Search for a resonance decaying into a scalar particle and a Higgs boson in the final state with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s}$=13 TeV with the ATLAS detector

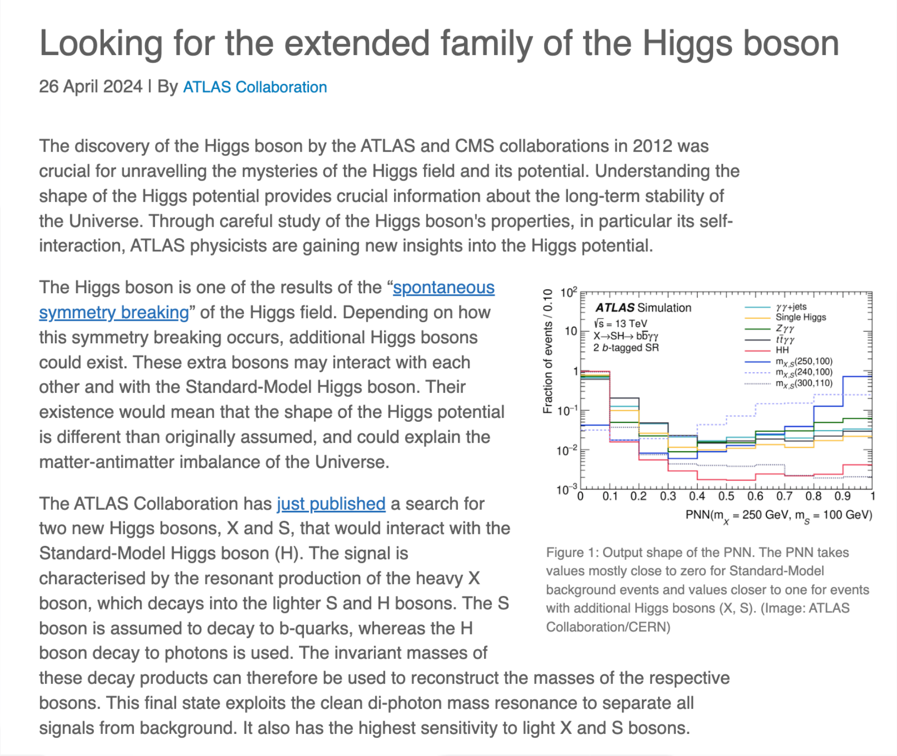
Xi Wang
SJTU, CPPM

FCPPL, Jun 11, 2024, BORDEAUX

饮 水 思 源 • 爱 国 荣 校

Contents

Analysis team



Looking for the extended family of the Higgs boson

26 April 2024 I By ATLAS Collaboration

The discovery of the Higgs boson by the ATLAS and CMS collaborations in 2012 was crucial for unravelling the mysteries of the Higgs field and its potential. Understanding the shape of the Higgs potential provides crucial information about the long-term stability of the Universe. Through careful study of the Higgs boson's properties, in particular its self-interaction, ATLAS physicists are gaining new insights into the Higgs potential.

The Higgs boson is one of the results of the "spontaneous symmetry breaking" of the Higgs field. Depending on how this symmetry breaking occurs, additional Higgs bosons could exist. These extra bosons may interact with each other and with the Standard-Model Higgs boson. Their existence would mean that the shape of the Higgs potential is different than originally assumed, and could explain the matter-antimatter imbalance of the Universe.

