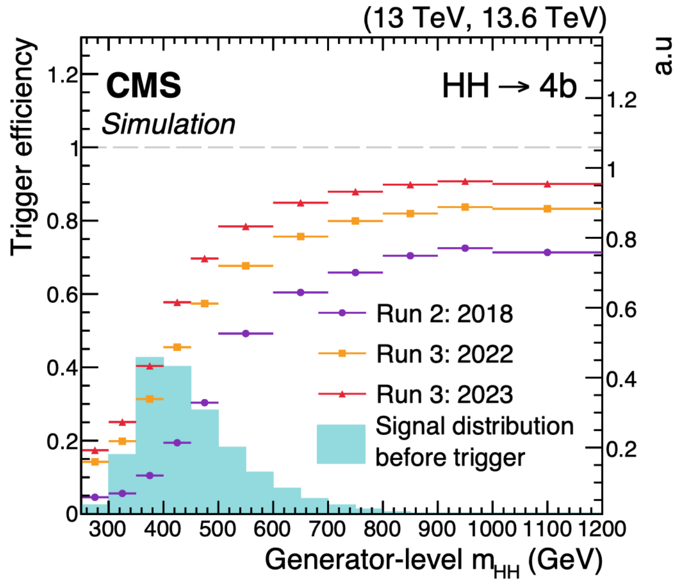
Application of Particle Net to the HLT allows for di-Higgs analyses to increase signal efficiency through improved trigger selections:

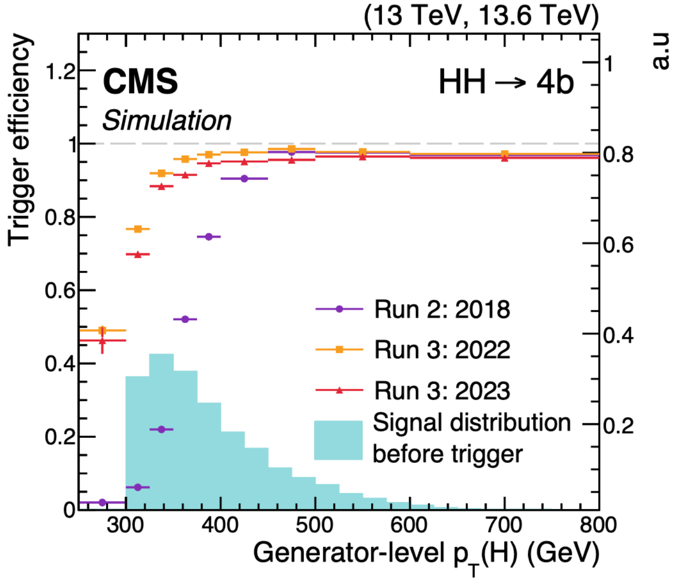
Trigger Requirements	Run 2	Run 3
Jet Multiplicity	4	4
Jet p_T Treshold	40 GeV	30 GeV
b-tagged jets	3	2

Example $HH \rightarrow 4b$ analysis:

Signal efficiency: **18% → 45%**

CMS-HIG-24-010 (2026)





Dedicated "Parking" strategy implemented in 2023 allowed for an increase in the trigger rate and signal efficiency increase of 22%.

In boosted topologies AK8 version of PNET@HLT allows for a significant improvement in signal efficiency (~90%).

Significant improvement compared to Run 2 triggers in the region $p_T(H) < 500$ GeV.

PNet applied to τ_h triggers, giving better ID than DeepTau (backup).

Unified Particle Transformer

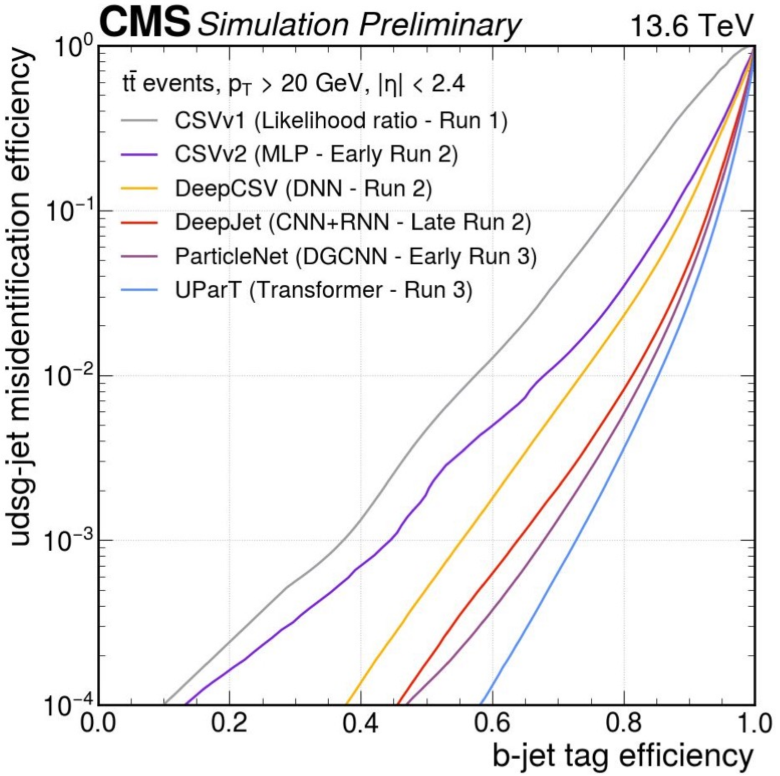
UParT is the latest unified jet-tagging model used by CMS, based on the **particle transformer architecture**. Performing several jet-based tasks:

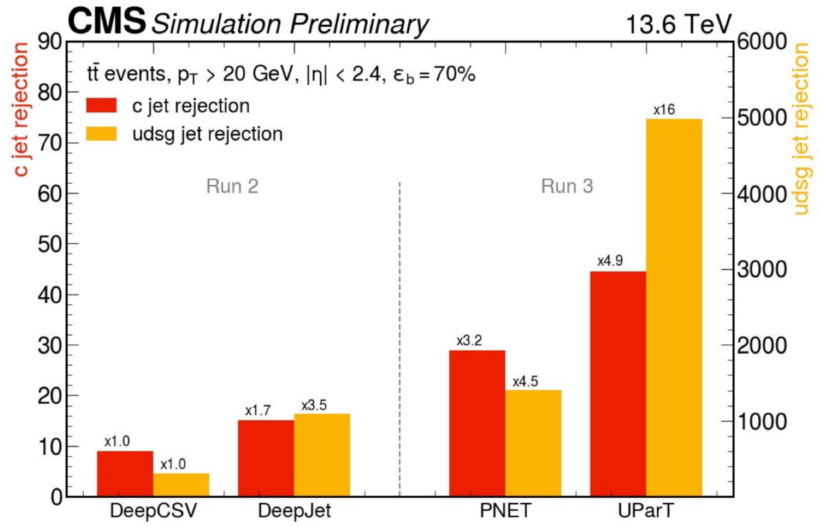
- Heavy flavour and τ_h identification.
- Flavour aware jet energy regression and jet energy resolution.
- A dedicated s-jet classifier, enabling s-quark jet tagging in CMS for the first time.

Adversarial training strategy used to enhance the robustness against mismodelling between simulation and data.

- Input featured perturbed in a direction that degrades model predictions.
- The model is trained to maintain performance on both nominal and perturbed inputs.
- Improves the robustness of the model while retaining similar performance on nominal samples.

Achieves **state of the art performance for both b-jet and c-jet identification** with improvements in rejection of relevant backgrounds.



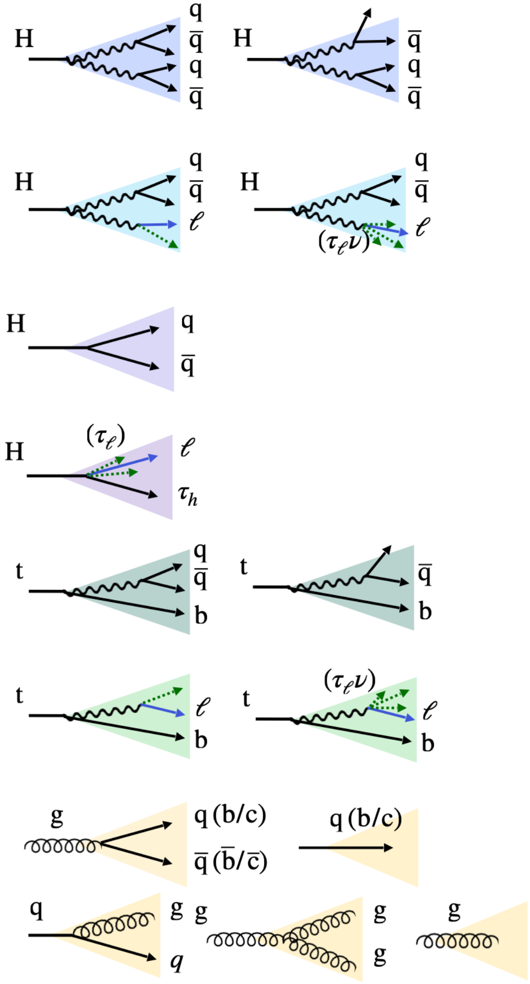


A unified approach for jet tagging in Run 3 (CMS DP -2024/066)

Global Particle Transformer

[Particle transformers for identifying Lorentz-boosted Higgs bosons decaying to a pair of W bosons (2026)]

Process	Final state		Flavor	Classes
H→WW (all-hadronic)	$q\bar{q}q\bar{q}$	⊗	0c / 1c / 2c	3
	$q\bar{q}q$			3
H→WW (semileptonic)	$e\nu q\bar{q}$	⊗	0c / 1c	2
	$\mu\nu q\bar{q}$			2
	$\tau_e\nu q\bar{q}$			2
	$\tau_\mu\nu q\bar{q}$			2
	$\tau_h\nu q\bar{q}$			2
H→$q\bar{q}$			$b\bar{b}$	1
			$c\bar{c}$	1
			$s\bar{s}$	1
			$q\bar{q}$ (q=u/d)	1
H→$\tau\tau$	$\tau_e\tau_h$			1
	$\tau_\mu\tau_h$			1
	$\tau_h\tau_h$			1
t→bW (hadronic)	$bq\bar{q}$	⊗	1b + 0c / 1c	2
	bq			2
t→bW (leptonic)	$be\nu$	⊗	1b	1
	$b\mu\nu$			1
	$b\tau_e\nu$			1
	$b\tau_\mu\nu$			1
	$b\tau_h\nu$			1
QCD			b	1
			$b\bar{b}$	1
			c	1
			$c\bar{c}$	1
			others (udsg)	1



Trained to classify jets from various all-hadronic and semileptonic Higgs and top quark decays from QCD backgrounds

Based on the Particle Transformer architecture.

