Search for non-resonant $HH \to b\bar{b}\gamma\gamma$ with Run3 data

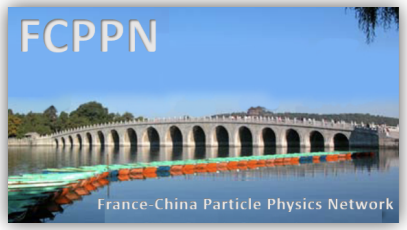
17^{th} Workshop of the FCPPN/L

On behalf of the CMS collaboration

Éva Didier-Estève

Institut de Physique des 2 Infinis
Université Claude Bernard - Lyon1

June 29^{th}, 2026



Introduction & motivations

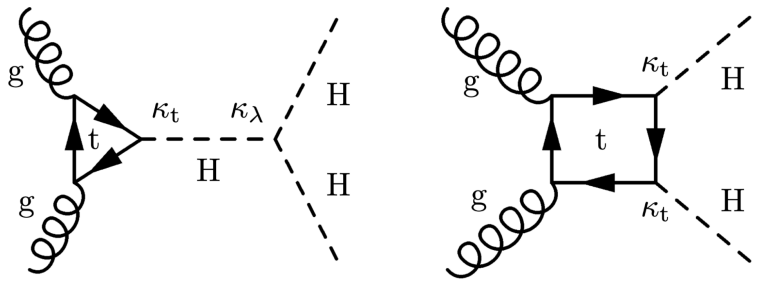
Higgs potential

$$V(\phi) \approx \lambda(\phi^2 + \nu^2)^2$$

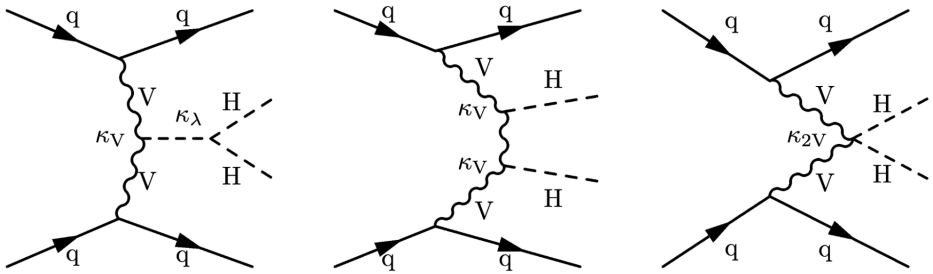
- Higgs **self-coupling λ: access via Higgs pair production** *HH*
- *HH* at LHC: dominant mode is ggF, followed by VBF
- $HH \to b\bar{b}\gamma\gamma$: large $\mathcal{B}(b)$, good mass resolution of $\gamma\gamma$

	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%

$\sigma(HH \to XXYY)$



κ_λ via ggF



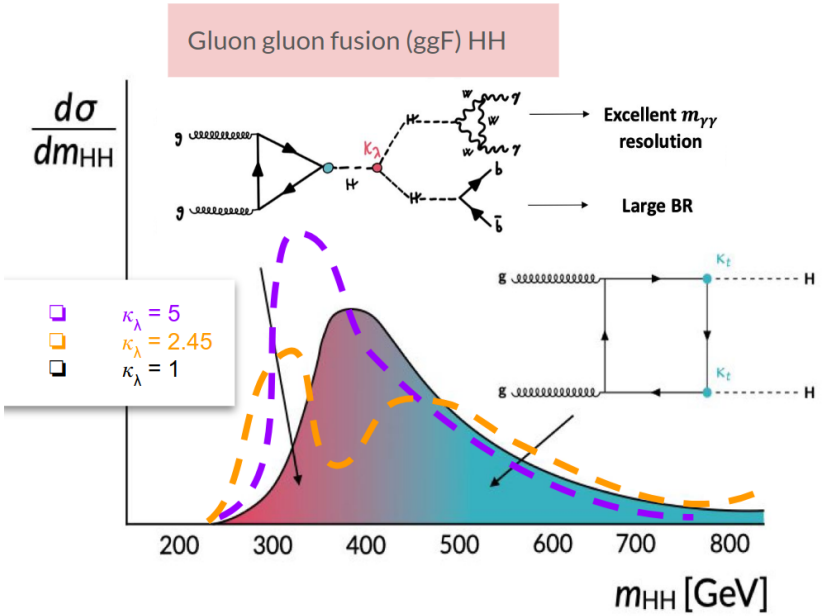
κ_λ via VBF

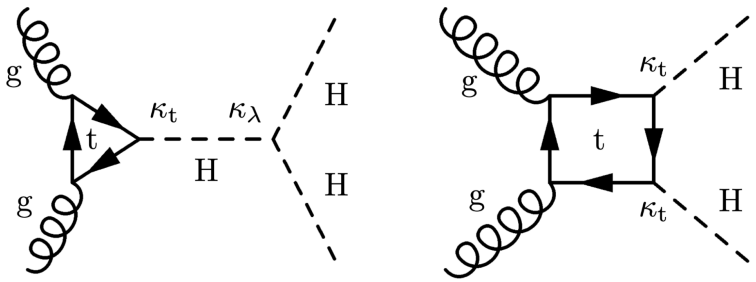
Introduction & motivations

Objective

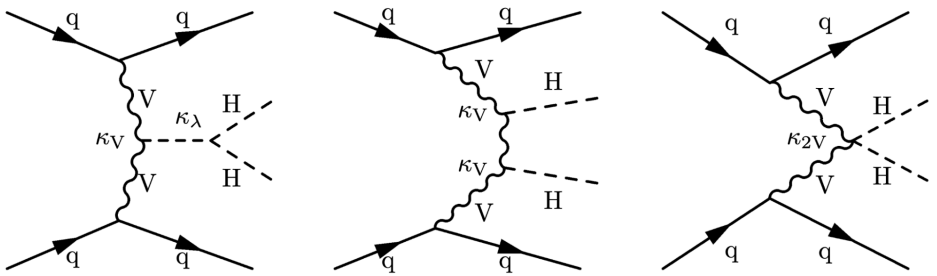
Constrain Higgs self-coupling modifier $\kappa_\lambda = \lambda/\lambda^{SM}$

- Partial Run3 (2022-2023): channel $pp \to HH \to b\bar{b}\gamma\gamma$: 61.9fb^{-1}
- 2 classes of bkg: resonant ($H \to \gamma\gamma$) and nonres ($\gamma\gamma$ not from H).



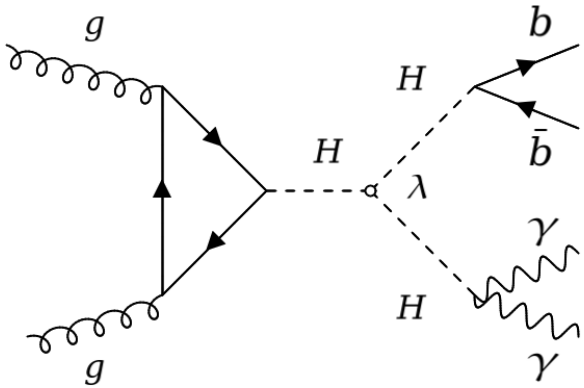


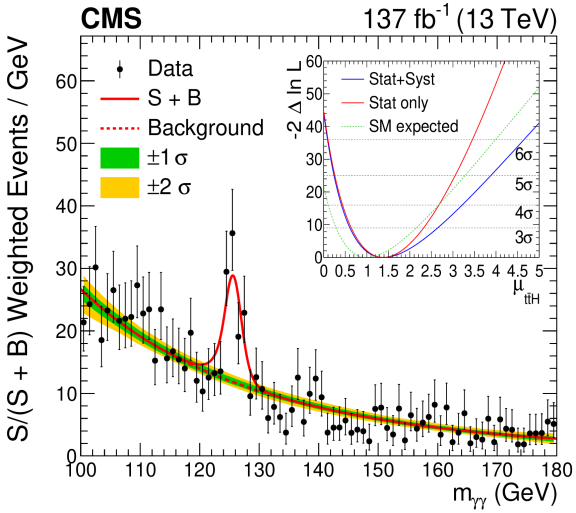
κ_λ via ggF



κ_λ via VBF

Measuring $HH \to b\bar{b}\gamma\gamma$





Example of a **single Higgs** S+B peak.