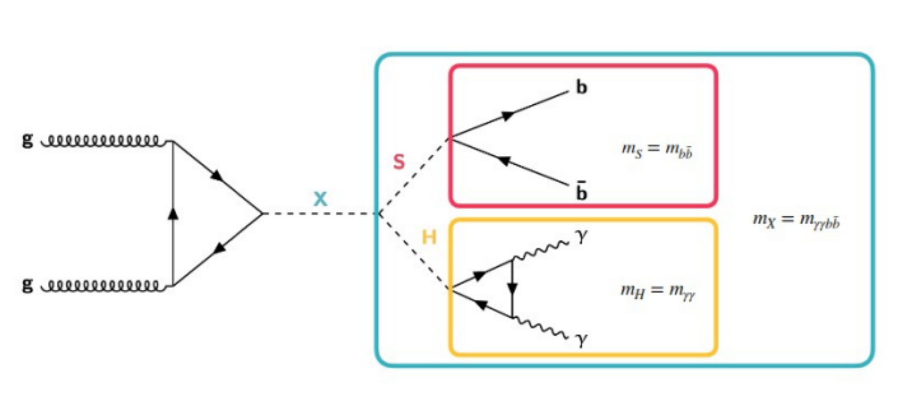
The ATLAS Collaboration has just published a search for two new Higgs bosons, X and S, that would interact with the Standard-Model Higgs boson (H). The signal is characterised by the resonant production of the heavy X boson, which decays into the lighter S and H bosons. The S boson is assumed to decay to b-quarks, whereas the H boson decay to photons is used. The invariant masses of these decay products can therefore be used to reconstruct the masses of the respective bosons. This final state exploits the clean di-photon mass resonance to separate all signals from background. It also has the highest sensitivity to light X and S bosons.

Figure 1: Output shape of the PNN. The PNN takes values mostly close to zero for Standard-Model background events and values closer to one for events with additional Higgs bosons (X, S). (Image: ATLAS Collaboration/CERN)

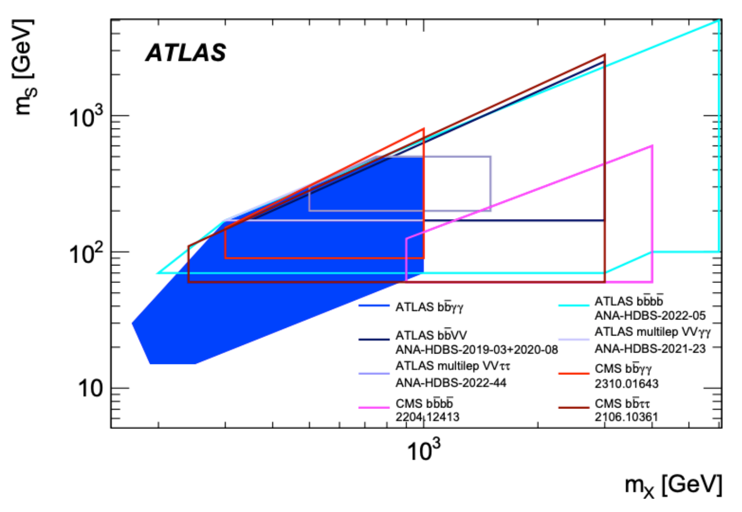
Physics briefing: Looking for the extended family of the Higgs boson

Introduction and Motivation

- Many BSM models predicts an extended Higgs sector where the 125 GeV Higgs is one of several physical states, such as 2-HDM, NMSSM and TRSM
- We search for two additional such scalars X and S with in the full Run 2 data
- We analyze the $bb\gamma\gamma$ final state with $H \rightarrow \gamma\gamma$ and $S \rightarrow bb$ in a mass grid of $170\ GeV < m_X < 1000\ GeV$ and $15GeV < m_X < 500\ GeV$, using the full Run2 data
- Limits set on $\sigma(X \rightarrow SH \rightarrow b\bar{b}\gamma\gamma)$



Gluon-gluon fusion production of a scalar X decaying into a Scalar S and a Standard Model Higgs boson



Phase space probed by the present $X \rightarrow S(b\bar{b})H(\gamma\gamma))$ analysis(in full blue), Compared to other CMS and ATLAS equivalents

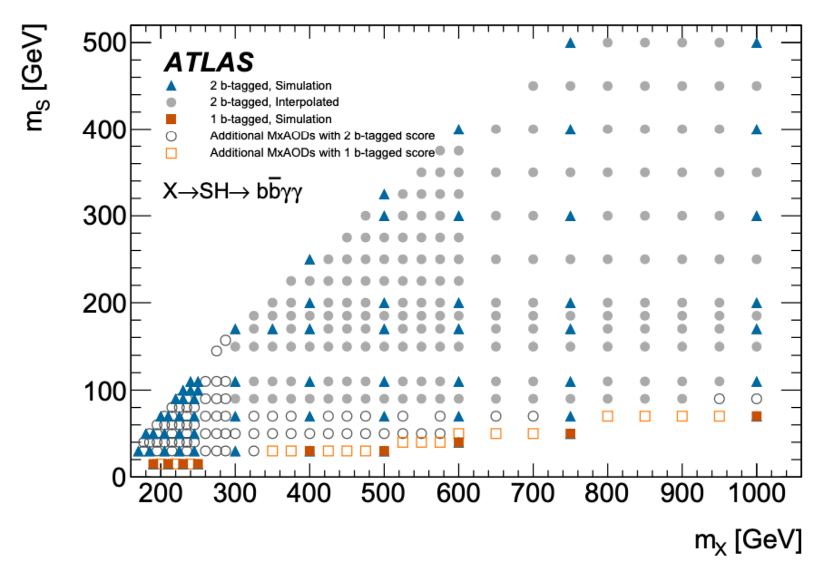
we can go to low masses thanks to the di-photon trigger