- Inputs: takes 128 Particle Flow candidates and 10 SVs per AK8 jet
- Provides 37 output classes
- Additionally trained to regress the jet mass
  $$\mathcal{L}\left(m_{\text{true}}, m_{\text{pred}}\right) = \ln(\cosh((m_{\text{pred}} - m_{\text{true}})/\text{GeV}))$$
- Mass decorrelated training
- Discrimination for $H \to b\bar{b}$ given by:

$$T_{Xb\bar{b}} = \frac{\mathcal{P}(X \to b\bar{b})}{\mathcal{P}(X \to b\bar{b}) + \mathcal{P}(\text{QCD})}$$

Jet p_T and Mass regression

Current taggers employed in CMS are trained to perform multiple tasks:

- Jet p_T and Mass regression: Implemented with a log-cosh log and used to predict the ratio between truth-level jet p_T (mass) to raw jet p_T (mass).
- Jet energy resolution: Implemented through quantile regression.

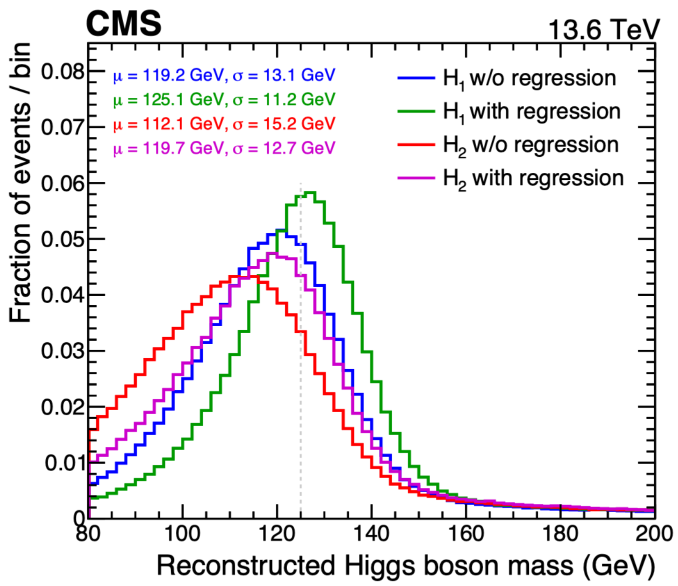
Classification	Regression	Quantile regression (resolution estimation)

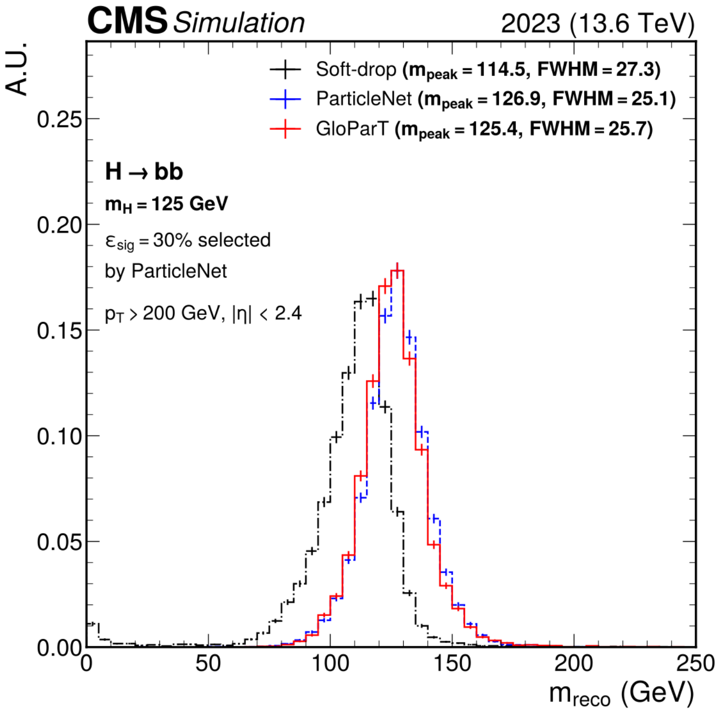
$$L = CatEntropy(x, x_{truth}) + \gamma_{regr} \cdot log(cosh(y - y_{truth})) + \gamma_{quantile} \cdot [p_{0.16}(z - z_{truth}) + p_{0.84}(z - z_{truth})]$$

Allows for a flavour aware prediction of the jet energy regression and resolution estimation from the network.

Regression improves the jet p_T resolution by 10-20%, and reduces the flavour dependence of jet energy corrections

- PNet jet p_T regression applied to $H \rightarrow b\bar{b}$ reconstruction improves the mass resolution by 20% in resolved topology.
- GloParT mass regression provides a 9% improvement in mass resolution for boosted topology.

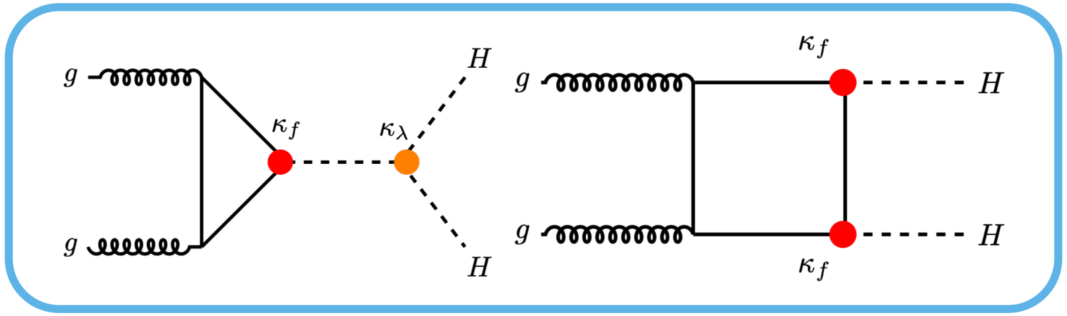




Methodological improvements for di-Higgs searches

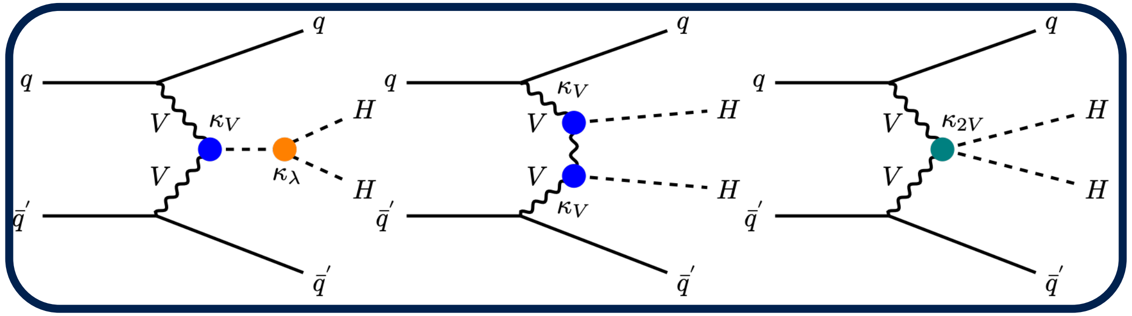
Search for Higgs pair production provides a direct probe of Higgs self interaction and shape of Higgs potential.

$$V(\mathcal{H}) = \frac{1}{2}{m_H}^2\mathcal{H}^2 + \lambda v \mathcal{H}^3 + \frac{1}{4}\lambda\mathcal{H}^4$$



Gluon fusion (ggf)

Improvements in tooling and **methodological improvements** contribute to **increased sensitivity** for searches for Higgs pair production for Run 3 analysis.



Vector boson fusion (VBF)

*	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

4b: Has the largest branching ratio, large multijet background.

$b\bar{b}WW$: Second highest branching ratio, final states with at least one lepton are cleaner.

$b\bar{b}\tau\tau$: Relatively large branching ration, cleaner final state.

Primarily produced via ggf

SM cross section ~1000 times smaller than single Higgs production

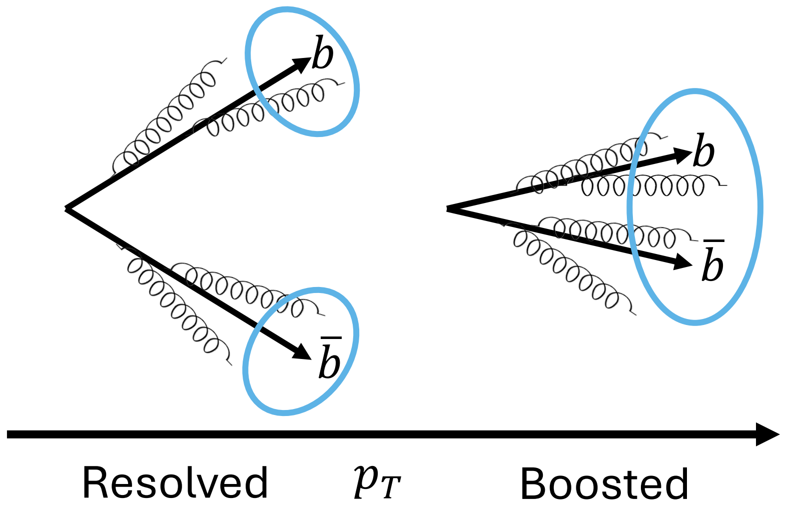
***Covered in this talk**

Improved results on HH → 4b | Overview

The $4b$ decay mode is amongst the most sensitive channels, with the largest branching fraction.

