



# Why two Higgs are better than one?

or...

Search for **Higgs boson pair production** in the two bottom quarks plus two photons final state in  $p p$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector

arXiv:2112.11876

Project of group E: Searches at the LHC.

C-CRET Collaboration (CERN Commission of Recreational Experiments outside)

CERN School of High Energy Physics, Chile, March 2023



# Di-Higgs Searches · Motivations

- Study the Higgs potential  $\longrightarrow$  **still untested!**

$$V(H) = -\mu^2 |H|^2 + \lambda |H|^4$$

EWBSB  $\hookrightarrow \lambda v h^3$

- Origin of the **EW scale** (Hierarchy Problem)
- **Baryogenesis**  $\rightarrow$  1<sup>st</sup> order phase transition
- **Metastability** of the Higgs vacuum
- Search for **new massive scalar** (Composite Higgs, 2HDM, SUSY)

+ . . .

Beyond Standard Model (BSM)

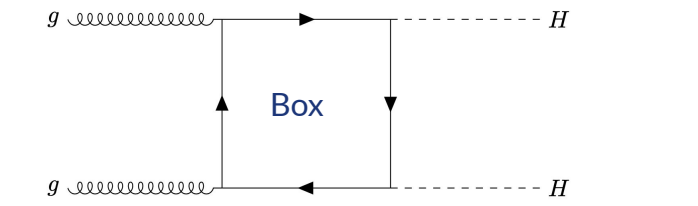
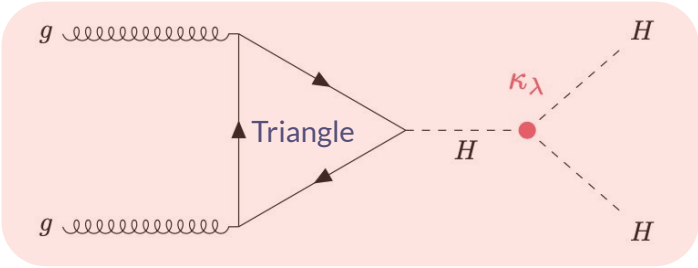
# Di-Higgs Searches · Signal ( $pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ )

## 1. Non-resonant search

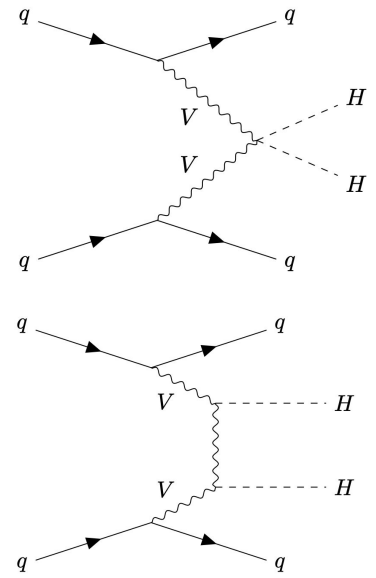
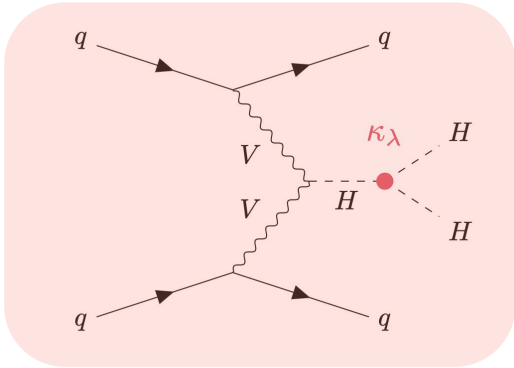
DEF

$$\kappa_\lambda = \lambda_{HHH} / \lambda_{HHH}^{\text{SM}}$$

ggF



VBF



# Di-Higgs Searches · Signal ( $pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ )

## 1. Non-resonant search

Search for BSM where the SM is suppressed

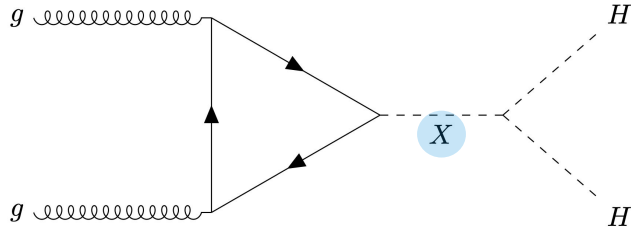
Destructive Interference

$$\mathcal{M}_{\text{triangle}} \sim +\kappa\lambda \log^2 \frac{m_{hh}^2}{m_t^2}$$

$$\mathcal{M}_{\text{box}} \sim -\log^2 \frac{m_{hh}^2}{m_t^2}$$

Off-shell Higgs: higher sensitivity to UV physics

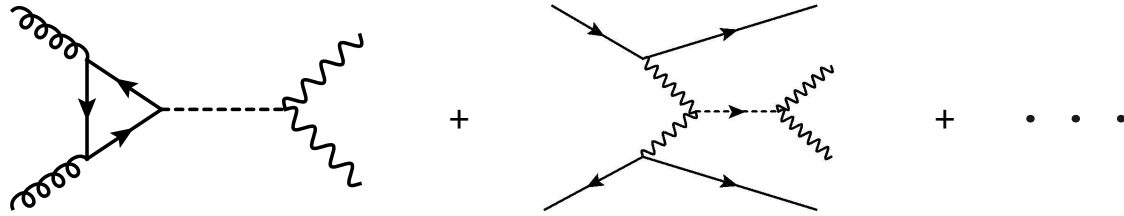
## 2. Resonant search



$$m_X > 2m_h$$

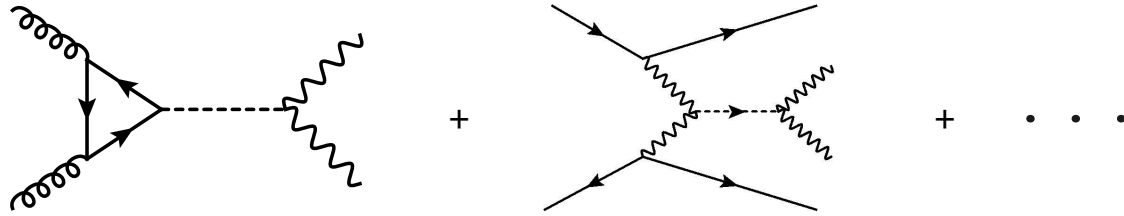
# Di-Higgs Searches · Background ( $pp \rightarrow b\bar{b}\gamma\gamma$ )