Samples, backgrounds

- Signal MC samples generated to cover a large range and ensure smooth interpolation
- Main background processes
  - Non-resonant backgrounds: $\gamma\gamma + jets$, $tt\gamma\gamma$ and $Z(\rightarrow q\bar{q})\gamma\gamma$
  - Resonant backgrounds: single Higgs, Di-Higgs

Process	Generator	PDF set	Showering	Tune
$X \to SH$	Pythia 8.2 [30]	NNPDF2.3lo [31]	Pythia 8.2 [30]	A14 [32]
$\gamma\gamma$+jets	Sherpa 2.2.4 [33]	NNPDF3.0nnlo [34]	–	–
$t\bar{t}\gamma\gamma$	MadGraph5_aMC@NLO [35]	NNPDF2.3lo	Pythia 8.2	A14
$Z(\to q\bar{q})\gamma\gamma$	Sherpa 2.2.11 [33]	NNPDF3.0nnlo	–	–
ggF H	Nnlops [36–38] [39, 40]	PDF4LHC15 [41]	Pythia 8.2	AZNLO [42]
VBF H	Powheg Box v2 [43–46]	PDF4LHC15	Pythia 8.2	AZNLO
WH	Powheg Box v2 [47, 48]	PDF4LHC15	Pythia 8.2	AZNLO
$qq \to ZH$	Powheg Box v2 [47, 48]	PDF4LHC15	Pythia 8.2	AZNLO
$gg \to ZH$	Powheg Box v2 [47, 48]	PDF4LHC15	Pythia 8.2	AZNLO
$t\bar{t}H$	Powheg Box v2 [49]	NNPDF3.0nlo	Pythia 8.2	A14
$b\bar{b}H$	Powheg Box v2 [37]	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
ggF HH	Powheg Box v2 +FT [46, 50, 51]	PDFLHC	Pythia 8.2	A14
VBF HH	MadGraph5_aMC@NLO	NNPDF3.0nlo	Pythia 8.2	A14