Key challenges:

1. Separating HH signal from overwhelming multijet background.
2. Precise modelling of background contribution.



Challenges addressed by advances in triggering, heavy flavour tagging and ML tools:

- Background modelling: Using DNNs to perform multidimensional reweighting in data driven method.
- Signal extraction: Using classifiers to increase signal to background ratio.

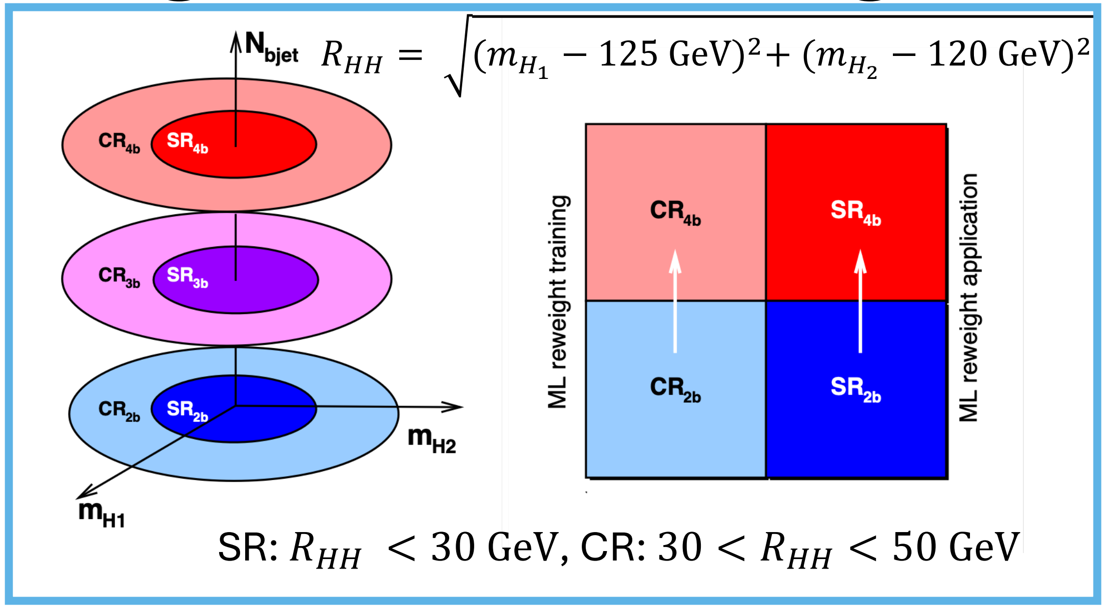
Analysis considers both **resolved** and **merged** topologies.

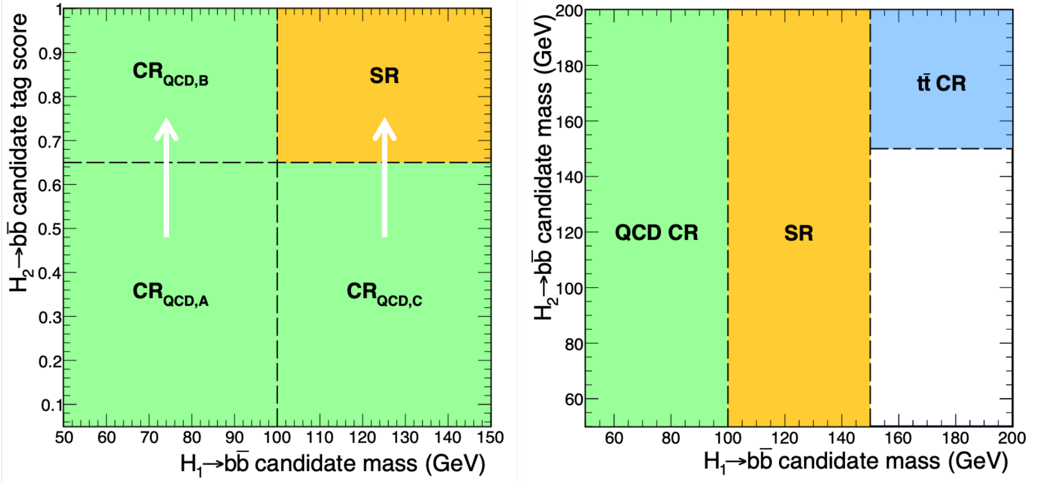
- Resolved topology has larger statistics but larger background.
- Merged topology has less statistics and less background, additionally enhances sensitivity to rare processes and $\kappa_{2V} \neq 1$.

Improved results on HH → 4b | Background Modelling

Background must be predicted to ~5% precision due to low expected signal contribution.

- Not possible to use MC, instead use data driven methods.
- Both analyses make use of DNNs to perform reweighting using an ABCD method.
- Use **ensemble of 20 models to reduce variance** from training.





$t\bar{t}$ forms ~20% of background

Resolved Topology:

Events are split based on b-tag multiplicity and reconstructed Higgs mass → directly impacted by improved tagging and regression.

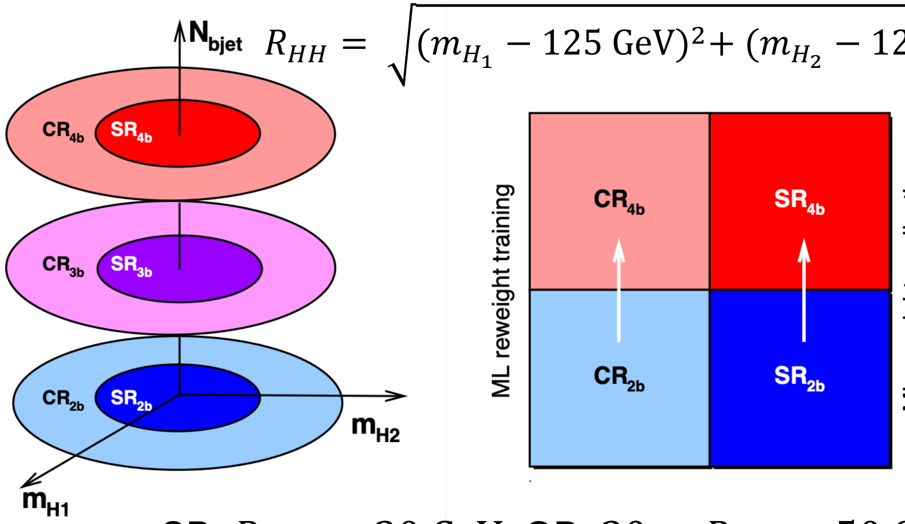
DNN learns to differentiate 2b from 4b events in control regions, weights applied to 2b data for prediction in 4b signal region.

$$w = \frac{\mathcal{P}(\mathrm{CR}4b)}{\mathcal{P}(\mathrm{CR}2b)}$$

Improved results on HH → 4b | Background Modelling

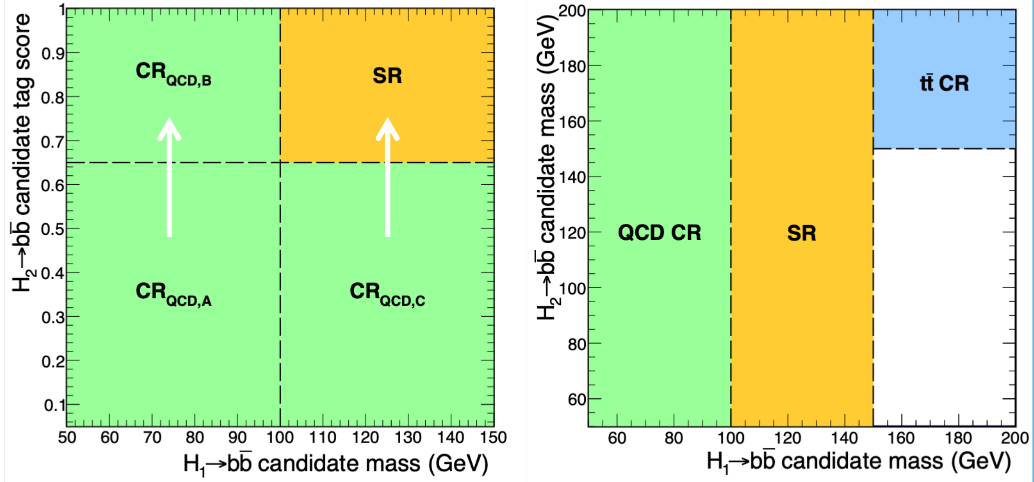
Background must be predicted to ~5% precision due to low expected signal contribution.

- Not possible to use MC, instead use data driven methods.
- Both analyses make use of DNNs to perform reweighting using an ABCD method.
- Use **ensemble of 20 models to reduce variance** from training.



$R_{HH} = \sqrt{(m_{H_1} - 125\text{ GeV})^2 + (m_{H_2} - 120\text{ GeV})^2}$

SR: $R_{HH} < 30$ GeV, CR: $30 < R_{HH} < 50$ GeV



$t\bar{t}$ forms ~20% of background

Merged Topology:

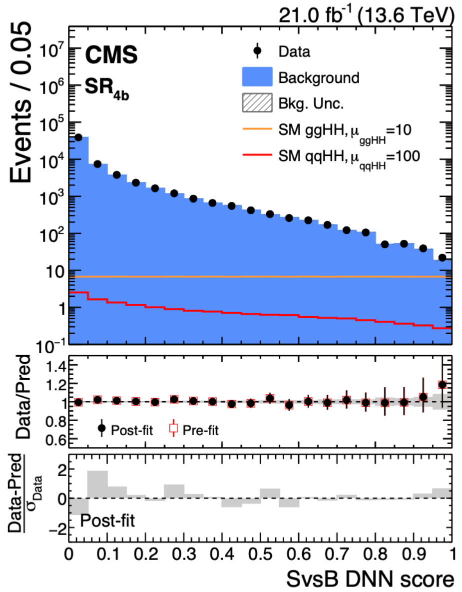
Events split on the tagging score of the Higgs candidate and regressed Higgs mass.