Single Higgs: ggF, VBF, Wh, Zh, tth, Zh ... (Monte-Carlo methods)

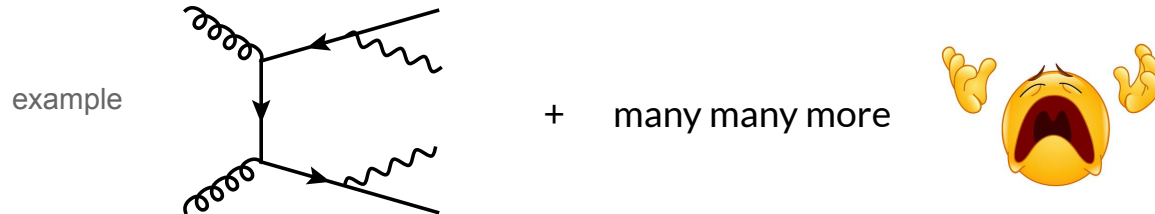


# Di-Higgs Searches · Background ( $pp \rightarrow b\bar{b}\gamma\gamma$ )

Single Higgs: ggF, VBF, Wh, Zh, tth, Zh ... (Monte-Carlo methods)



Others:  $\gamma\gamma$  + jets (data-driven method)

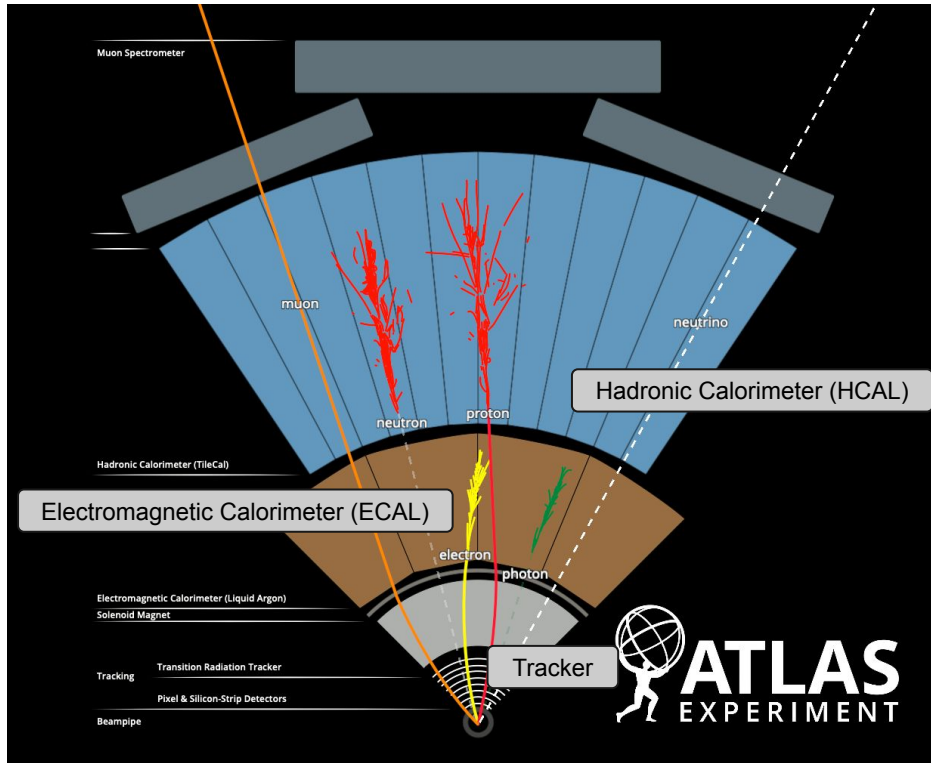


# Analysis strategy

---

- The analysis considers the **full Run 2** data set of 139 fb<sup>-1</sup> at 13 TeV.
- Processes from the SM produce a smoothly falling di-photon invariant mass spectrum.
- Different resonant and non-resonant regions and selections.
- **BDTs** (Boosted Decision Trees) are used to :
  - **Reject background** processes in (non)resonant  $HH$  searches.
  - Divide data into **invariant mass categories** to target different  $\kappa_\lambda$  ranges.
- Statistical results are obtained from a **fit  $m_{\gamma\gamma}$** .
- The **selection criteria** depend on the mass of the probed scalar particle.

# Overview of the ATLAS detector and object selection



## Object selection

- Trigger on di-photon events.
- Photons:
  - Tight Identification.
  - Loose Isolation.
- Jets:
  - Particle flow.
  - anti-kT algorithm.
  - b-tagger

# Event selection

## Preselection

$N_\gamma$	$\geq 2$
$m_{\gamma\gamma}$	[105,160] GeV
$p_T^\gamma(\text{leading})$	$> 0.35 \times m_{\gamma\gamma}$
$p_T^\gamma(\text{subleading})$	$> 0.25 \times m_{\gamma\gamma}$
$N_\mu + N_e$	0
$N_{b\text{-jets}}$	2
b-jets working point	77%
$N_{jets}( \eta  < 2.5)$	$< 6$

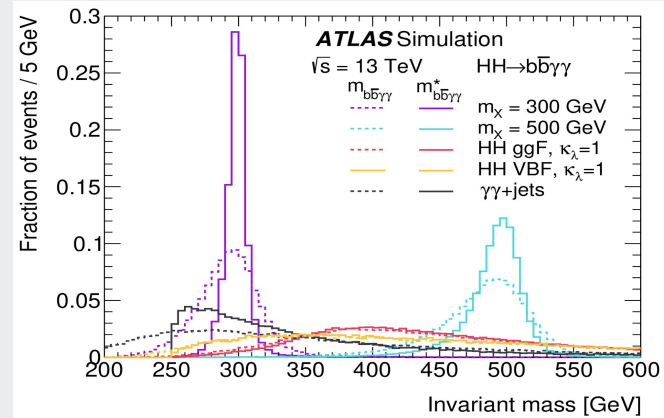
Acceptance x efficiency:

- SM ggF HH: 14%.
- $m_X = 300$  : 8.5%
- $m_X = 500$  : 14%

Alternative 4-body invariant mass is used:

$$m_{bb\bar{\gamma}\gamma}^* = m_{bb\bar{\gamma}\gamma} - m_{bb} - m_{\gamma\gamma} + 250 \text{ GeV}$$

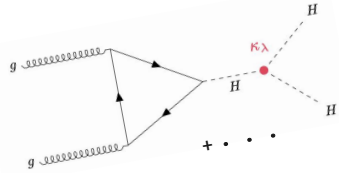
- Improves mass resolution.
- Detector resolutions effects cancel out.