Main signal and background samples, split by production mode



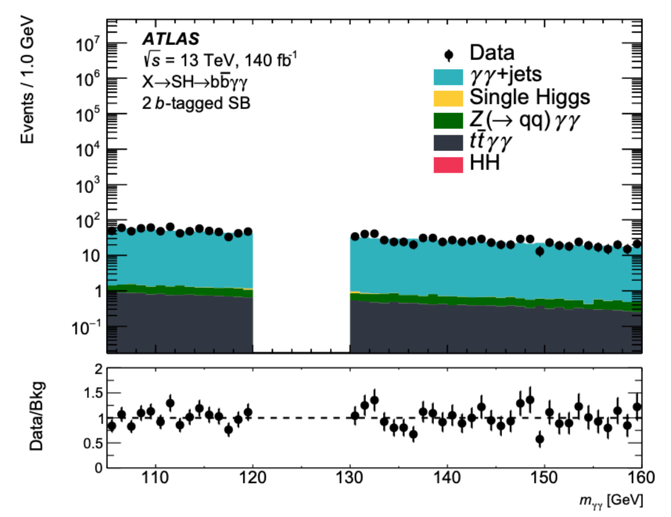
MC samples and interpolated mass points

Event selection

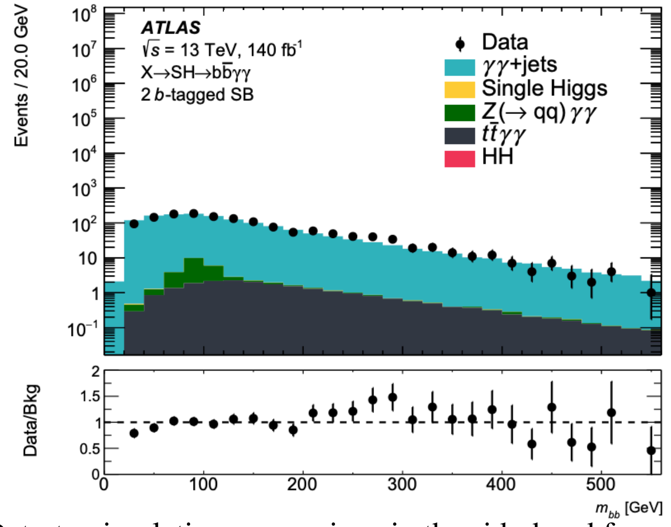
- Event preselection
  - Trigger: g35_medium_g25_medium
  - $E_T/m_{\gamma\gamma}$ > 0.35(0.25) on the leading/sub-leading photons
  -

	2 b-tagged	1 b-tagged
Number of 'tight' and isolated photons	≥ 2	
$m_{\gamma\gamma}$ [GeV]	$\in [105, 160]$	
Number of leptons	$= 0$	
Number of central jets	$\in [2, 5]$	
Number of b-tagged jets @ 77% WP	$= 2$	$= 1$

- Signal region: $m_{\gamma\gamma}\ [GeV] \in (120,\ 130)$
- Side-bands: $m_{\gamma\gamma}\ [GeV] \in 105,\ 120] \cup [130,\ 160]$
- Good agreement between data and MC samples in sidebands



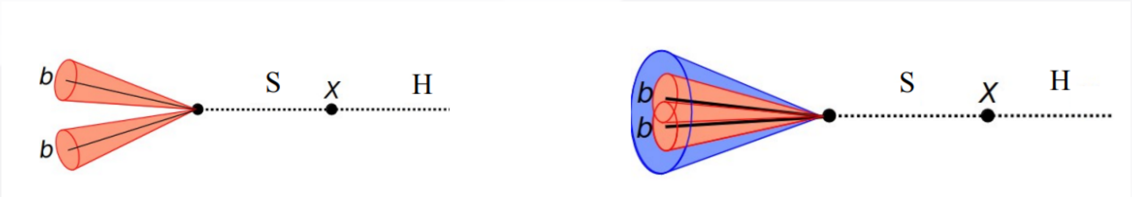
Data-to-simulation comparison in the side-band for $m_{\gamma\gamma}$



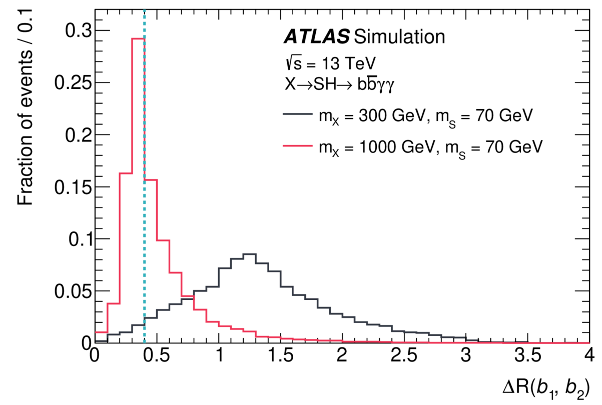
Data-to-simulation comparison in the side-band for $m_{b\bar{b}}$

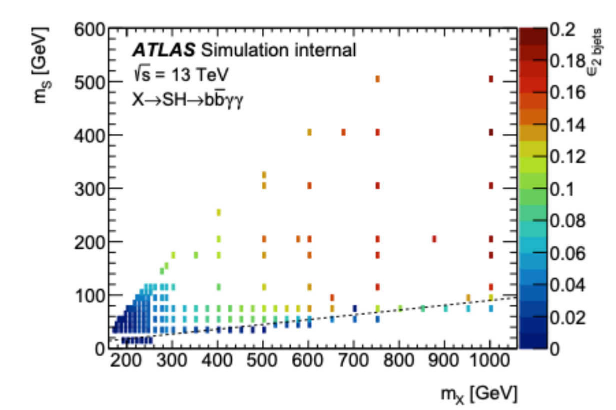
Event selection: two categories

- When $m_X >> m_S + m_H$, the particle S becomes so boosted that its decay products, the 2 b-quarks are very collimated and hard to be resolved within $\Delta R = 0.4$.