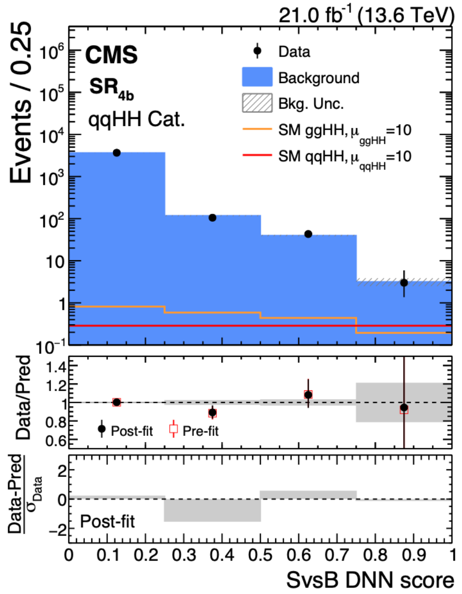
DNNs trained in mass sideband region $50 < m_{reg}(H_1) < 100$ to separate high Xbb score from low Xbb score data, weights applied to control region in mass window.

$$w = \frac{P(\mathrm{CR_{data,B}}) - P(\mathrm{CR}_{t\bar{t},\mathrm{B}})}{P(\mathrm{CR_{data,A}}) - P(\mathrm{CR}_{t\bar{t},\mathrm{A}})}$$

CMS-HIG-24-010 (2026)

Improved results on HH → 4b | Signal Extraction





Events split based on production mode, either $ggHH$ or $qqHH$

- **$ggHH$**: Classifier to separate HH → 4b signal from background prediction, composed of two feature encoders:
  1. DNN: Takes event level features as input.
  2. Transformer: Takes 4-vectors of 5 leading b-tagged jets and pairwise features.
- **$qqHH$**: DNN trained to separate $ggHH$ and $qqHH$ signals along with background, three output nodes ($\mathcal{P}_{ggHH}, \mathcal{P}_{qqHH}, \mathcal{P}_{bkg}$):
  1. $\mathcal{D}_{ggHH} = \mathcal{P}_{qqHH}/(\mathcal{P}_{qqHH} + \mathcal{P}_{ggHH})$, used to separate $qqHH$ and $ggHH$ events.
  2. $\mathcal{D}_{qqHH} = \mathcal{P}_{qqHH}/(\mathcal{P}_{qqHH} + \mathcal{P}_{bkg})$, used to reject background events.

In the Resolved topology the model outputs are used directly as inputs in maximum likelihood fit.

Models trained with signals containing a mixture of simulated κ_λ values to retain sensitivity to non-SM κ_λ values.

Merged topology discussed in backup.

CMS-HIG-24-010 (2026)

Improved results on HH → 4b | Results

Results benefit from the application of novel methods to the trigger, event selection, background modelling and signal extraction.

Upper limit on μ_{HH} of 4.4 at 95% CL.

- Improvement of more than a factor or two with an equivalent integrated luminosity to Run 2.

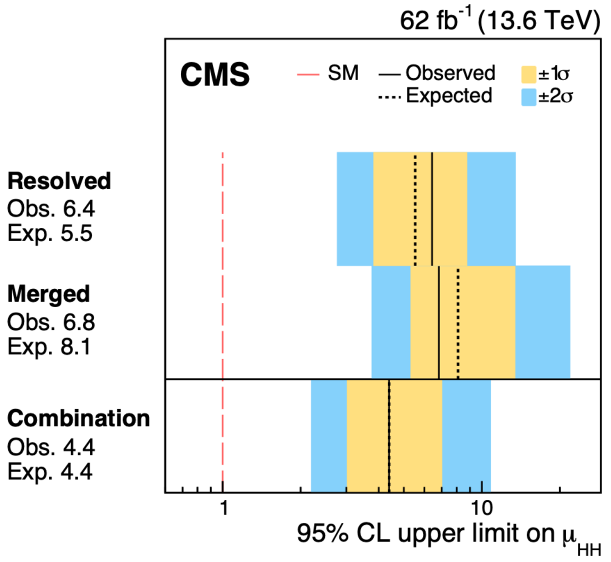
Cross section limit used to constrain at 95% CL:

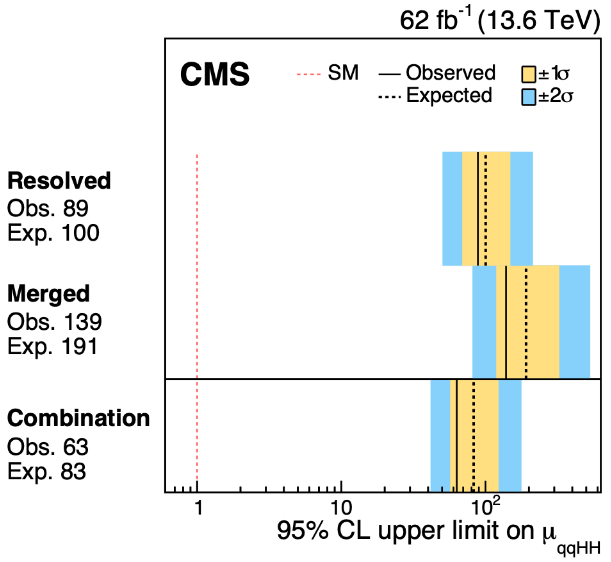
- κ_λ to $[-3.3, 9.7]([-3.4, 10.0])$
- κ_{2V} to $[0.63, 1.43]([0.54, 1.51])$

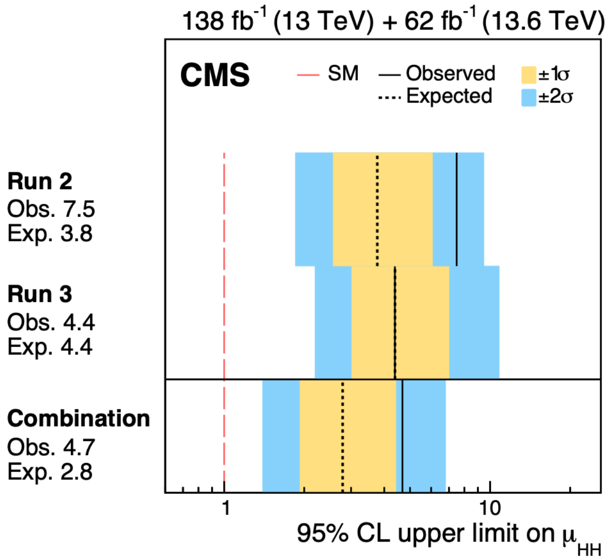
Combining the results with previous Run 2 measurements:

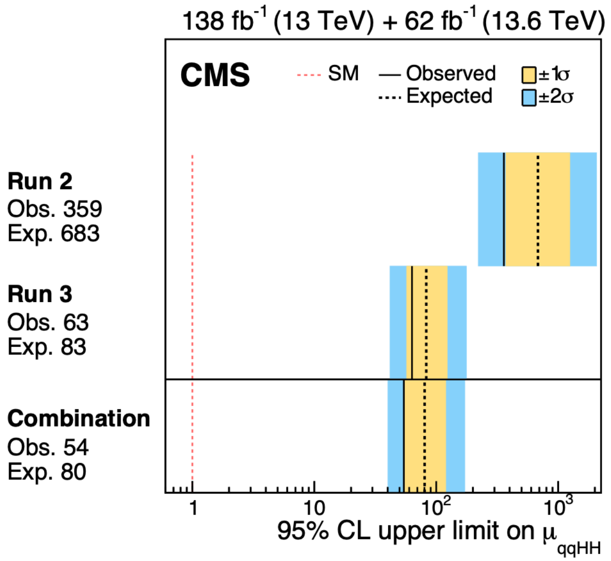
- upper limit on μ_{HH} of 4.7
- κ_λ to $[-4.1, 9.6]([-3.2, 9.8])$
- κ_{2V} to$[0.77\ 1.27]([0.69\ 1.35])$

These represent the most stringent constraints in the 4b final state to date.









Improvements in the search for $\mathrm{HH} \rightarrow \mathrm{b\bar{b}WW}$

Measurement in channel where one Higgs decays to $b\bar{b}$ and one to WW pair that decays leptonically.