# Selection nonresonant & resonant

Sidebands regions definition from the  $m_{\gamma\gamma}$  distribution:

- Sidebands (CRs):  $105 \leq m_{\gamma\gamma} \leq 120$  GeV and  $130 \leq m_{\gamma\gamma} \leq 160$  GeV.
- Signal region (Higgs peak):  $120 < m_{\gamma\gamma} < 130$  GeV



Nonresonant

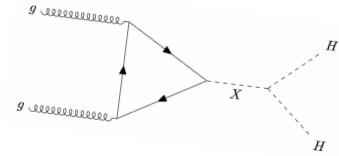
BDT

$$m_{bb\gamma\gamma}^* < 350 \text{ GeV}$$

BDT

$$m_{bb\gamma\gamma}^* > 350 \text{ GeV}$$

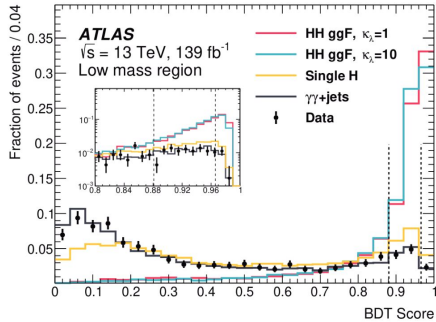
Resonant



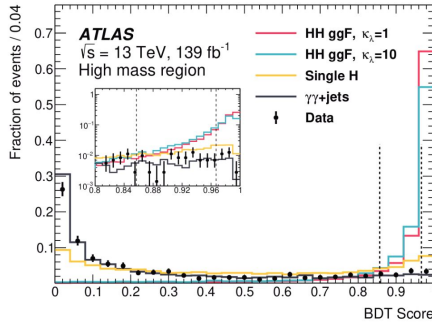
BDT for  $\gamma\gamma$  background.

BDT for single-Higgs background.

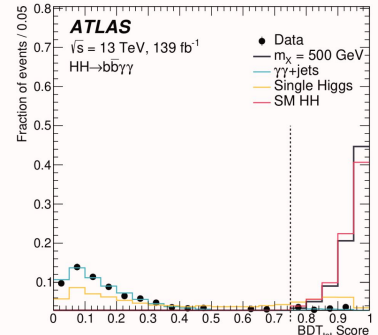
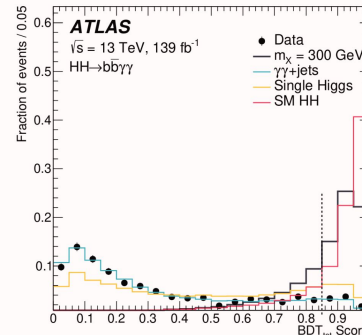
The two BDT scores are combined and BDT cut optimized by maximizing the signal significance.



(a) Low mass region



(b) High mass region



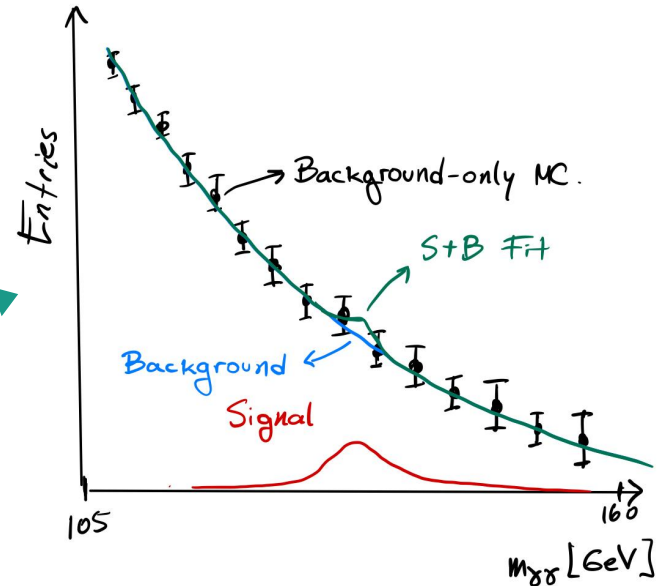
# Signal & Background parameterization

The signal and backgrounds are extracted by fitting analytic functions to the diphoton invariant mass distributions in the range of 105 -160 GeV.

- Di-Higgs and single-Higgs  $m_{\gamma\gamma}$  shapes are modelled using double sided Crystal Balls (gaussian core and power-law tails).
- Background is modelled with simple exponential function:

$$f_B = \exp(a \cdot m_{\gamma\gamma})$$

- Has least number of parameters.
- Passes spurious signal tests.



# Statistical analysis

- Maximum likelihood fit of the  $m_{\gamma\gamma}$  ( $105 < m_{\gamma\gamma} < 160$  GeV) has been performed for both **Nonresonant** & **Resonant** searches.

$$\mathcal{L} = \prod_c \left( \text{Pois}(n_c | N_c(\boldsymbol{\theta})) \cdot \prod_{i=1}^{n_c} f_c(m_{\gamma\gamma}^i, \boldsymbol{\theta}) \cdot G(\boldsymbol{\theta}) \right)$$

- Fit done simultaneously over all categories.
- Expected number of events  $N_c$  considers:
  - Di-Higgs boson production.
  - Single Higgs boson production.
  - Non-resonant background.
  - Spurious signals from background modelling.