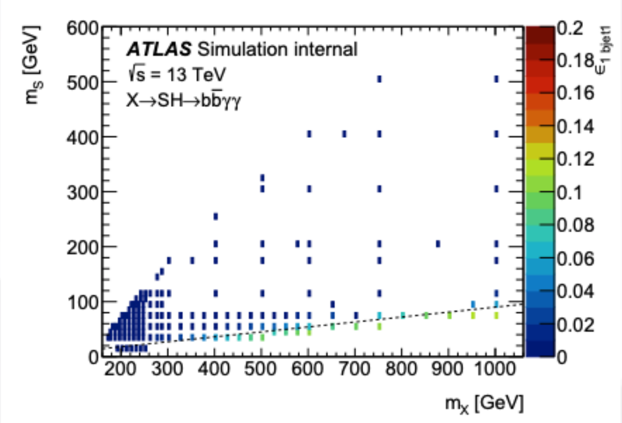


Topology of Boosted particle S





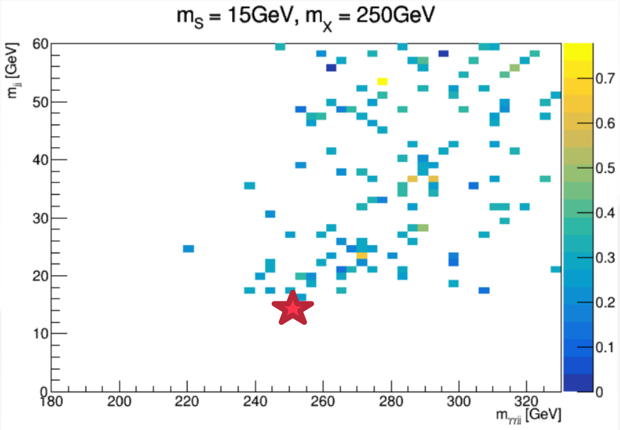
(a) 2 b-tagged efficiency



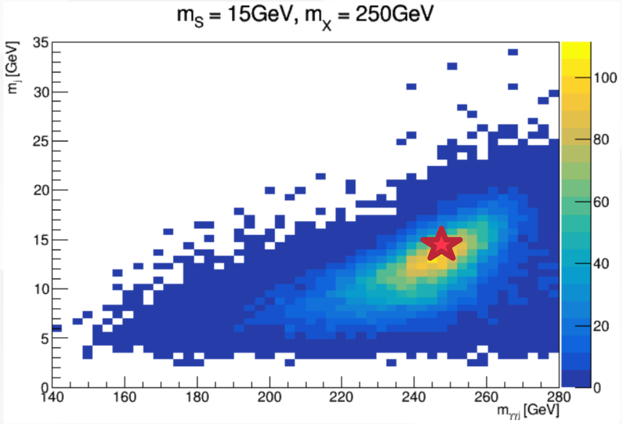
(b) 1 b-tagged efficiency

- 2 b-tagged category and 1 b-tagged category are split

  - expected limits are used to choose the category



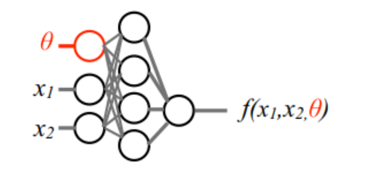
2 b-tagged category

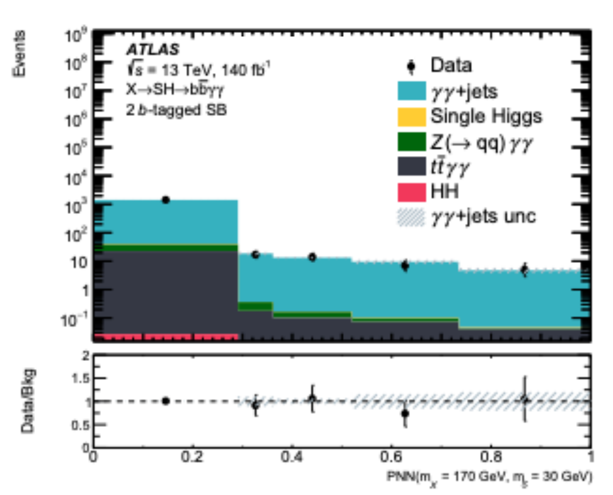


1 b-tagged category

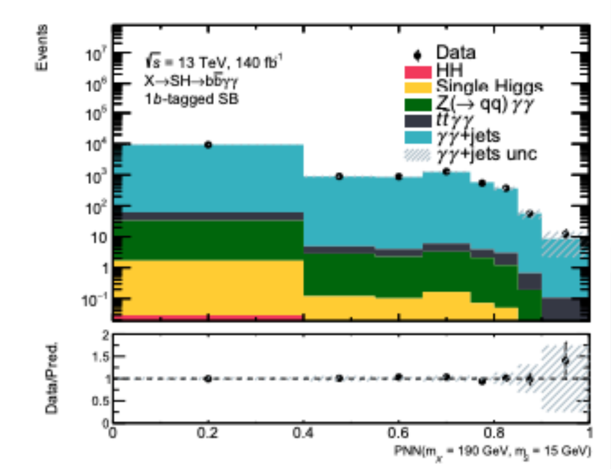
Analysis strategy: PNN

- Parameterised Neural Networks(PNN) is used to distinguish signal and background in the signal region. The PNN shapes depend on which m_X, m_S are chosen to target signals.