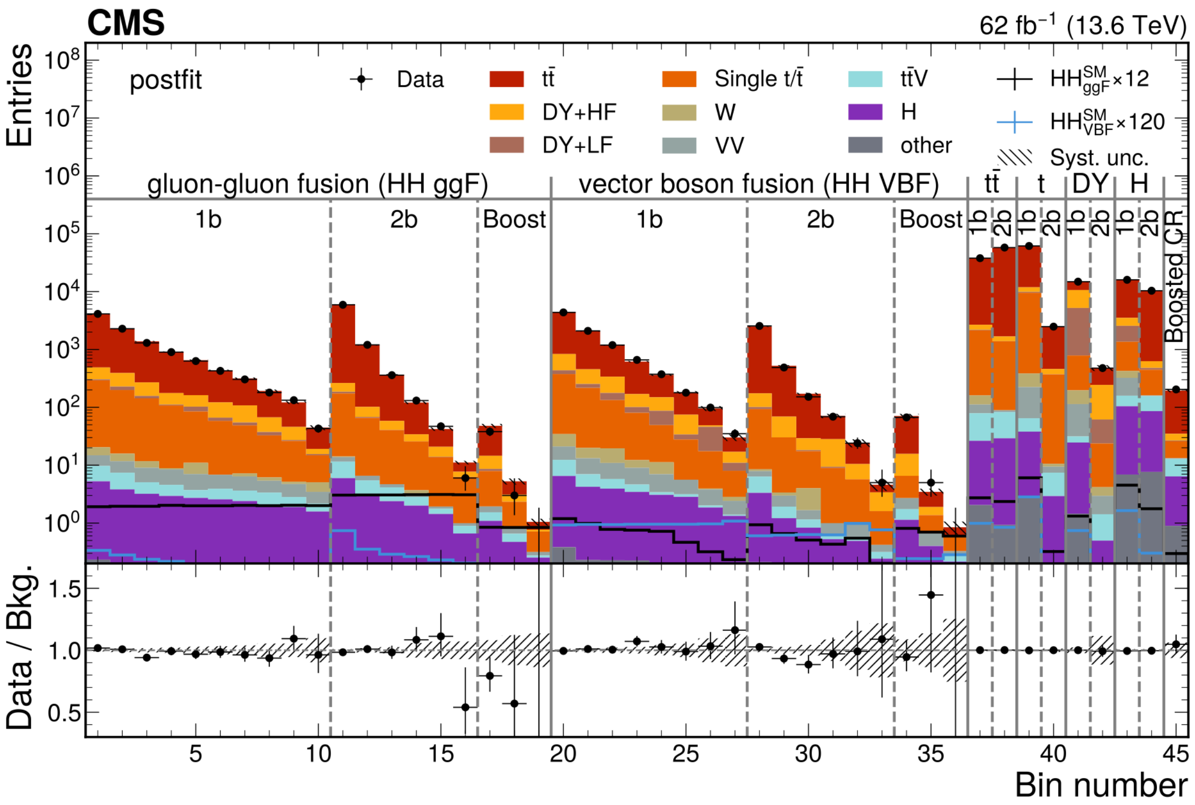
Benefits from improved event classification and updated trigger strategy.

Analysis makes use of a new classification strategy compared to Run 2 measurement:

- Multiclass NN: Designed to separate different signal and background processes giving two SRs and four background CRs
- Binary NNs: Two networks trained, one to separate HH ggF from backgrounds and one for HH VBF.

Improved classification strategy gave an improvement in sensitivity of around 50%.

Improvement in the e^+e^- and $e^{\pm}\mu^{\pm}$ trigger gives a 5% increase in inclusive signal and up to 17% in more signal like bins



Binning chosen so the output distribution of the signal is uniform.

Results of $\mathrm{HH} \rightarrow \mathrm{b\bar{b}WW}$ search

Measurement sets upper limit on μ_{HH} of 12.0 (18.5) observed (expected) at 95% CL.

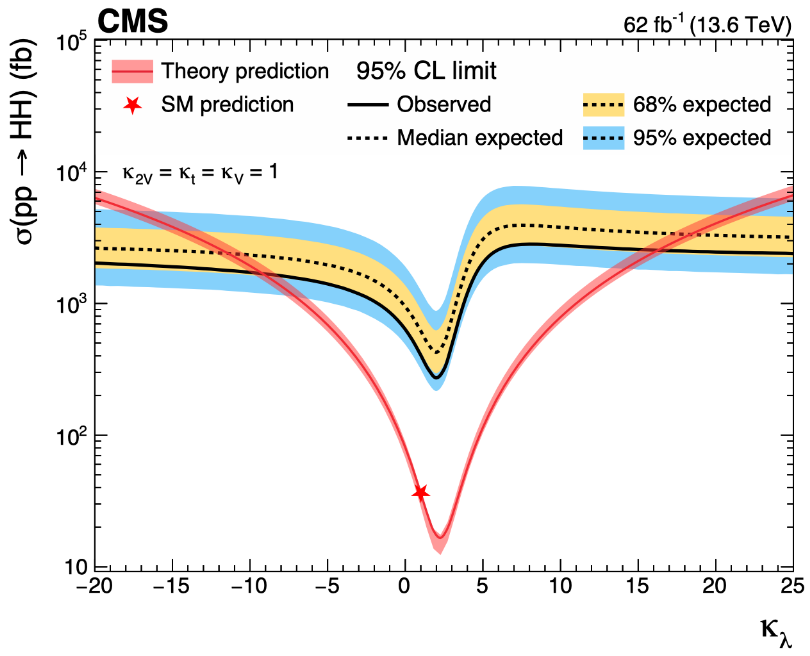
- Run 2 measurement set upper limit on μ_{HH} of 14.0 (18.0) observed (expected) at 95% CL

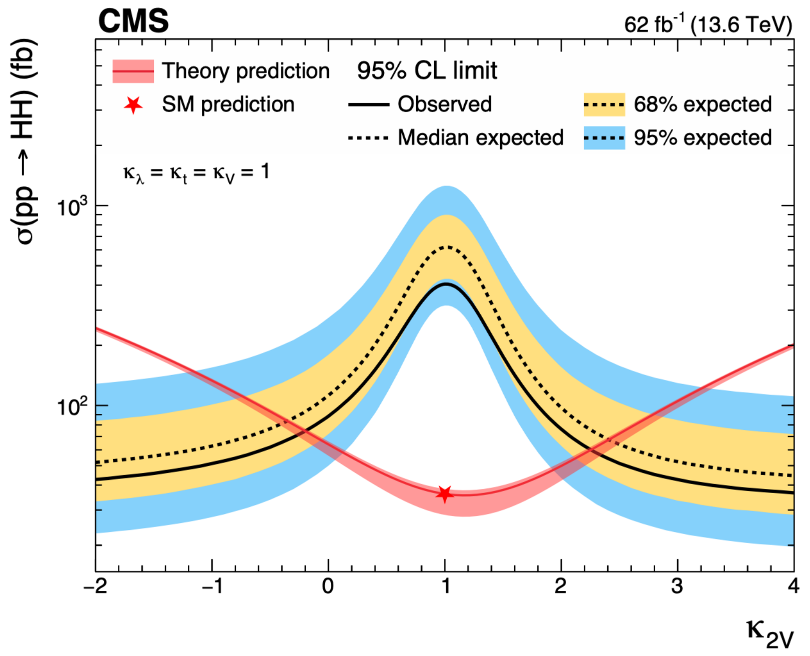
Cross section limit used to constrain κ_λ to $[-9.1, 15.7]$($[-13.3, 19.8]$) and κ_{2V} to $[-0.20, 2.25]$($[-0.48, 2.54]$) at 95% CL.

The measurement provides **comparable overall expected sensitivity to Run 2** despite the smaller dataset used.

- Improvements come from refined classification strategy, new triggers and improvements in b-tagging algorithms.

Largest uncertainty on the analysis comes from the limited statistical precision due to the small expected signal.



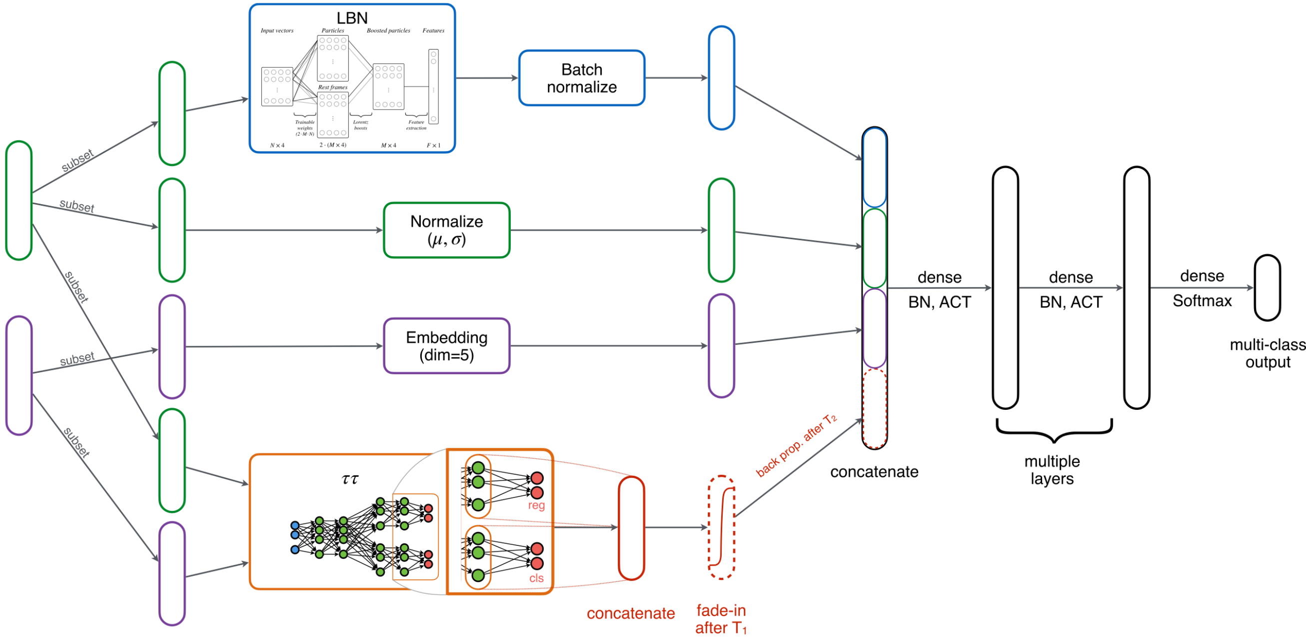


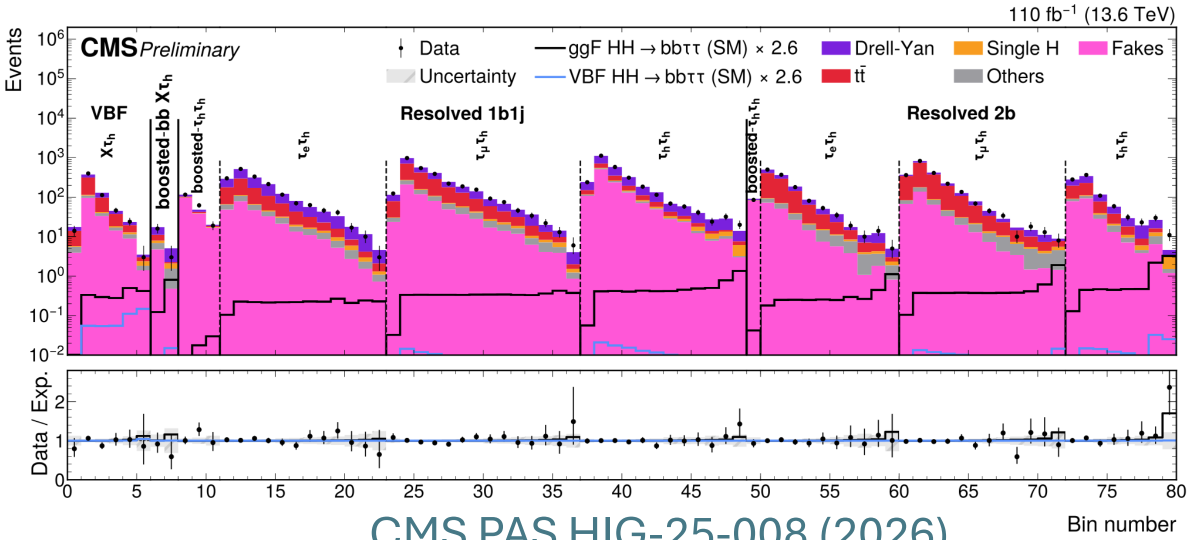
Improvements in the search for $\mathrm{HH} \rightarrow \mathrm{b\bar{b}\tau\tau}$