# Systematic uncertainties

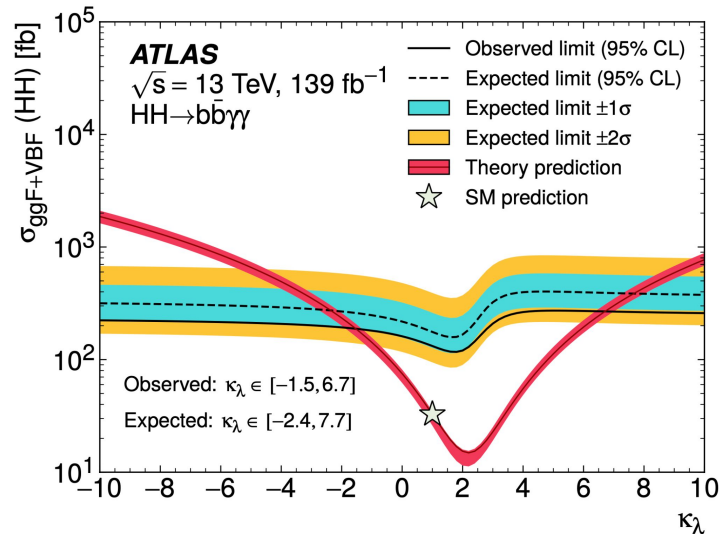
- The analysis is limited by statistical precision.
- The uncertainty on the integrated luminosity is 1.7%, obtained using the LUCID-2 detector.
- Dominant systematics:
  - Spurious signal
  - Parton showering model
  - Photon energy resolution.

Source	Type	Relative impact of the systematic uncertainties [%]	
		Nonresonant analysis <i>HH</i>	Resonant analysis $m_X = 300 \text{ GeV}$
Experimental			
Photon energy resolution	Norm. + Shape	0.4	0.6
Jet energy scale and resolution	Normalization	< 0.2	0.3
Flavor tagging	Normalization	< 0.2	0.2
Theoretical			
Factorization and renormalization scale	Normalization	0.3	< 0.2
Parton showering model	Norm. + Shape	0.6	2.6
Heavy-flavor content	Normalization	0.3	< 0.2
$\mathcal{B}(H \rightarrow \gamma\gamma, b\bar{b})$	Normalization	0.2	< 0.2
Spurious signal	Normalization	3.0	3.3

Impact of systematic uncertainties under 2% on the upper limit on the cross section.

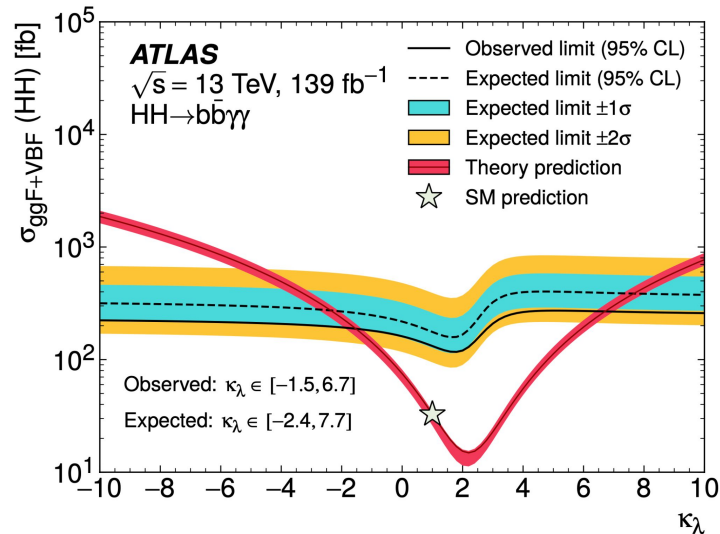
# Results. Nonresonant

- 95% CL upper limit on the **nonresonant** HH production cross section: 130 fb (expected 180 fb)
- Signal strength: observed (expected) = 4.2 (5.7).
- Upper limits on the  $HH$  production cross section are also computed as a function of  $\kappa_\lambda$ :
  - Single-Higgs-boson production cross section, Higgs decay BR and Higgs coupling strength set to SM values.

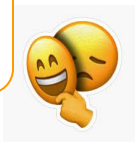


# Results. Nonresonant

- 95% CL upper limit on the **nonresonant** HH production cross section: 130 fb (expected 180 fb)
- Signal strength: observed (expected) = 4.2 (5.7).
- Upper limits on the  $HH$  production cross section are also computed as a function of  $\kappa_\lambda$ :
  - Single-Higgs-boson production cross section, Higgs decay BR and Higgs coupling strength set to SM values.

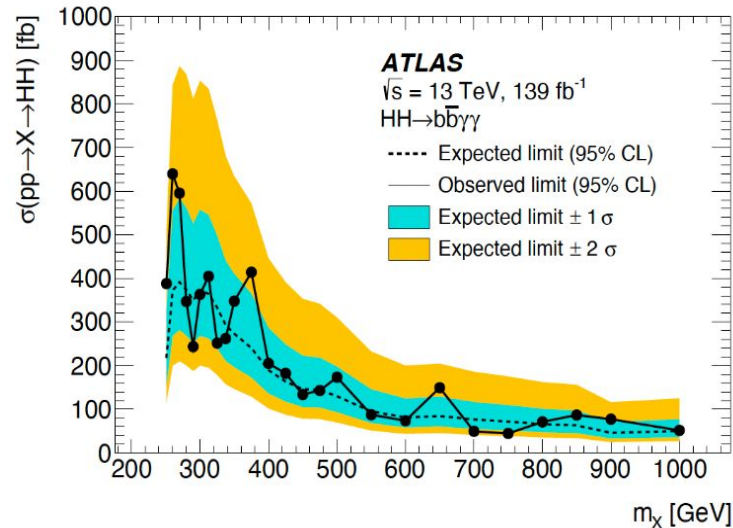


*No significant excess over the SM background expectations is found*



# Results. Resonant

- 95% CL upper limit on the production XS of a narrow-width resonance of a hypothetical scalar particle X.
- Limits in range  $251 \leq m_X \leq 1000$  GeV :
  - Observed: 640 - 44 fb.
  - Expected: 391 - 46 fb.
- Improvement up to a factor of 3 compared to previous ATLAS analysis.