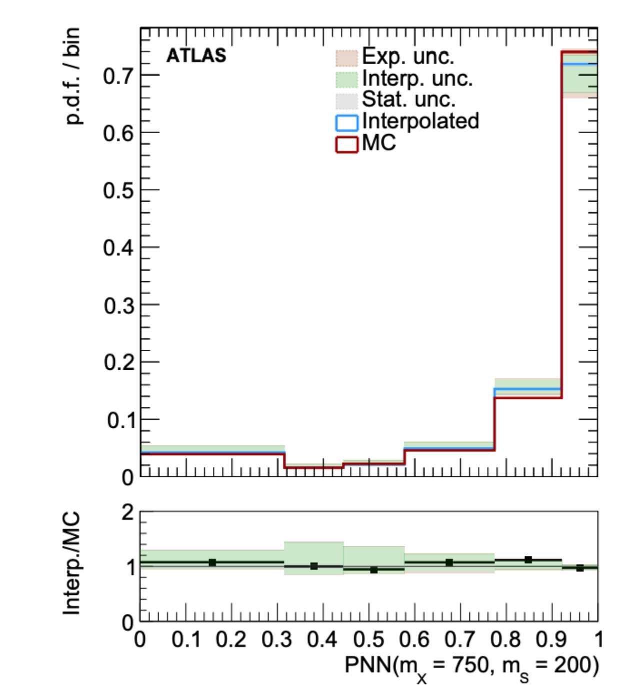
- Two separate PNNs
  - 2 b-tagged: $$m^*_{bb\gamma\gamma} = m_{bb\gamma\gamma} - (m_{\gamma\gamma} - 125\text{ GeV})$$
    - Parameter $\theta = m_X, m_S$
    - Training samples : signal, ttH, ggH, ZH and γγ + jets backgrounds (no HH – too signal like and confuses the network)
    - Training variables : m_{bb} and $m^*_{bb\gamma\gamma}$
  - 1 b-tagged: $$m^*_{b\gamma\gamma} = m_{b\gamma\gamma} - (m_{\gamma\gamma} - 125\text{ GeV})$$
    - Parameter $\theta = m_X$
    - Training samples : signal, VBFH, HH, ttH, ggH, ZH and γγ + jets backgrounds
    - Training variables : p_b^T and $m^*_{b\gamma\gamma}$ (m_b is not calibrated)