Measurement makes use of Run 3 data collected from **2022 – 2024**, giving a total integrated luminosity of **172 fb^{-1}**.

Analysis benefits from:

- **Improved trigger strategy**:
  - New trigger targeting two τ_h's + 1 jet.
  - New parking trigger implemented and extended to $b\bar{b}\tau\tau$ decay mode (4 jets with 1 τ_h and 1 b-tagged jet)
  - Dedicated triggers targeting $qqHH$ production.
- **Dedicated DNN selection of VBF jets**:
  - Uses same design as HH-bTag algorithm developed for Run 2 analysis but to select VBF jets.
- **Improved signal extraction** with new DNN architecture consisting of **three networks**:
  - Feature engineering Lorentz Boost Network
  - $\tau\tau$-mass regression network
  - Classifying DenseNet.





CMS PAS HIG-25-008 (2026)

Results of $\mathrm{HH} \to \mathrm{b\bar{b}}\tau\tau$ search

Analysis further improves the performance of boosted $\mathrm{H} \to \tau_h\tau_h$ events via jet substructure techniques:

- Identify two subjets within selected AK8 jets.
- If subjets satisfy kinematic conditions, pass to HPS algorithm as seeds for τ_h reconstruction.

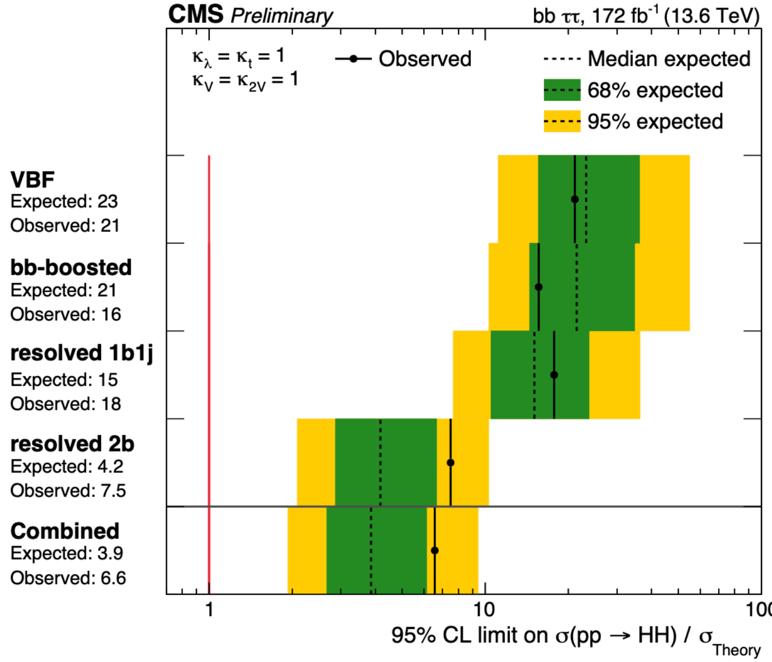
$$\boxed{p_T > 10\text{ GeV and } \frac{\max\left(m_{s_{j1}}, m_{s_{j2}}\right)}{m_{CA8\,jet}} < 2/3}$$

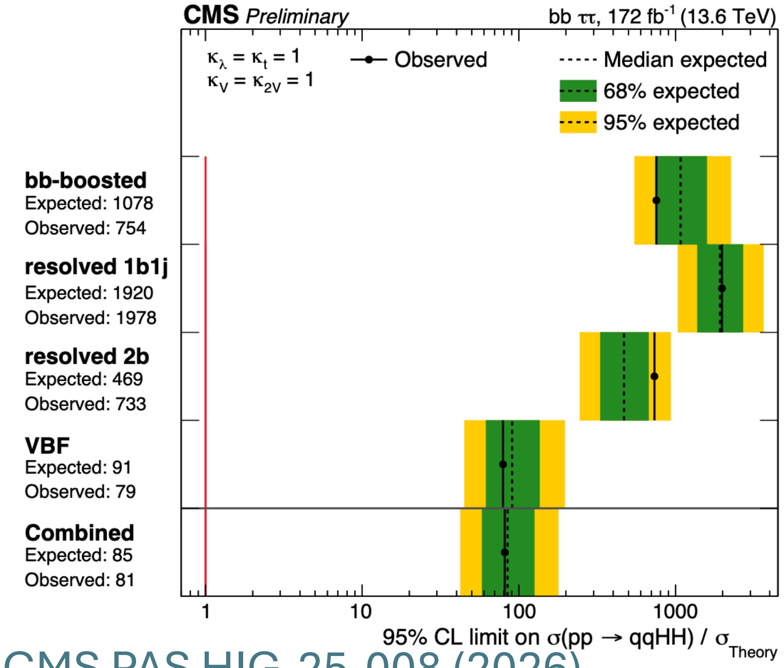
Measurement sets upper limit on μ_{HH} of 6.6 (3.9) observed (expected) at 95% CL.

- Corresponds to a 20% improvement when scaled to Run 2 luminosity.

Cross section limit used to constrain at 95% CL:

- κ_λ to $[-4.1, 11.2]([-2.2, 9.2])$
- κ_{2V} to $[-0.2, 2.3]([-0.2, 2.4])$





CMS PAS HIG-25-008 (2026)

Summary & Outlook

Latest generation of jet flavour taggers, including PNet and UParT, show state of the art performance on flavour tagging in resolved and boosted regimes.

- Taggers perform multiple tasks including jet p_T regression, which improves mass resolution and used for event classification in analysis.

Improved trigger strategies enhance signal acceptance within measurements.

- Incorporating PNet to perform tagging for the HLT.
- Dedicated parking streams optimised for HH topologies.

Machine learning techniques improve signal extraction through event classification and signal vs background separation.

All analyses show improvements exceeding luminosity gain!

Thanks For Listening!

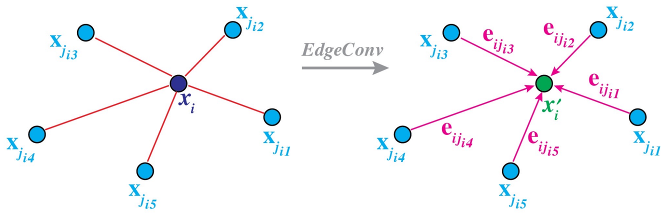
Backup

Particle Net | architecture

Jet Tagging via Particle Clouds (H Qu, L Gouskos 2020)

Particle Net is a Dynamic Graph Convolutional Network (DGCN) trained for jet tagging.

Jets are treated as an unordered set of particles, called a "**particle cloud**"



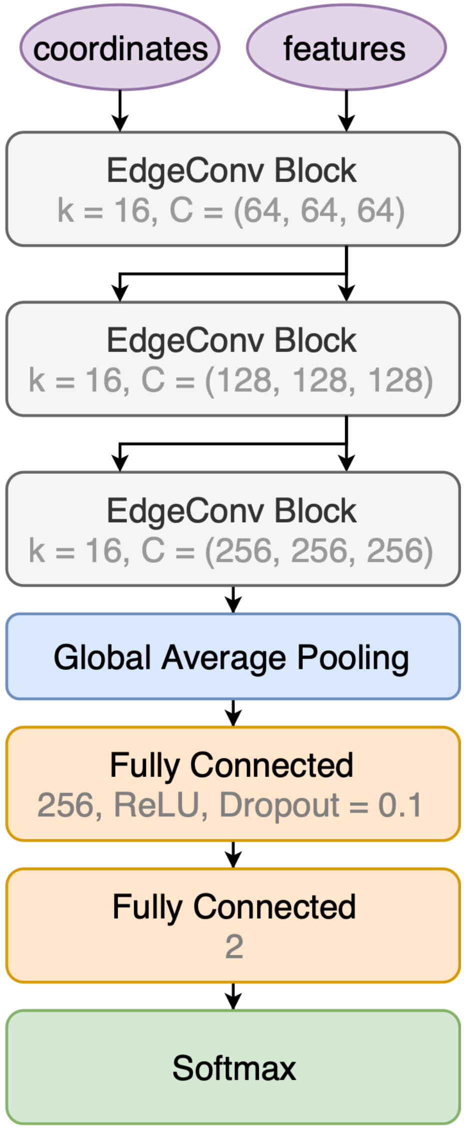
The network makes use of the Edge convolution block.

Consists of 3 EdgeConv blocks, which builds a graph from the k nearest neighbours using the particles "coordinates"

- 1st block uses physical coordinates in terms of η, ϕ.
- Subsequent blocks use learnt feature vectors.

Following this a channel wide global averaging is performed to aggregate features from across the particle cloud.