*No significant excess over the SM background expectations is found*

# Summary & outlook

- **Di-Higgs searches**: probes the **Higgs potential** and is a window to **BSM physics**
- Search for nonresonant & resonant Higgs boson pair production in the  **$bb\gamma\gamma$**  final state with Run2 dataset has been discussed.
- Signal is modelled with double sided Crystal Ball, and background with one-parameter exponential
- BDT has been used to separate Signal from backgrounds in order to improve the sensitivity.
- No significant excess over the SM expectations has been found.
- The results improves the ATLAS limits on  $HH\rightarrow bb\gamma\gamma$  production cross section by up to **factor of five**, and allowed the  $\kappa_\lambda$  range shrink by about a **factor of two**.

## Outlook

- HL-LHC or Higgs factory will be essential to improve the precision of the measurement of the Higgs self-couplings!

---

# Back-up

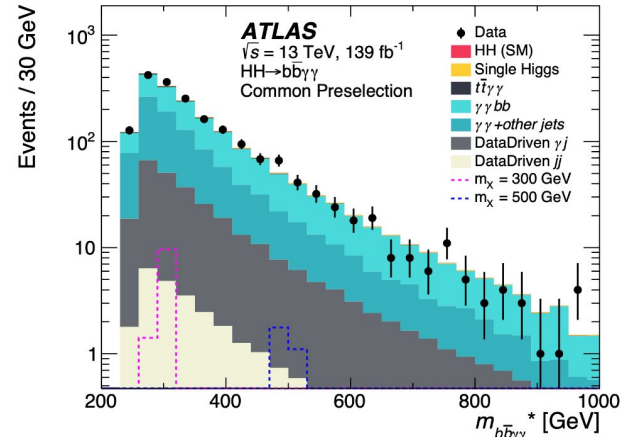
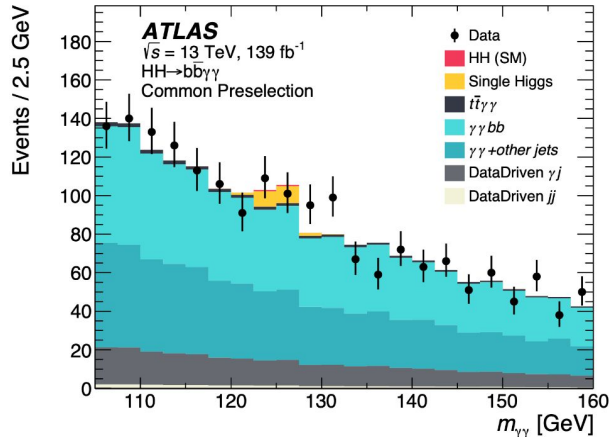
# BDT categories



Category	$\sigma_{\text{DSCB}}$ [GeV]
High mass BDT tight	$1.33 \pm 0.01$
High mass BDT loose	$1.47 \pm 0.02$
Low mass BDT tight	$1.50 \pm 0.06$
Low mass BDT loose	$1.64 \pm 0.03$
Resonant $m_X = 300$ GeV	$1.78 \pm 0.02$
Resonant $m_X = 500$ GeV	$1.46 \pm 0.01$

# Jet faking photons estimation

- Signal region is mainly composed of  $\gamma\gamma$ ,  $\gamma$ -jet and dijet events, where one or two jets are mis-identified as photons.
- Contribution of each one of them is estimated using the data driven method (ABCD) in the sidebands regions:
  - $\gamma\gamma$ : 85%
  - $\gamma$ -jet: 15%
  - Dijet: negligible.
- MC is normalized by the  $\gamma\gamma$ ,  $\gamma$ -jet and dijet fractions in the data sidebands regions.



# Simulation Samples

Process	Generator	PDF set	Showering	Tune
Nonresonant ggF $HH$	POWHEG Box v2 +FT [41] [42, 43]	PDFLHC [44]	PYTHIA 8.2 [67]	A14 [68]
Nonresonant VBF $HH$	MADGRAPH5_AMC@NLO [47]	NNPDF3.0NLO [69]	PYTHIA 8.2	A14
Resonant ggF $HH$	MADGRAPH5_AMC@NLO	NNPDF2.3LO	HERWIG 7.1.3 [53, 54]	H7.1 - Default [70]
ggF $H$	NNLOPS [71–73] [74, 75]	PDFLHC	PYTHIA 8.2	AZNLO [76]
VBF $H$	POWHEG Box v2 [41, 72, 77–83]	PDFLHC	PYTHIA 8.2	AZNLO
$WH$	POWHEG Box v2	PDFLHC	PYTHIA 8.2	AZNLO
$qq \rightarrow ZH$	POWHEG Box v2	PDFLHC	PYTHIA 8.2	AZNLO
$gg \rightarrow ZH$	POWHEG Box v2	PDFLHC	PYTHIA 8.2	AZNLO
$t\bar{t}H$	POWHEG Box v2 [78–80, 83, 84]	NNPDF3.0NLO	PYTHIA 8.2	A14
$bbH$	POWHEG Box v2	NNPDF3.0NLO	PYTHIA 8.2	A14
$tHq$	MADGRAPH5_AMC@NLO	NNPDF3.0NLO	PYTHIA 8.2	A14
$tHW$	MADGRAPH5_AMC@NLO	NNPDF3.0NLO	PYTHIA 8.2	A14
$\gamma\gamma$ +jets	SHERPA 2.2.4 [58]	NNPDF3.0NNLO	SHERPA 2.2.4	–
$t\bar{t}\gamma\gamma$	MADGRAPH5_AMC@NLO	NNPDF2.3LO	PYTHIA 8.2	–

# Simulation Samples



- Nonresonant ggF HH signal obtained from reweighting method based on simulations with different  $\kappa_\lambda$  values.
- VBF HH at LO and rescaled to N3LO.  $\kappa_\lambda$  dependence interpolated from second-order polynomial fit of MC samples.
- Resonant ggF HH generated at LO with 25 different  $m_\chi$  values in the range of [251, 1000] GeV.