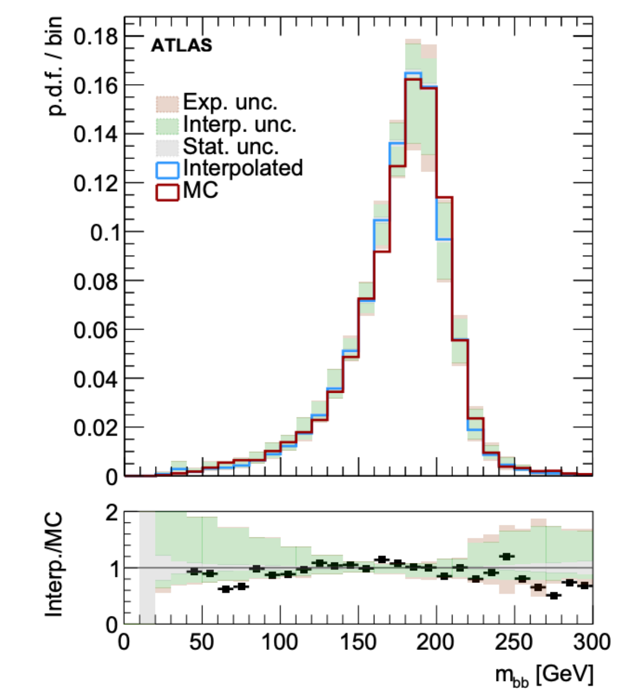


- Good agreement between data and MC in side bands.

Analysis strategy: Interpolation

- To fully exploit PNN and set continuous limits, we interpolated between signal mass points:
  - Rescaling of 4-vectors of S and H in c.o.m. frame to match kinematics of nearby points
  - Resolution interpolation using Bukin probabilities interpolated through Delaunay triangulation
- Additional interpolation systematic added
- Grid granularity chosen after thorough signal interpolation study
- At low mass, the quality of PNN shape after interpolation is too low to be useful
- For the 1 b-tagged category, resolution interpolation not possible
- Implemented for $m_X > 300\_GeV, m_S > 70\ GeV$





Comparison of MC direct prediction and interpolation for PNN and $m_{b\bar{b}}$

Systematic uncertainties

- Systematic uncertainties affect the overall normalization of backgrounds, but also the shape of the PNN distribution.
- Main systematic uncertainties: Flavor tagging(up to 6% impact), Jets(up to 15% at low mass), Combined signal theory uncertainties(up to 10% at high mass), Modeling uncertainty(up to 20%)

		Signal	HH ggF	HH VBF	ttH & ZH	Other Single Higgs	Continuum $\gamma\gamma$+jets
Theory	Normalisation	$BR(H \to \gamma\gamma)$	$BR(H \to \gamma\gamma)$ $BR(H \to b\bar{b})$ PDF+α_S Scales + m_t	$BR(H \to \gamma\gamma)$ $BR(H \to b\bar{b})$ PDF+α_S Scales	$BR(H \to \gamma\gamma)$	$BR(H \to \gamma\gamma)$ PDF+α_S Scales	$\gamma\gamma$ transfer factor
	Shape+Norm.	Scales, PDF+α_S Parton shower Interpolation	Parton Shower		Scales, PDF+α_S Parton Shower		Scales, PDF+α_S Modelling
Exp.	Shape+Norm.	Pile-up modelling Diphoton trigger efficiency Photon identification and isolation efficiency Photon energy scale and resolution Jet energy scale and resolution Jet vertex tagger efficiency Flavour tagging efficiency	(all exp. systematics are neglected for bbH, tH and VBF H)				