Fully connected layers follow and output for jet classification.



Particle Transformer | architecture

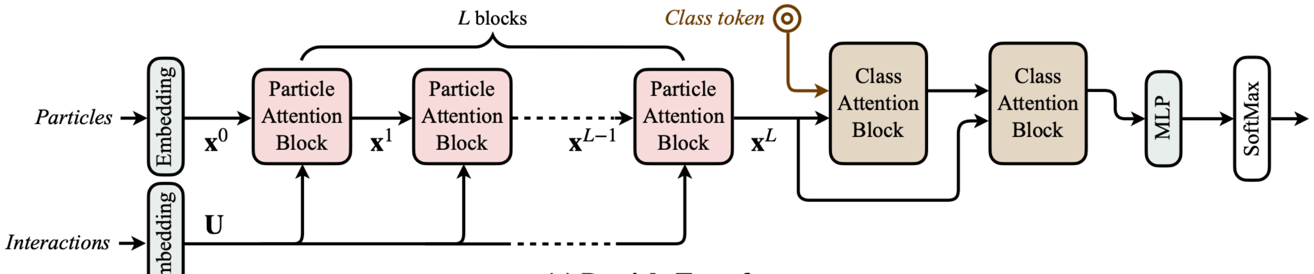
[Particle Transformer for Jet Tagging (H Qu, C Li, S Qian 2022)](#)

A transformer-based model for jet tagging which incorporates pairwise interaction features and a class token for enhanced classification performance.

- **Particle Features**: includes a set of C features for every particle of shape (N,C)
- **Particle Interaction**: includes C' features per pair of particles of shape (N, N, C')

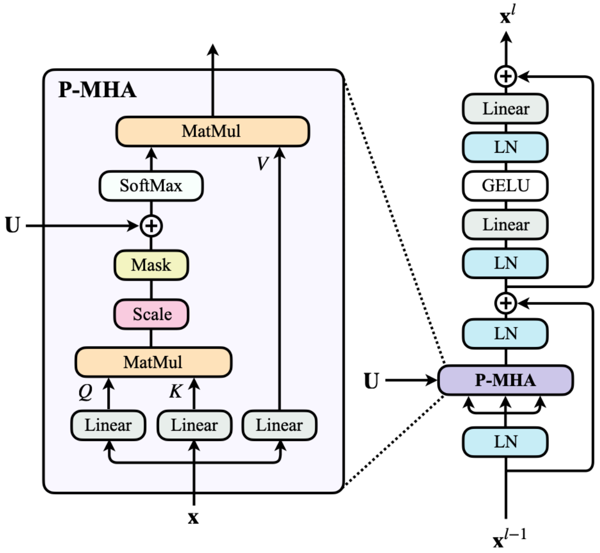
Pairwise Interactions: Architecture can support any interaction features; in this case features are derived from 4-vectors. These modify the attention weights in the particle attention block.

Class Token: Used as the query in the class attention block, attending to a concatenated class and particle embedding and giving an updated class token



(a) Particle Transformer

$$\Delta = \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2},$$
$$k_{\mathrm{T}} = \min(p_{\mathrm{T},a}, p_{\mathrm{T},b})\Delta,$$
$$z = \min(p_{\mathrm{T},a}, p_{\mathrm{T},b})/(p_{\mathrm{T},a} + p_{\mathrm{T},b}),$$
$$m^2 = (E_a + E_b)^2 - \|\mathbf{p}_a + \mathbf{p}_b\|^2,$$

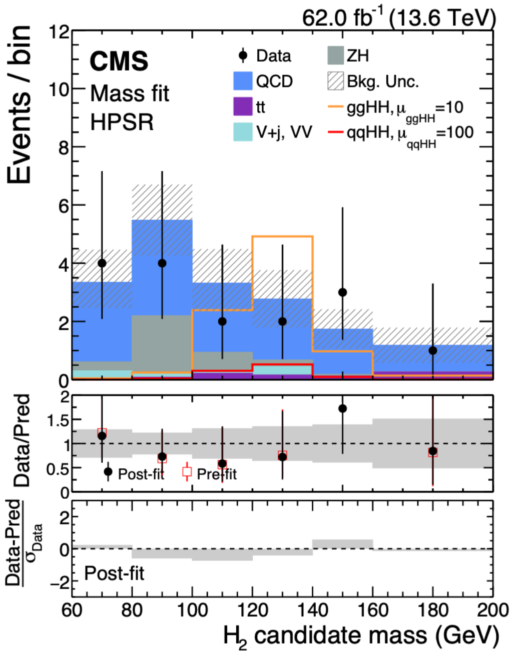


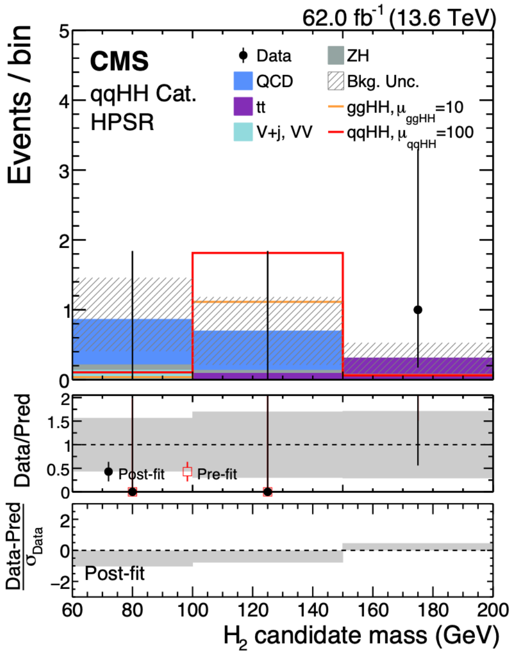
(b) Particle Attention Block

$$\text{P-MHA}(Q, K, V) = \mathrm{SoftMax}(QK^T/\sqrt{d_k} + \mathbf{U})V,$$

Improved results on $HH \rightarrow 4b$ | Signal Extraction

Region Name	$\mathcal{P}_{HH}$ Selection	$T_{xb\bar{b}}$ Selection
HPSR	$\mathcal{P}_{HH} > 0.9$	$T_{xb\bar{b}}(H_2) > 0.95$
LPSR	$\mathcal{P}_{HH} > 0.9$	$0.95 > T_{xb\bar{b}}(H_2) > 0.85$
VR (Veto events in the HPSR and LPSR)	$\mathcal{P}_{HH} > 0.7$	$T_{xb\bar{b}}(H_2) > 0.7$





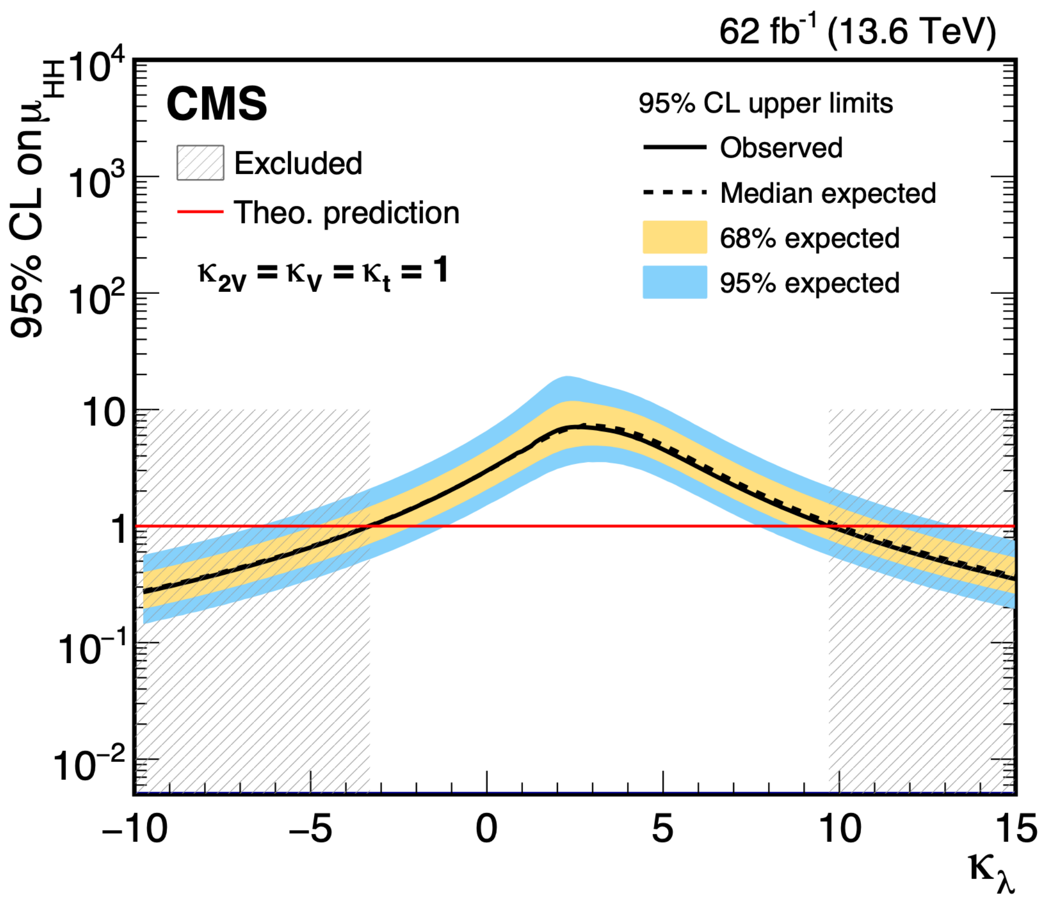
Two categories defined, one for ggHH and one for $qqHH$, based on the presence of additional AK4 jets.