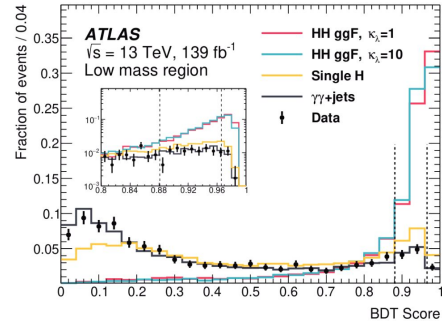
# Selection nonresonant & resonant

Sidebands regions definition from the  $m_{\gamma\gamma}^*$  distribution:

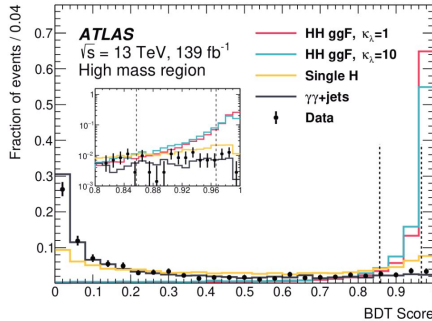
- Sidebands (CRs):  $105 \leq m_{\gamma\gamma}^* \leq 120$  GeV and  $130 \leq m_{\gamma\gamma}^* \leq 160$  GeV.
- Signal region (Higgs peak):  $120 < m_{\gamma\gamma}^* < 130$  GeV

## Nonresonant

- Events are divided in two mass regions: low mass ( $m_{bb\gamma\gamma}^* < 350$  GeV) and high mass ( $m_{bb\gamma\gamma}^* > 350$  GeV)
- A BDT is trained for each mass region.
- Two categories for each mass region defined based on the BDT score by maximizing the combined number-counting significance.



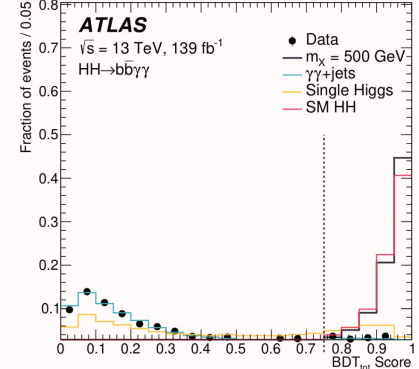
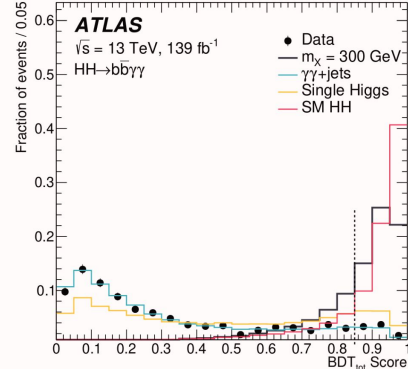
(a) Low mass region



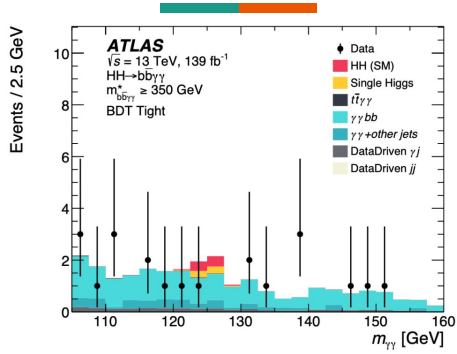
(b) High mass region

## Resonant

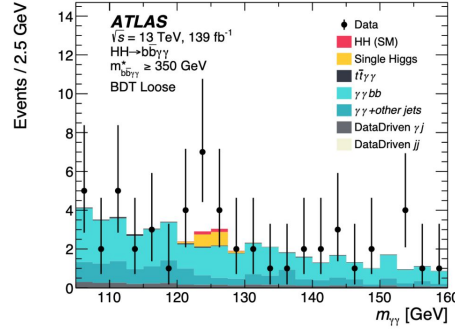
- Signal are reweighed to match the  $m_{bb\gamma\gamma}^*$  background distribution.
- Two BDT trained to separate the signal against two main sources of backgrounds:  $\gamma\gamma$  and single-Higgs backgrounds.
- The two BDT scores are combined and BDT cut optimized by maximizing the signal significance.



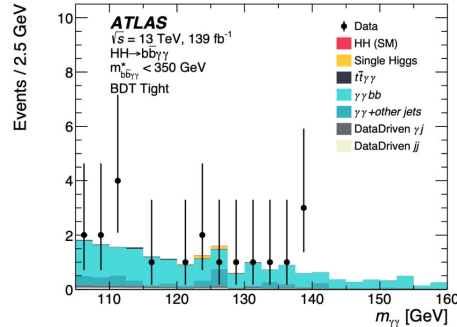
# Non-resonant regions definitions



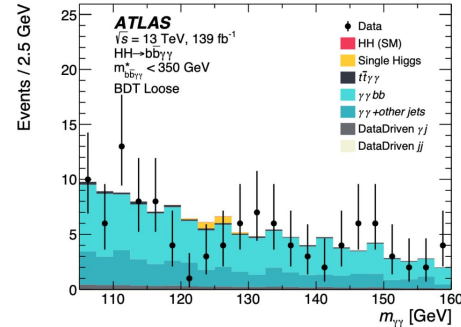
(a) High mass BDT tight selection



(b) High mass BDT loose selection



(c) Low mass BDT tight selection



(d) Low mass BDT loose selection

Category	Selection criteria
High mass BDT tight	$m_{b\bar{b}\gamma\gamma}^* \geq 350$ GeV, BDT score $\in [0.967, 1]$
High mass BDT loose	$m_{b\bar{b}\gamma\gamma}^* \geq 350$ GeV, BDT score $\in [0.857, 0.967]$
Low mass BDT tight	$m_{b\bar{b}\gamma\gamma}^* < 350$ GeV, BDT score $\in [0.966, 1]$
Low mass BDT loose	$m_{b\bar{b}\gamma\gamma}^* < 350$ GeV, BDT score $\in [0.881, 0.966]$