From (HIG-19-013)

- Very small $\sigma_{pp\to HH}$ (1000 $\times$ smaller than single H)
- Small expected signal, large bkg contribution.

Strategy

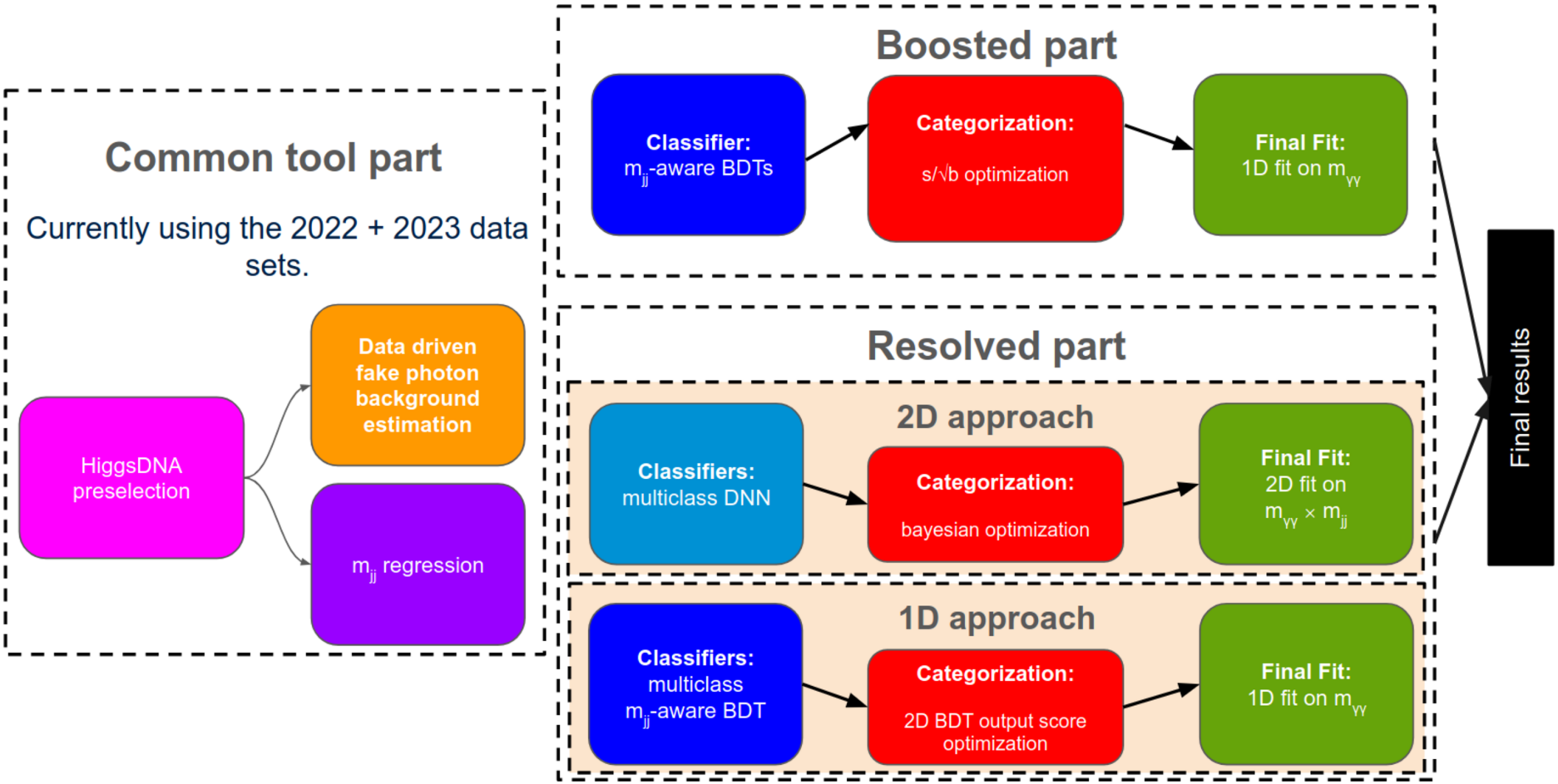
Signal extraction based on $m_{\gamma\gamma}$ shape fit:

Search for **narrow diphoton peak**

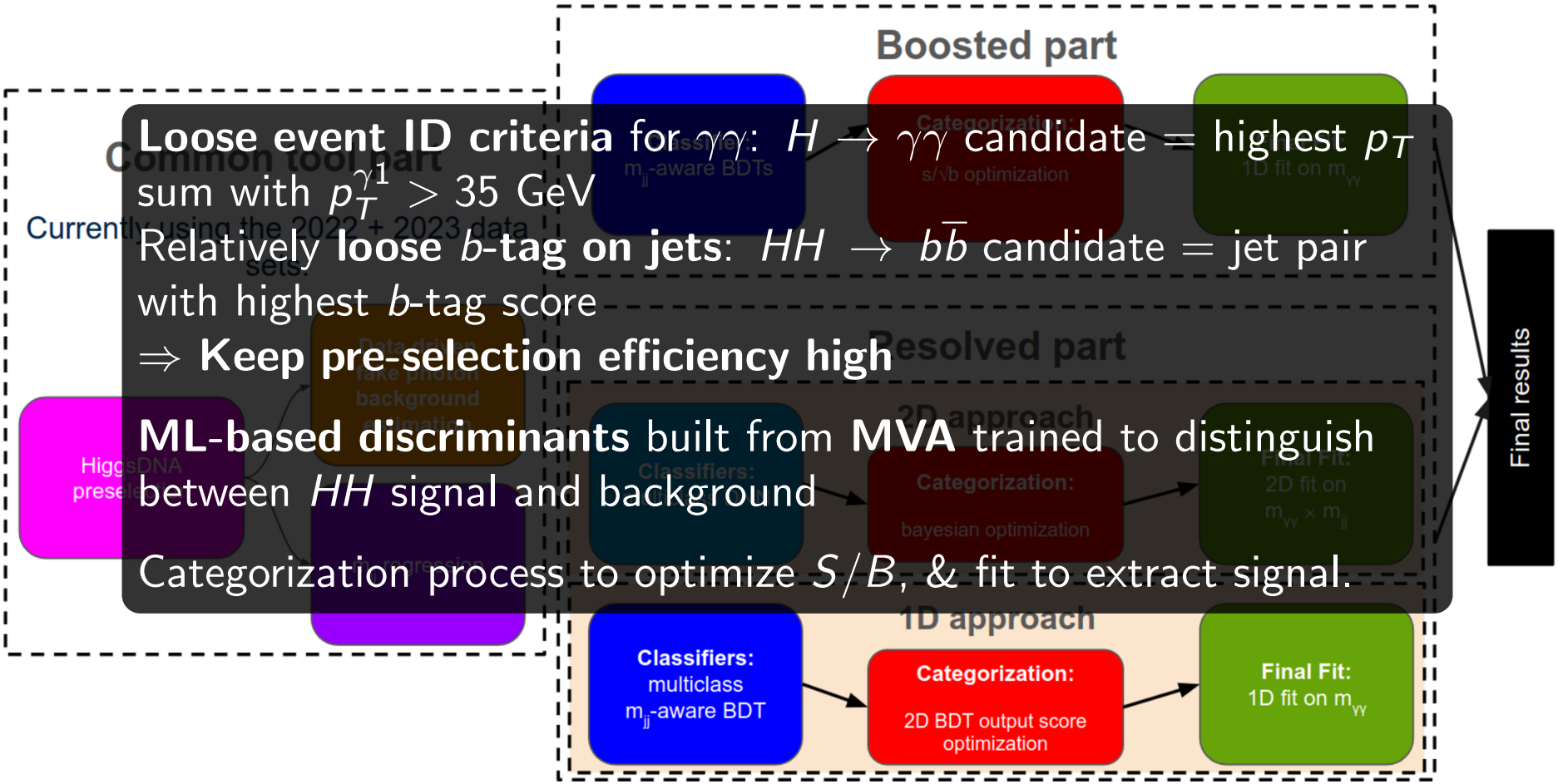
Smoothly decreasing background

Use of **MultiVariate Analysis** (MVA) to select $H \to \gamma\gamma$ & $H \to b\bar{b}$

Analysis Workflow



Analysis Workflow



Loose event ID criteria for $\gamma\gamma$: $H \to \gamma\gamma$ candidate = highest p_T sum with $p_T^{\gamma 1} > 35$ GeV

Relatively **loose** b**-tag on jets**: $HH \to b\bar{b}$ candidate = jet pair with highest b-tag score

$\Rightarrow$ **Keep pre-selection efficiency high**

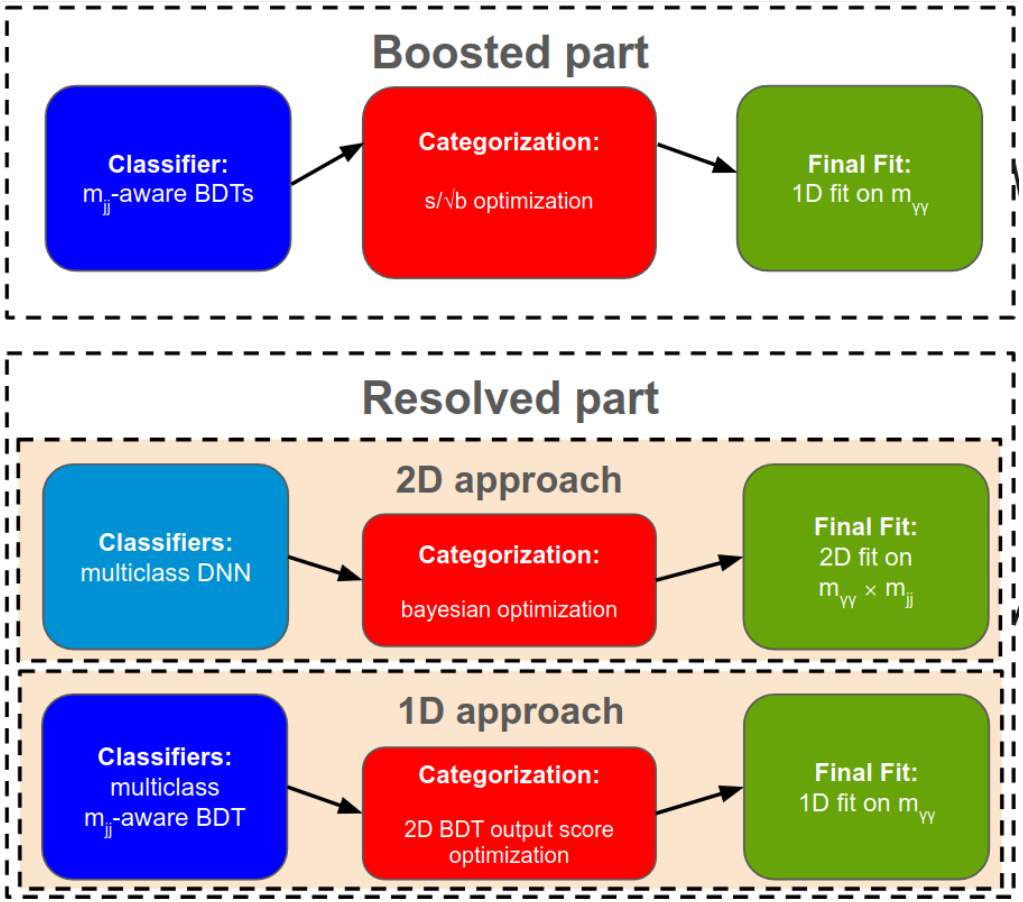
ML-based discriminants built from **MVA** trained to distinguish between HH signal and background

Categorization process to optimize S/B, & fit to extract signal.

Analysis Strategy

- After pre-selection: train **ML-based discriminant** to separate signal & background.
- Two approaches:
  1D fit ($m_{\gamma\gamma}$);
  2D ($m_{\gamma\gamma}$, m_{jj})
- 3 categories for each + new "boosted" category

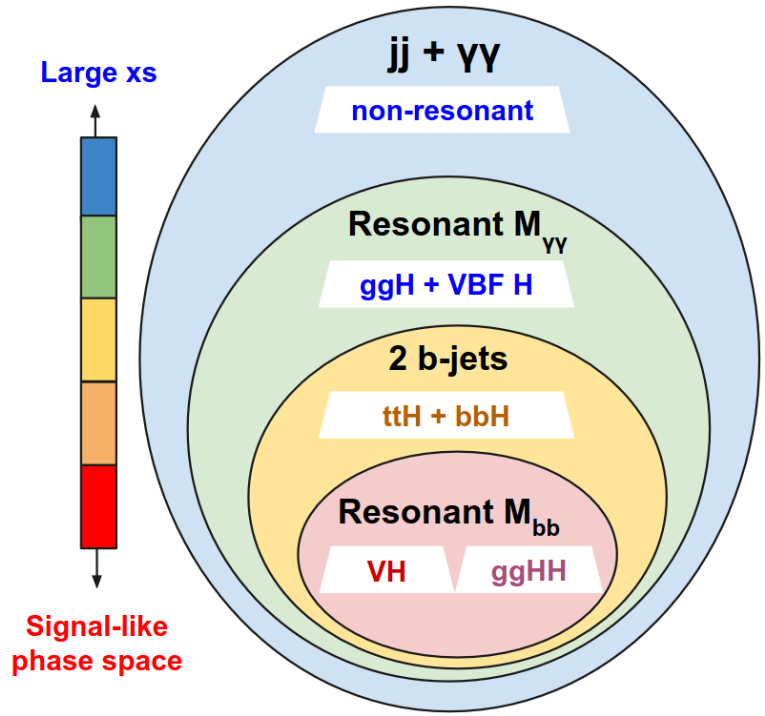
$\Rightarrow$ 2 approaches are **complementary**



1D approach

Multiclass BDT with m_{jj} as input
Signal extracted via 1D-fit on $m_{\gamma\gamma}$

- Trained with several kinematic variables
- BDT trained to optimize S/B for 1D fit over $m_{\gamma\gamma}$: suppression of backgrounds resonant in $m_{\gamma\gamma}$
- BDT trained with 4 classes:
  - **ggF HH (signal)**
  - VH
  - $t\bar{t}H + b\bar{b}H$
  - QCD (non-res + ggH + VBF H)



1D approach: fitting

Discriminators from BDT:

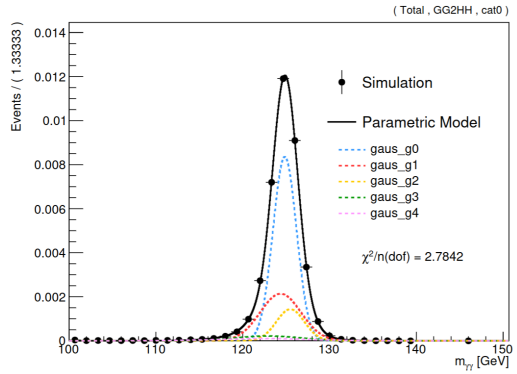
$$D_{sig\_vs\_ttH} = \frac{P_{ggF\ HH}}{P_{ggF\ HH} + P_{ttH}}$$

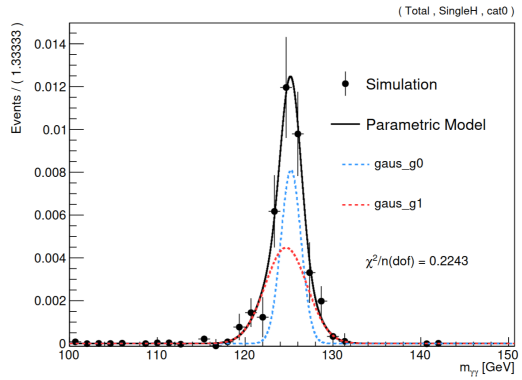
$$D_{sig\_vs\_nonres} = \frac{P_{ggF\ HH}}{P_{ggF\ HH} + P_{VH} + P_{QCD}}$$

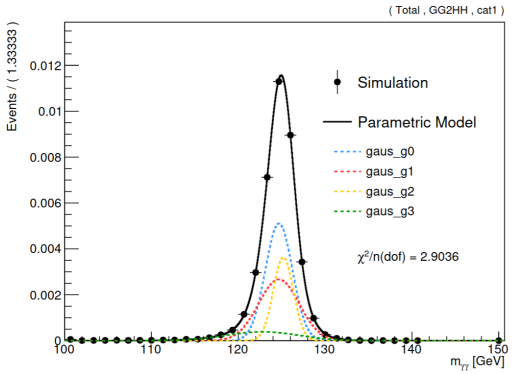
Used to build 3 categories:

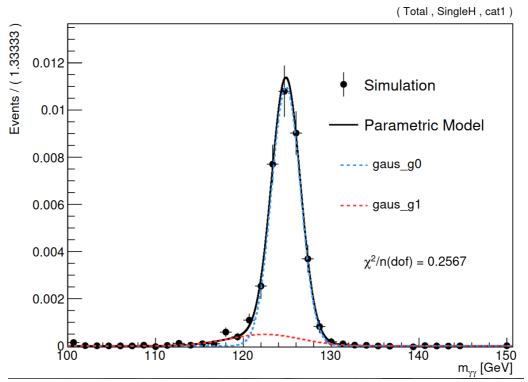
- ⇒ Cut on $D_{sig\_vs\_ttH}$ to get 85% ttH rejection
- ⇒ Scan $D_{sig\_vs\_nonres}$ to get best S/B

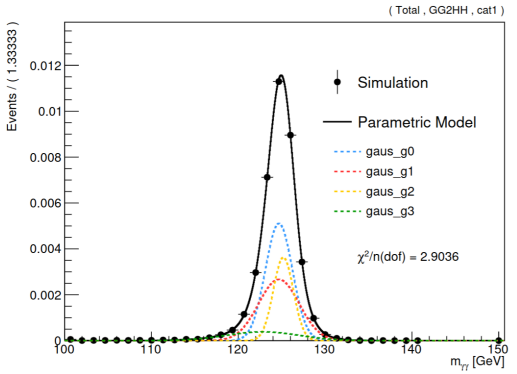
Use of FinalFit to extract signal.

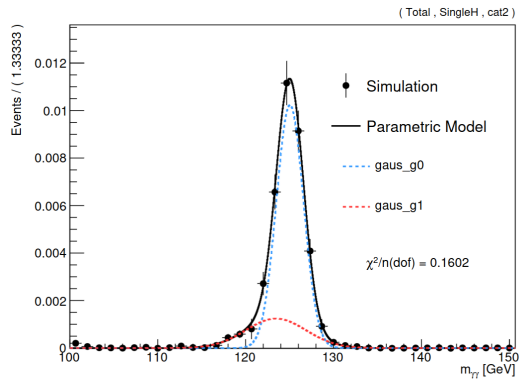












$m_{\gamma\gamma}$ fit for "signal"

$m_{\gamma\gamma}$ fit for "resonant"

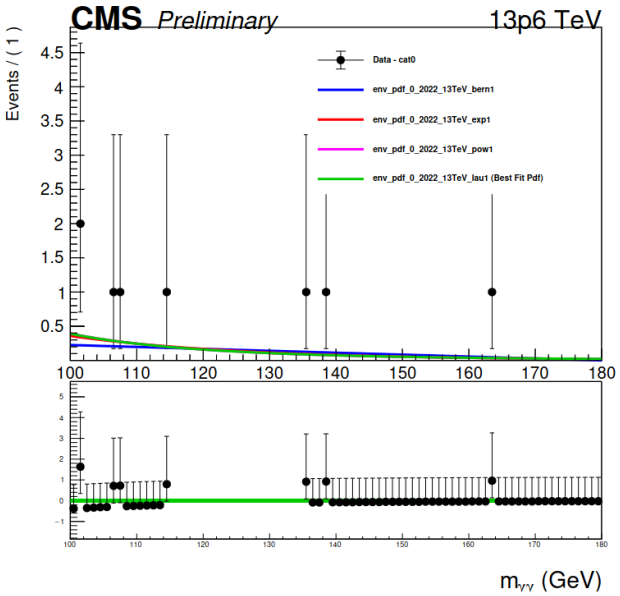
1D approach: fitting

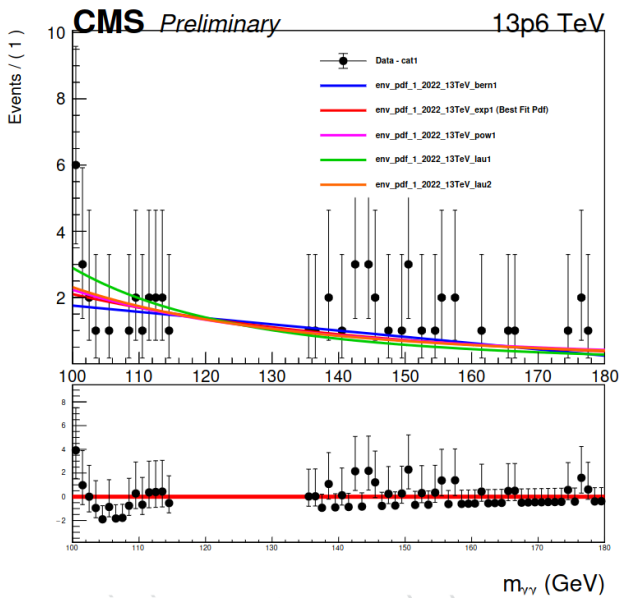
Discriminators from BDT:

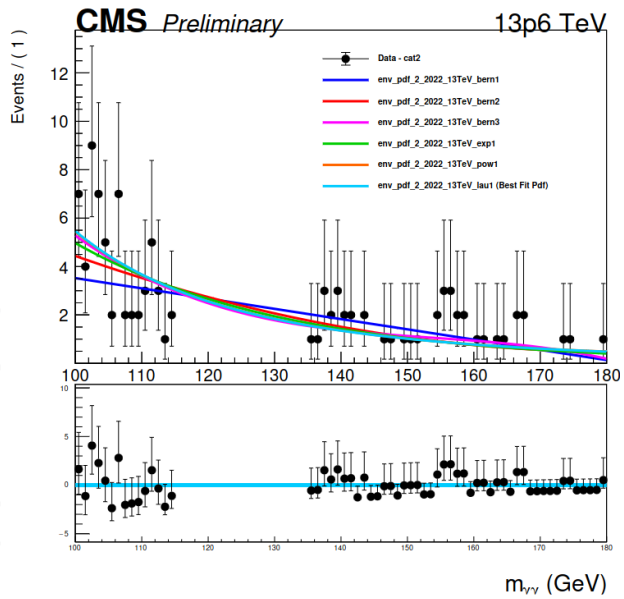
$$D_{sig\_vs\_ttH} = \frac{P_{ggF\ HH}}{P_{ggF\ HH} + P_{ttH}}$$

$$D_{sig\_vs\_nonres} = \frac{P_{ggF\ HH}}{P_{ggF\ HH} + P_{VH} + P_{QCD}}$$

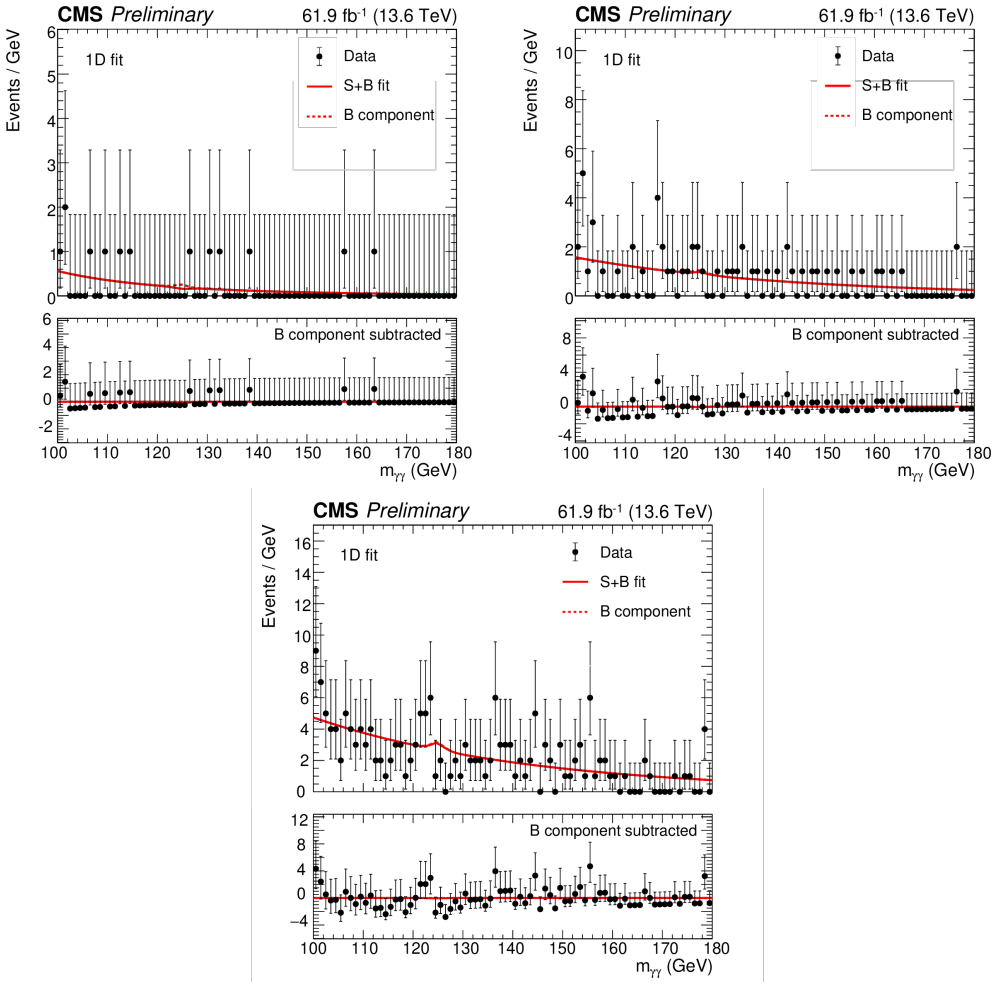
Sideband F-test fits in ($m_{\gamma\gamma}$) distribution in "data", for non-resonant background: the signal region $(115, 135)$ GeV has been **blinded**







1D approach: results



1D strategy $S + B$ fits in $m_{\gamma\gamma}$

- Category 1, 2, and 3 fits in $m_{\gamma\gamma}$: $S + B$ fit vs. data.
- Expected 95% CL upper limits are calculated using asymptotic approximation.

$\Rightarrow$ All categories yield **observed 95% CL upper limit of 7.4 times the SM cross section.**

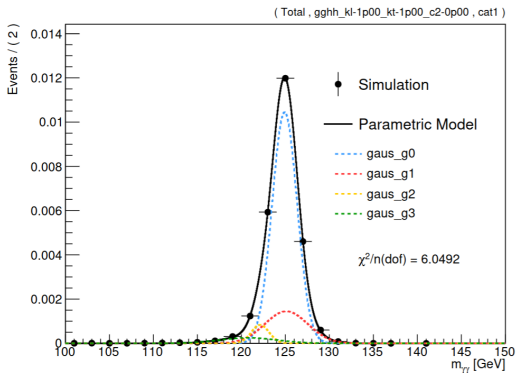
2D approach

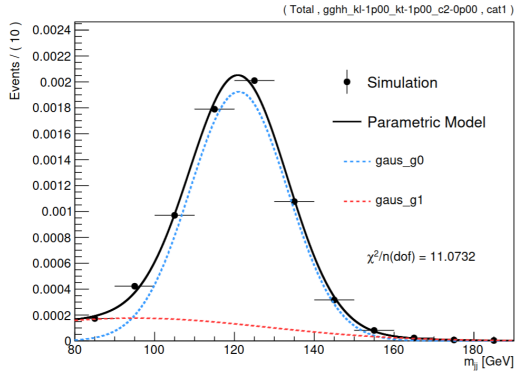
2D analysis strategy

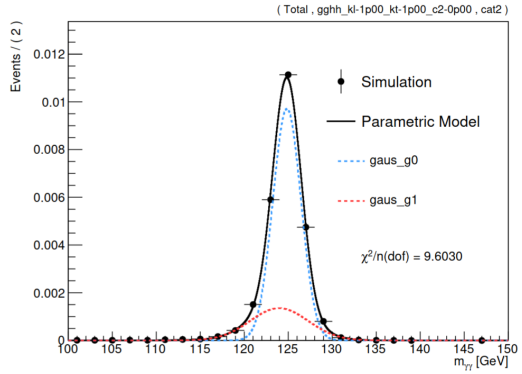
Multiclass DNN to suppress bkg
m_{jj} **not** a DNN variable:
Signal extracted via 2D-fit on ($m_{\gamma\gamma}$, m_{jj})

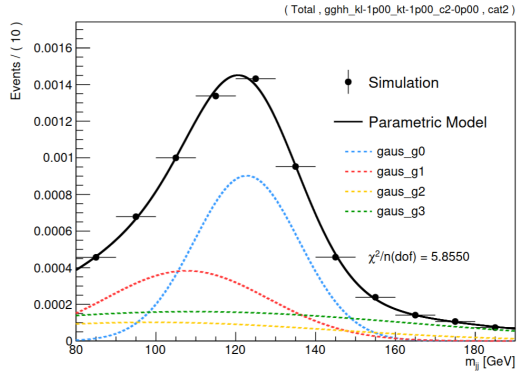
DNN trained to suppress all bkgs.
4 classes:

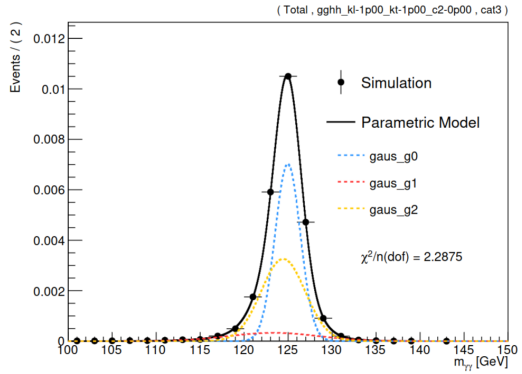
- ggF *HH*
- ttH
- Other single *H* bkgs (*ggH*, bbH, VBF *H*, VH)
- Non-resonant (QCD + γ+ j; $\gamma\gamma$ + tt/j)

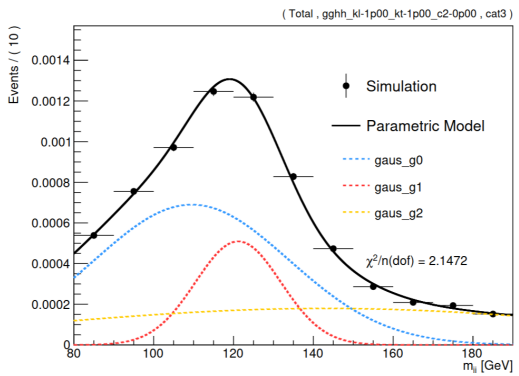












$m_{\gamma\gamma}$ fit for "signal"

m_{jj} fit for "signal"

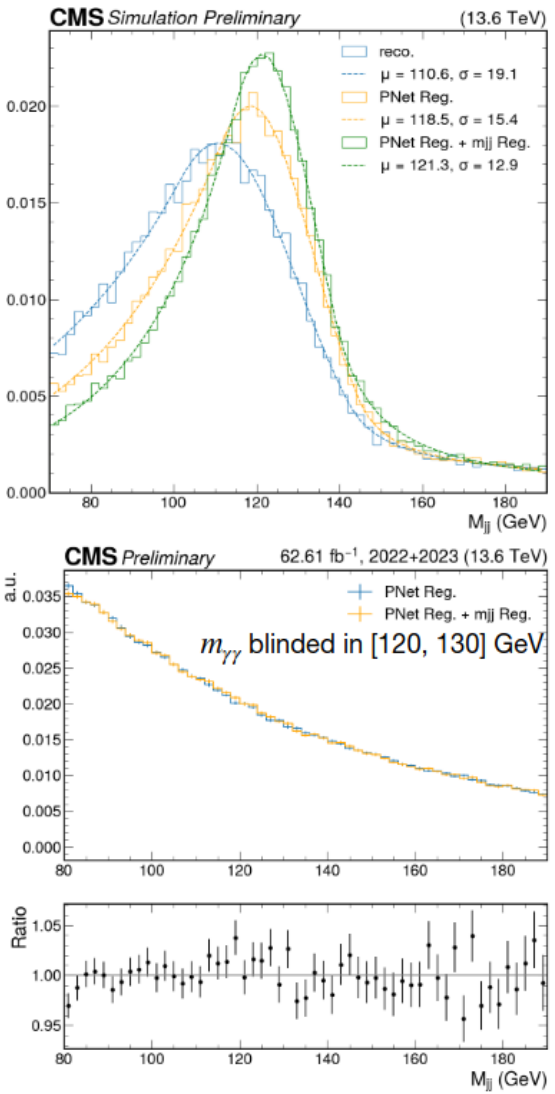
Dijet mass regression

m_{jj} regressor

DNN-based regression to further improve jet mass resolution m_{jj} of b-jet candidate.

- Variable is a **correction term**, on top of Particle Transformer for Jet Tagging b-jet P_T regression: $(m_{jj}^{gen} - m_{jj}^{reco})/m_{jj}^{reco}$
- Training set from 2022/2023, using $ggHH \to b\bar{b}\gamma\gamma$ MC sample

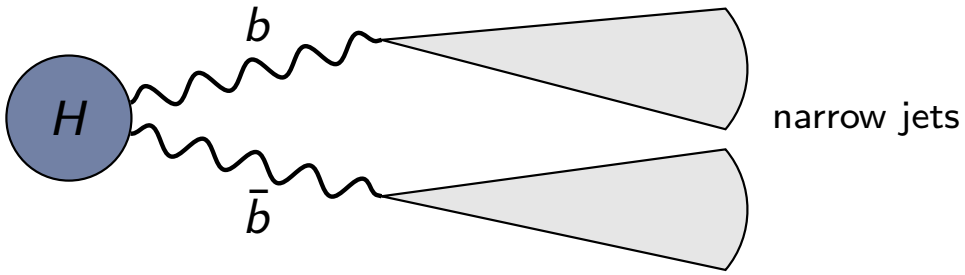
$\Rightarrow$ **10-20% improvement observed in $m_{H\to b\bar{b}}$!**



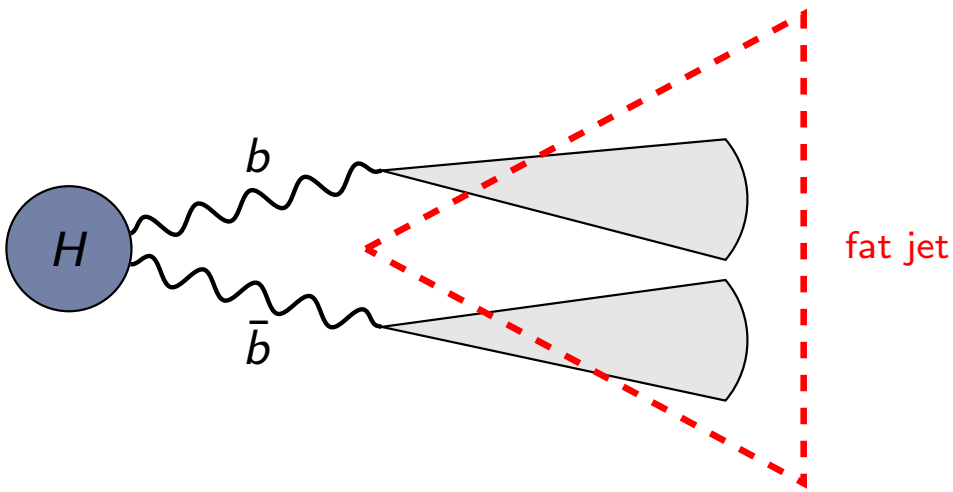
Boosted Higgs

The "boosted" category

Novelty: boosted Higgs-tagging to improve significance:
allows to exploit high-p_T $H \to b\bar{b}$