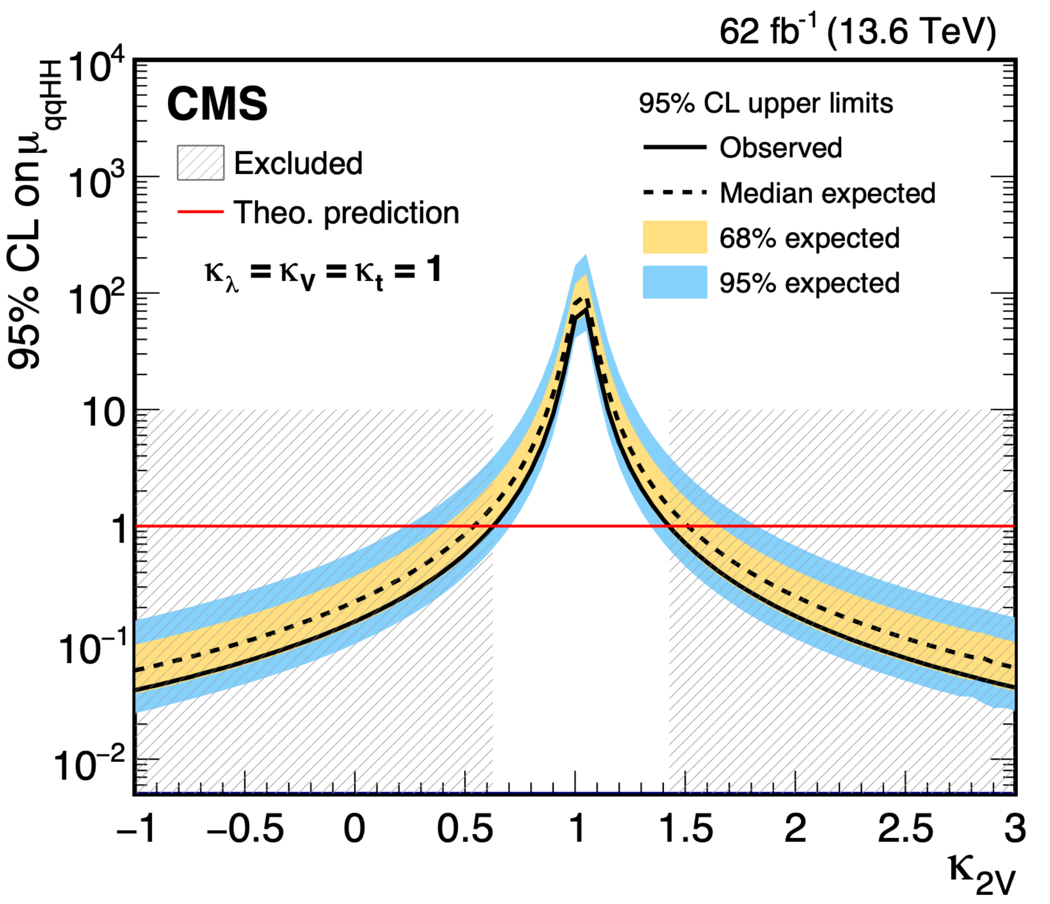
DNN trained to separate signal from backgrounds defining $\mathcal{P}_{HH}$ using:

- p_T and η of the Higgs candidate.
- $T_{xb\bar{b}}(H_1)$ and $m_{\text{reg}}(H_1)$.
- (p_T, η, m_{HH}) of the HH system.
- ΔR and mass of each Higgs candidate and nearest AK4 jet.

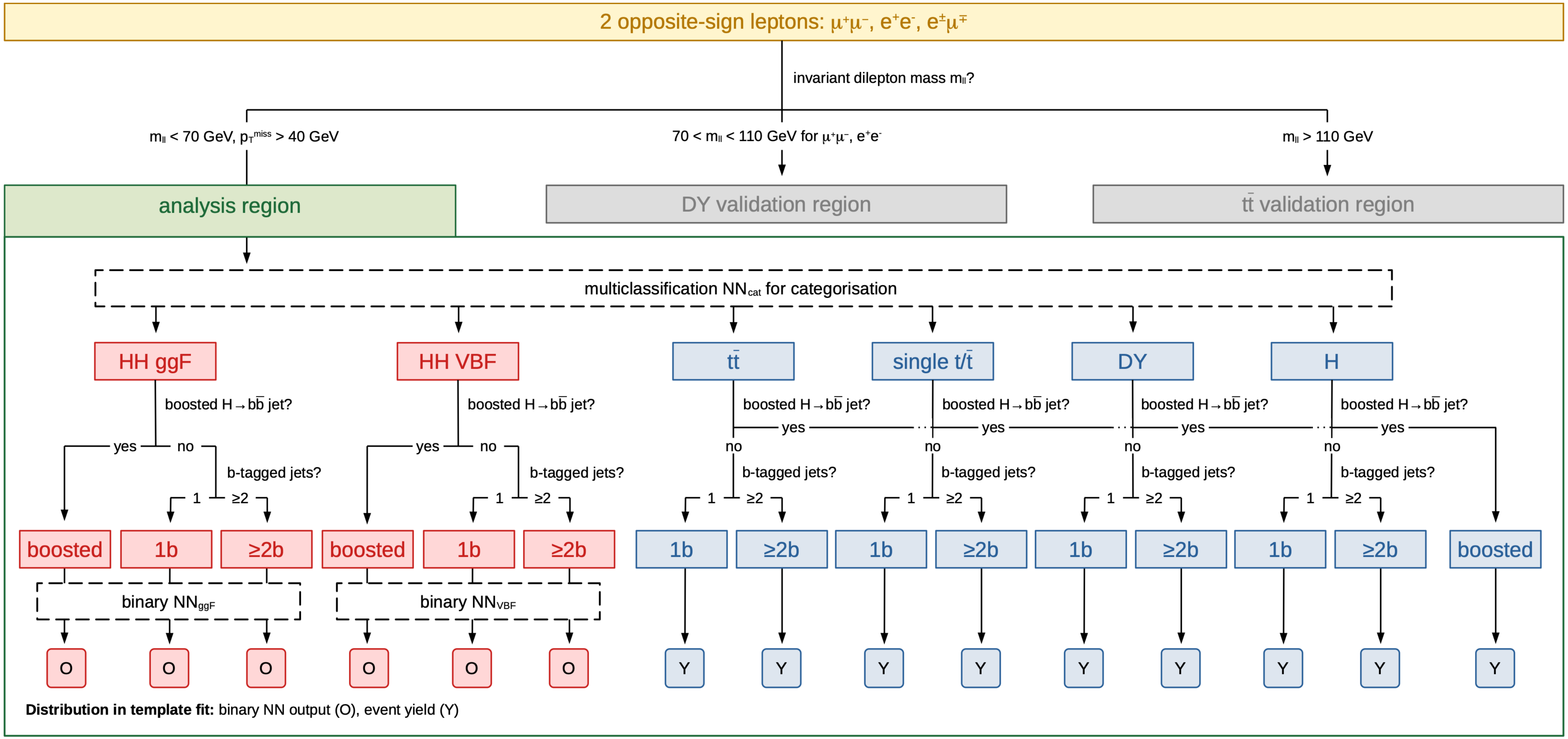
Distribution of $m_{\text{reg}}(H_2)$ used as the fitted variable.

Improved results on $\mathrm{HH} \rightarrow 4\mathrm{b}$ | Results

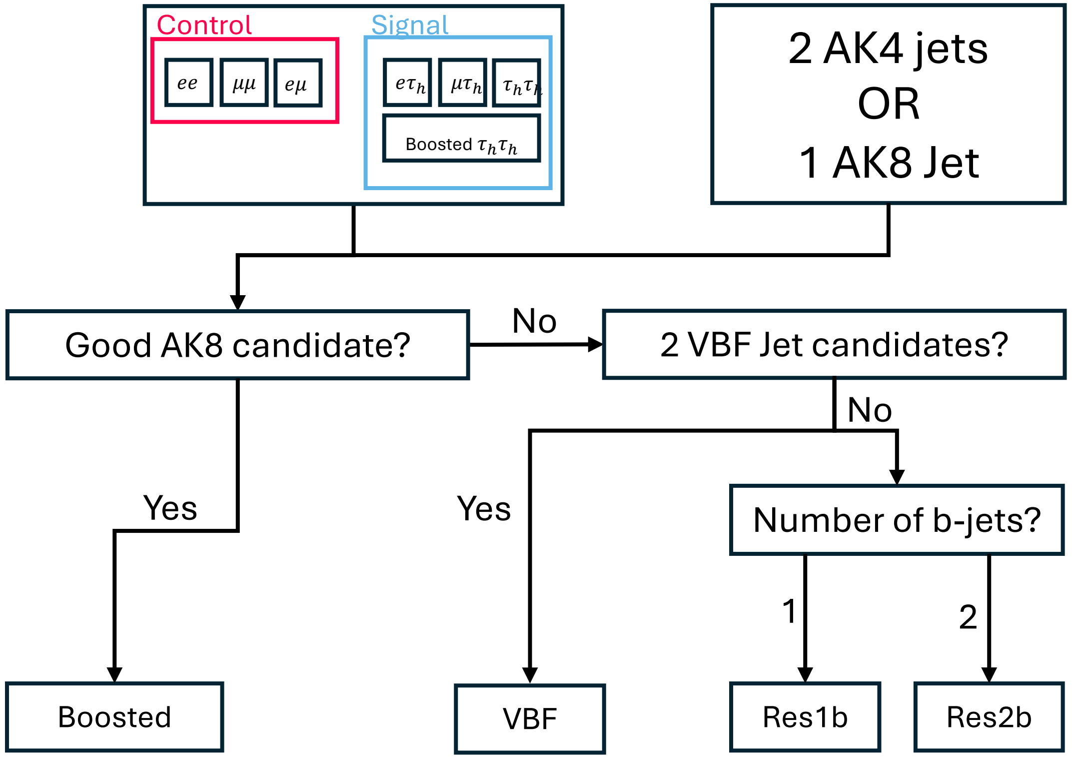




Event classification for $HH \rightarrow b\bar{b}WW$



Event Classification $\mathrm{HH} \rightarrow \mathrm{b\bar{b}\tau\tau}$



Improved τ_h triggers in 2024