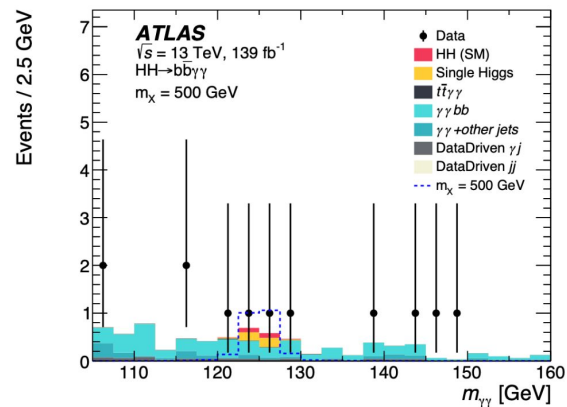
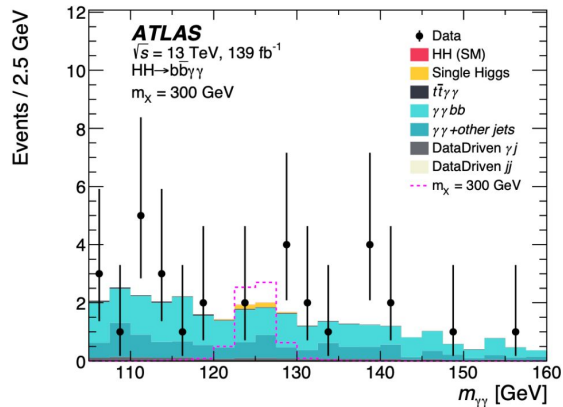
# Resonant regions

Resonant regions:

- Two BDTs are trained depending on the final state signature: single-Higgs or di-photon.
- BDTs are combined in quadrature as a function of  $C_1$  and  $C_2$  coefficients.

$$\text{BDT}_{\text{tot}} = \frac{1}{\sqrt{C_1^2 + C_2^2}} \sqrt{C_1^2 \left( \frac{\text{BDT}_{\gamma\gamma} + 1}{2} \right)^2 + C_2^2 \left( \frac{\text{BDT}_{\text{Single}H} + 1}{2} \right)^2}$$

- $C_1, C_2$  and BDT cut is optimized in a two step process:
  - Get signal significance for each value of  $C_1, C_2$  and BDT cut, for each mass point.
  - Obtain those  $C_1$  and  $C_2$  coefficients that are in a range of 5% around the maximum significance.



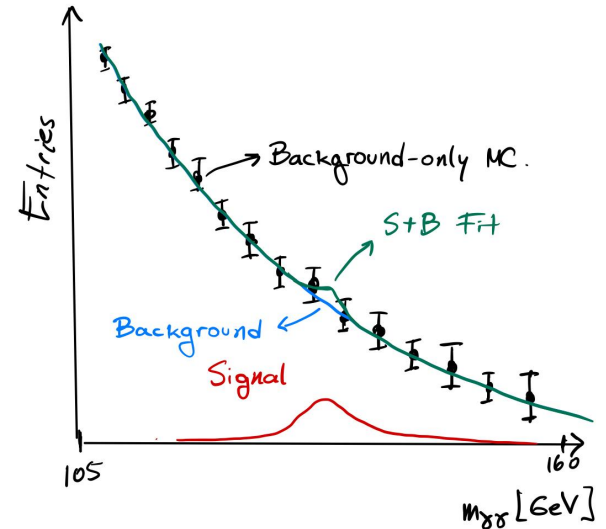
# Signal & Background parameterization

The signal and backgrounds are extracted by fitting analytic functions to the diphoton invariant mass distributions in the range of 105 -160 GeV.

- Di-Higgs and single-Higgs  $m_{\gamma\gamma}$  shapes are modelled using double sided Crystal Balls (gaussian core and power-law tails),
- MC  $m_{\gamma\gamma}$  distribution is fitted to obtain the shape parameters for each category.
  - Insensitive to specific signal processes.
- Background is modelled with simple exponential function:

$$f_B = \exp(a \cdot m_{\gamma\gamma})$$

- Passes spurious signal tests: fit with a S+B function a background-only MC distribution and extract the fitted signal.
- Has least number of parameters.



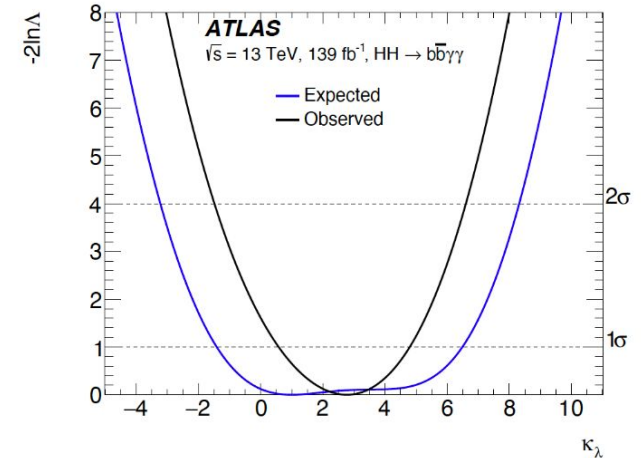
# Results . Nonresonant

The number of data events observed in the  $120 < m_{\gamma\gamma} < 130$  GeV window, the number of  $HH$  signal events expected for  $\kappa\lambda = 1$  and for  $\kappa\lambda = 10$