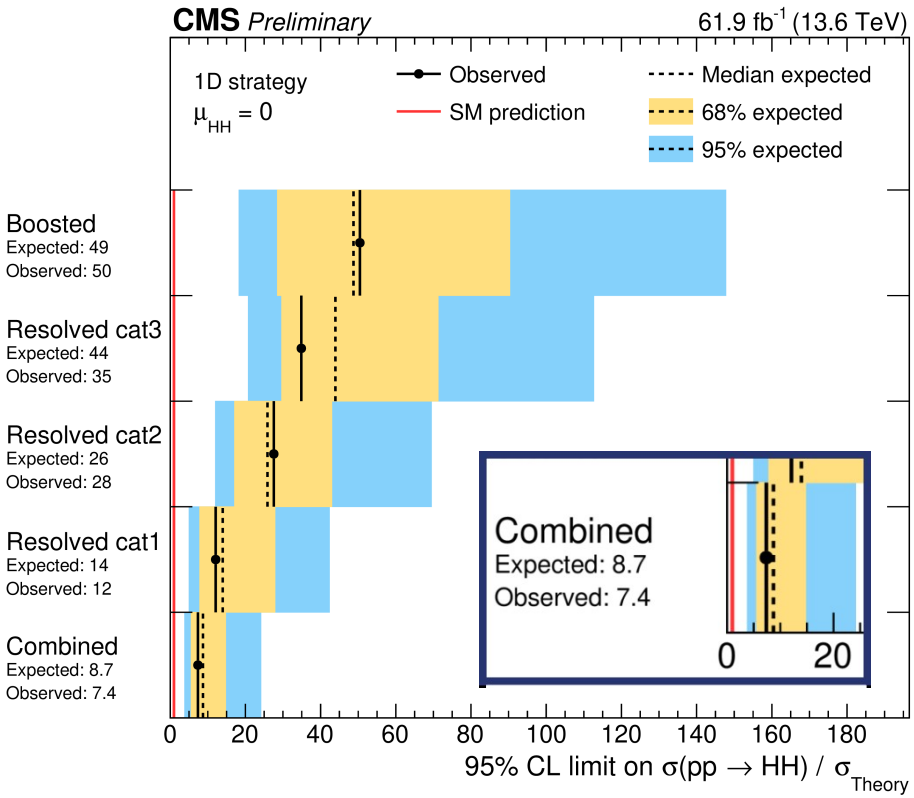


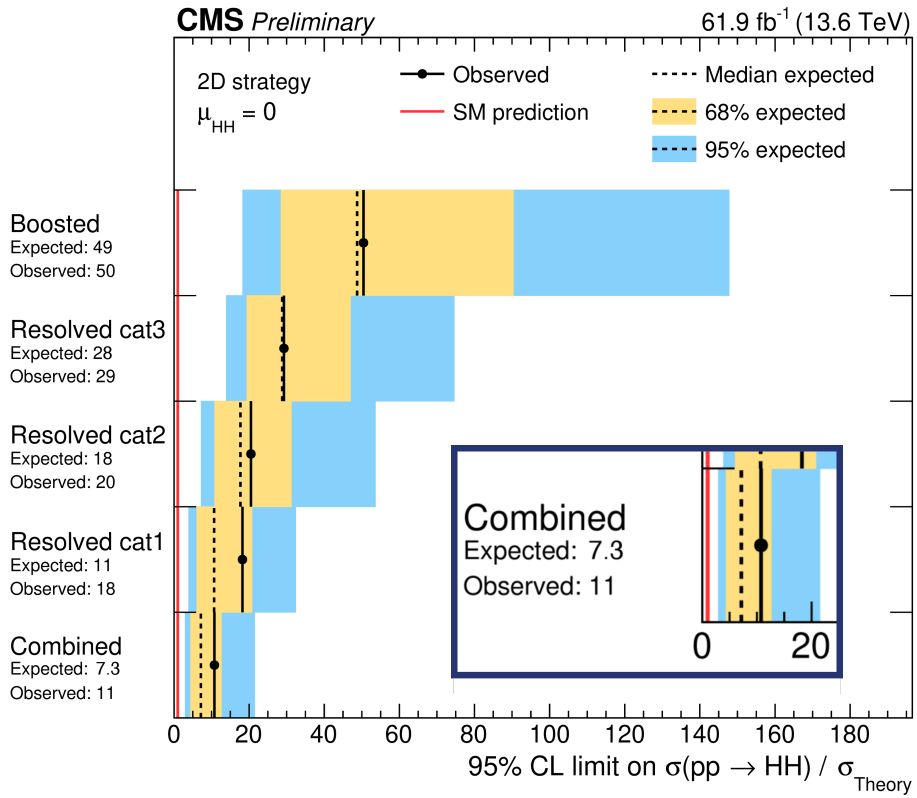
- Jet merging due to boost along $\vec{z}$: exploit fat jets, & final selection with BDT discriminants.
- Expected bkg is **low**: events in "boosted" are prioritized & removed fromp non-boosted categories.
- Low signal + bkg: perform 1D-fit in $m_{\gamma\gamma}$ to extract signal.



Results: cross-section

Resolved (1D, 2D) combined with "boosted": boosted impact $\sim 5\%$

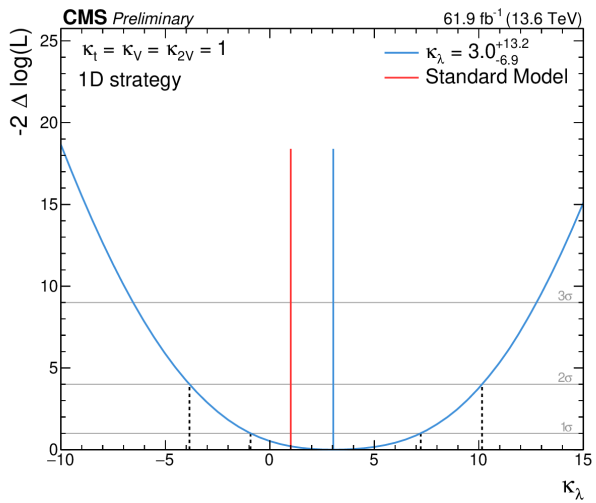


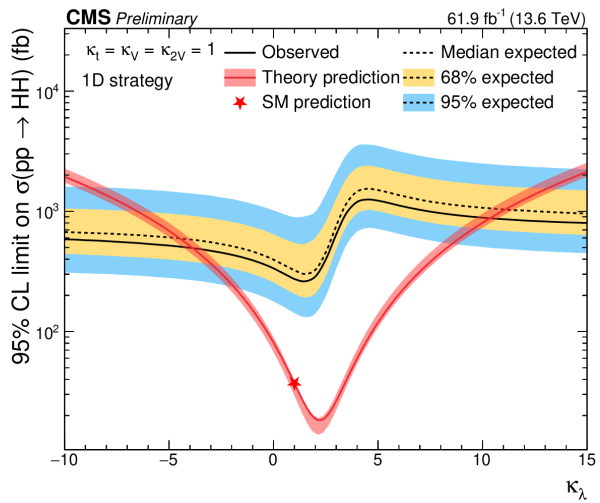


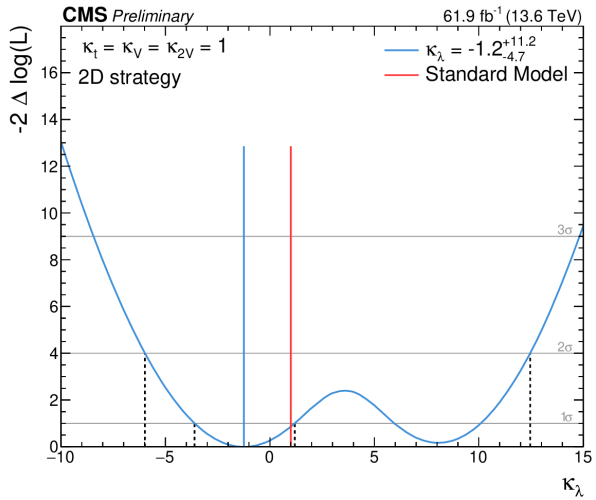
1D

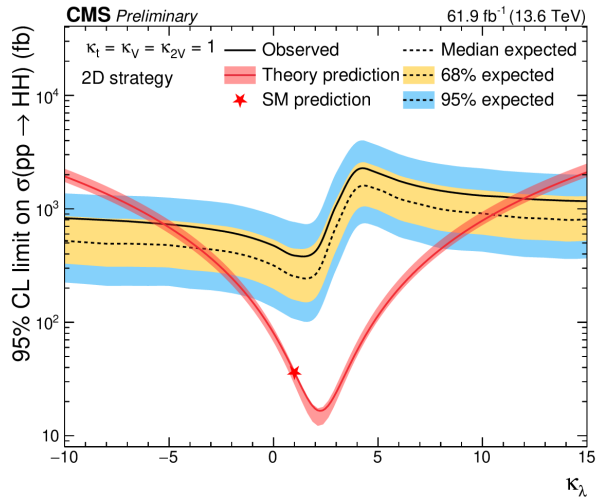
2D

Results: κ_λ scans









1D fit:

$\kappa_\lambda \in [-3.9, 10.1]$ (observed)

$\kappa_\lambda \in [-4.5, 11.1]$ (expected)

2D fit:

$\kappa_\lambda \in [-6.1, 12.8]$ (observed)

$\kappa_\lambda \in [-3.6, 10.5]$ (expected)

Summary & next steps

Nonresonant Run3 (2022+2023) search for ggF $HH \to b\bar{b}\gamma\gamma$.
1D approach and 2D approach are compatible:

- 1D: Observed (expected) 95% CL upper limit of 7.4 (8.7) $\times \sigma_{SM}$ and allowed κ_λ range [-3.9,10.1] ([-4.5,11.1]);
- 2D: Observed (expected) 95% CL upper limit of 11.0 (7.3) $\times \sigma_{SM}$ and allowed κ_λ [-6.1,12.8] ([-3.6,10.5])

$\Rightarrow$ Published in `CMS PAS HIG-25-007`

Summary & next steps

Nonresonant Run3 (2022+2023) search for ggF $HH \to b\bar{b}\gamma\gamma$.
1D approach and 2D approach are compatible:

- 1D: Observed (expected) 95% CL upper limit of 7.4 (8.7) $\times \sigma_{SM}$ and allowed κ_λ range [-3.9,10.1] ([-4.5,11.1]);
- 2D: Observed (expected) 95% CL upper limit of 11.0 (7.3) $\times \sigma_{SM}$ and allowed κ_λ [-6.1,12.8] ([-3.6,10.5])

$\Rightarrow$ Published in `CMS PAS HIG-25-007`

Next:

Work on: Run2 & rest of Run3
Work on EFT interpretation
$\Rightarrow$ Measure of κ_λ is one of the major goals of HL-LHC.

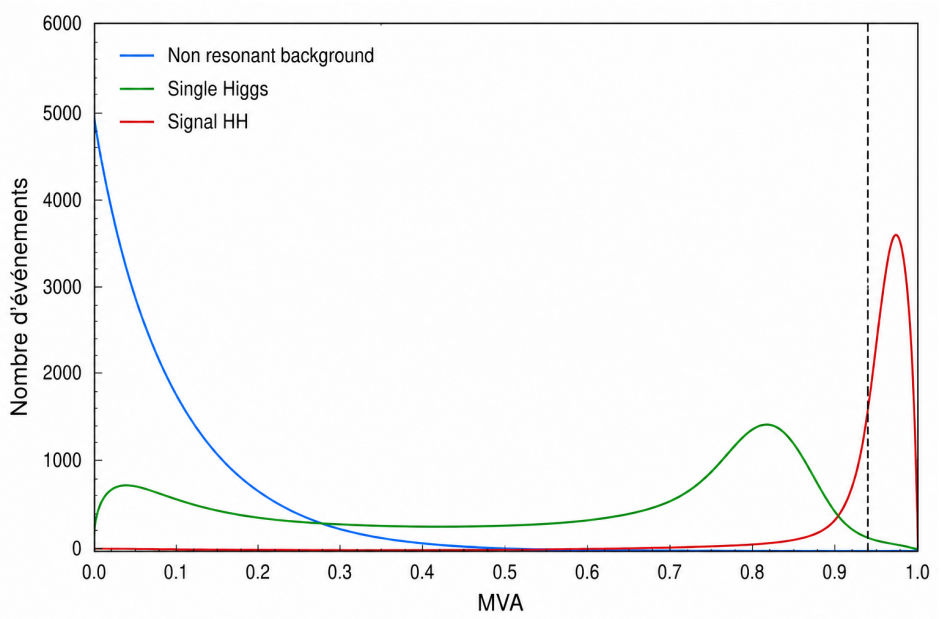
Thank you for listening!

Backup

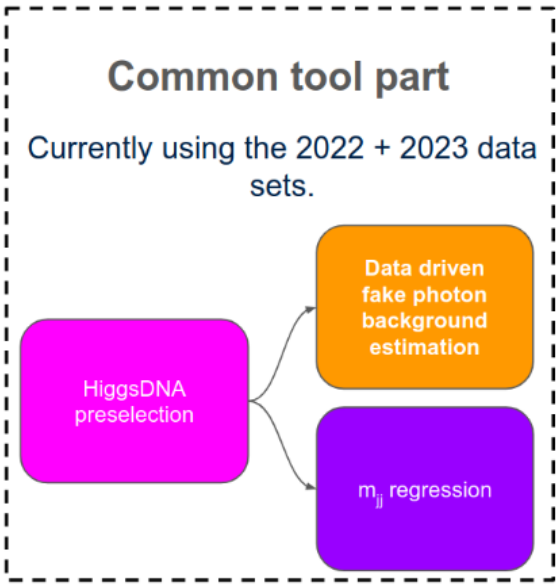
Summary of systematic uncertainties in the signal and background yields

Uncertainty source	uncertainty Size	1D fit	2D fit	boosted
b tagging	6–9%	norm	norm	norm
Trigger efficiency	1–2%	norm	norm	norm
Preselection efficiency	1–2%	norm	norm	norm
Electron veto efficiency	1–2%	norm	norm	norm
Pileup	1–4%	norm	norm	norm
Luminosity	1.4%	norm	norm	norm
Photon energy scale	$< 10\%$	shape	shape	shape
Photon energy resolution	$< 10\%$	shape	shape	shape
Jet energy scale	4–9%	norm	shape	norm
Jet energy resolution	1–4%	norm	shape	norm
Higgs boson branching ratio	2 %	norm	norm	norm
ggH, VH, VBF H cross section	1–4%	norm	norm	norm
$t\bar{t}H$ cross section	9%	norm	norm	norm
PDF	1–3%	norm	norm	norm
α_S	1–3%	norm	norm	norm
HH theory cross section	23%	norm	norm	norm
bbH modeling	50%	norm	norm	norm
Bias uncertainty	3–30%	–	norm	–
MC statistics	1–20%	norm	norm	norm

MVA representation



Backup: pre-selection



Selections, corrections, & systematics based on HiggsDNA.

- Photons: default $H \to \gamma\gamma$ preselection

  $p_T > 25$ GeV
  Within barrel and endcap acceptance
  $H/E < 0.08$
  γ MVA ID > -0.9
  Selection to mimic online selection

  $\Rightarrow$ $H \to \gamma\gamma$ candidate: highest P_T sum with $p_T^{\gamma 1} > 35$ GeV

- Loose jet preselection:

  AK4, PUPPI, ParticleNet corrected jets with $p_T(j) > 20$ GeV, $|\eta| < 2.5$, $\Delta R(\gamma, j) > 0.4$

  $\Rightarrow$ $H \to b\bar{b}$ candidate: jet pair with highest ParticleNet b-tag score

Fake photon background

Data-driven method for fake γ

Struggle to estimate non-prompt γ from MC samples: issues in yields & shape agreements.
$\Rightarrow$ **Estimate from data instead**

Adapted from original method (HIG-19-013):

- Build sidebands from data, with low γ-ID MVA score
- Build PDF from this sideband
- Use it to **re-weight data events** in the pre-selection region.