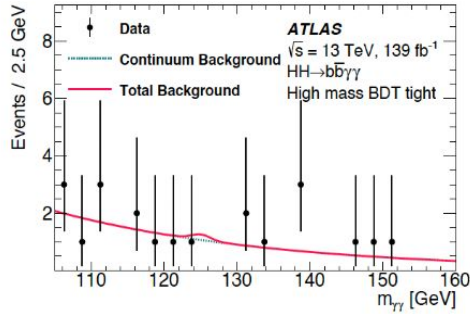
	High mass BDT tight	High mass BDT loose	Low mass BDT tight	Low mass BDT loose
Continuum background	$4.9^{+1.1}_{-1.3}$	$9.5^{+1.5}_{-1.7}$	$3.7^{+0.9}_{-1.1}$	$24.9^{+2.3}_{-2.5}$
Single Higgs boson background	$0.67^{+0.29}_{-0.13}$	$1.6^{+0.6}_{-0.2}$	$0.23^{+0.09}_{-0.03}$	$1.40^{+0.33}_{-0.16}$
$ggF+bbH$	$0.26^{+0.28}_{-0.16}$	$0.4^{+0.5}_{-0.2}$	$0.07^{+0.08}_{-0.04}$	$0.27^{+0.27}_{-0.16}$
$i\bar{i}H$	$0.19^{+0.03}_{-0.03}$	$0.49^{+0.09}_{-0.07}$	$0.107^{+0.022}_{-0.017}$	$0.75^{+0.13}_{-0.11}$
$ZH$	$0.142^{+0.035}_{-0.025}$	$0.48^{+0.09}_{-0.07}$	$0.040^{+0.020}_{-0.014}$	$0.27^{+0.06}_{-0.04}$
Rest	$0.074^{+0.032}_{-0.014}$	$0.16^{+0.07}_{-0.03}$	$0.012^{+0.008}_{-0.004}$	$0.111^{+0.030}_{-0.012}$
SM $HH(\kappa_\lambda = 1)$ signal	$0.87^{+0.10}_{-0.18}$	$0.37^{+0.04}_{-0.07}$	$0.049^{+0.006}_{-0.010}$	$0.078^{+0.008}_{-0.015}$
$ggF$	$0.86^{+0.10}_{-0.18}$	$0.35^{+0.04}_{-0.07}$	$0.046^{+0.006}_{-0.010}$	$0.072^{+0.008}_{-0.015}$
VBF	$(12.6^{+1.3}_{-1.2}) \cdot 10^{-3}$	$(16.1^{+1.4}_{-1.2}) \cdot 10^{-3}$	$(3.2^{+0.4}_{-0.4}) \cdot 10^{-3}$	$(6.9^{+0.5}_{-0.6}) \cdot 10^{-3}$
Alternative $HH(\kappa_\lambda = 10)$ signal	$6.5^{+1.0}_{-0.8}$	$3.6^{+0.6}_{-0.4}$	$4.5^{+0.7}_{-0.6}$	$8.5^{+1.3}_{-1.0}$
Data	2	17	5	14

An alternative statistical analysis consists in determining the best-fit value of the  $\kappa\lambda$  coupling modifier.

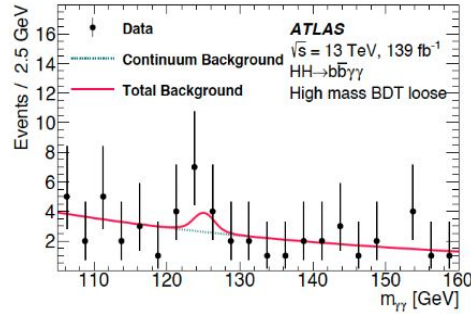
The best-fit value of  $\kappa\lambda$  and its uncertainty are obtained by means of a negative log-likelihood scan.



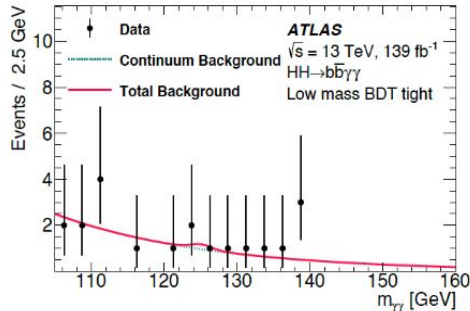
# Results . Nonresonant



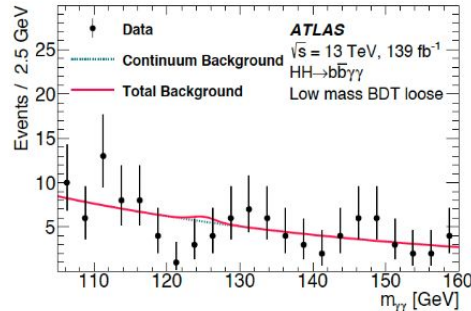
(a) High mass BDT tight



(b) High mass BDT loose



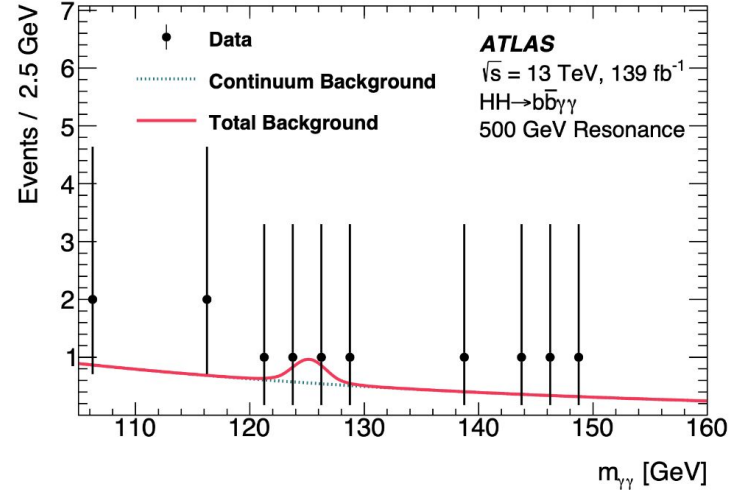
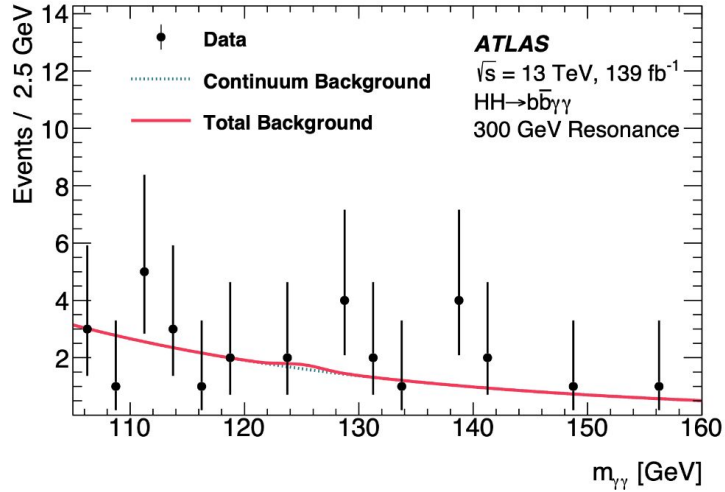
(c) Low mass BDT tight



(d) Low mass BDT loose

Comparison of the background only-fit with data for the four categories of the nonresonant search

# Results . Resonant



Comparison of the Data with the background-only fit for

- $m_X = 300 \text{ GeV}$  (left)
- $m_X = 500 \text{ GeV}$  (right)