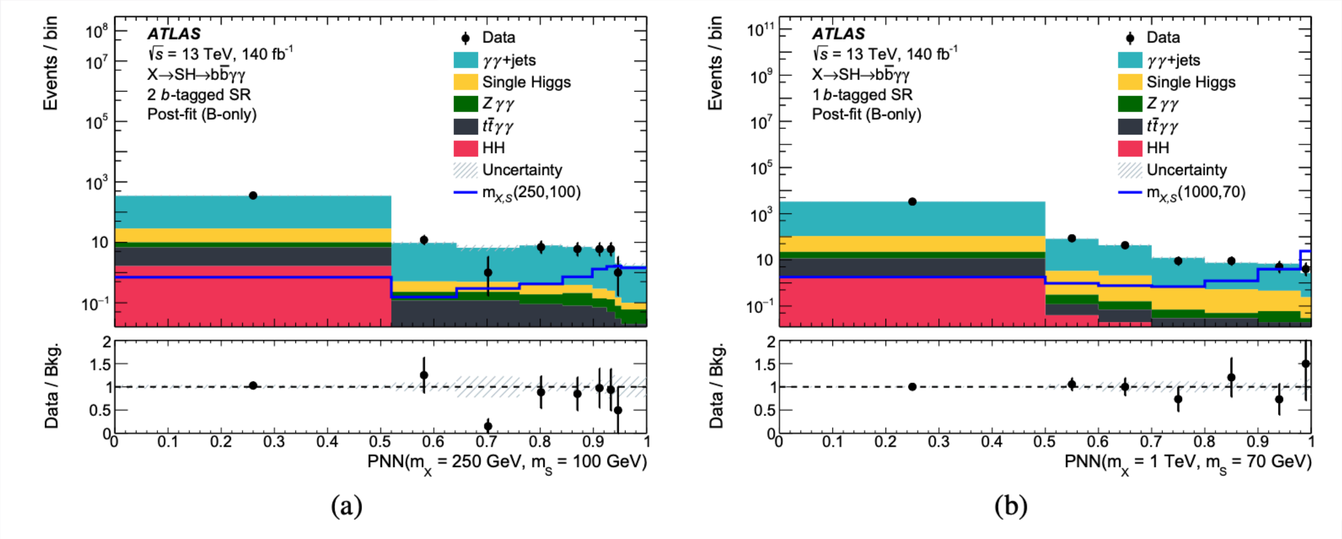
Summary of the systematic uncertainties included in the final results.

Results: maximum-likelihood fit

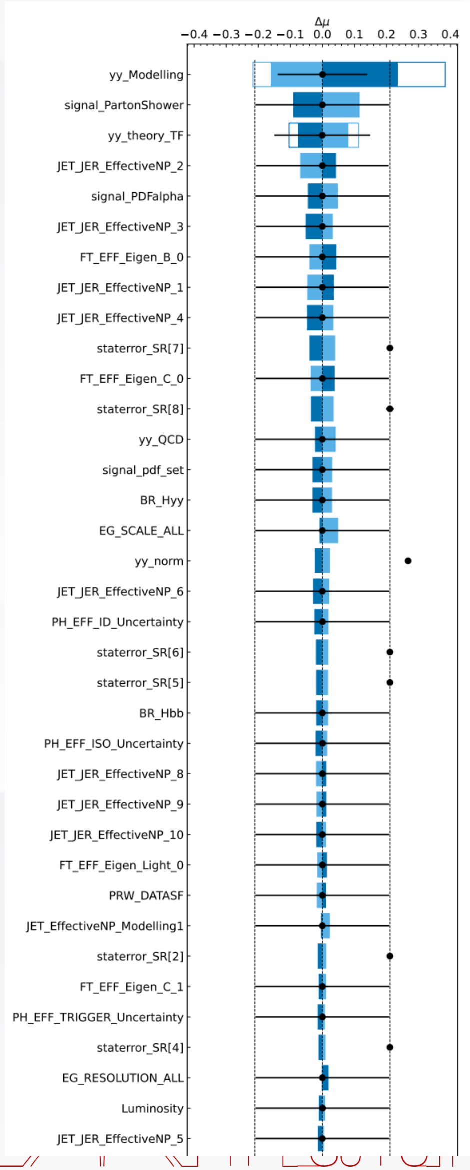
- a maximum-likelihood fit of the binned PNN output distribution

$$\mathcal{L} = \mathrm{Pois}\left(n_{\mathrm{SB}} \middle| \mu_{\gamma\gamma} N_{\mathrm{SB}}^{\gamma\gamma}(\boldsymbol{\theta}) + \sum_{p} N_{\mathrm{SB}}^{p}(\boldsymbol{\theta})\right) \cdot \prod_{i} \mathrm{Pois}\left(n_{\mathrm{SR},i} \middle| \mu_{\gamma\gamma} N_{\mathrm{SR}}^{\gamma\gamma}(\boldsymbol{\theta}) f_i^{\gamma\gamma}(\boldsymbol{\theta}) + \sum_{p} N_{\mathrm{SR}}^{p}(\boldsymbol{\theta}) f_i^{p}(\boldsymbol{\theta})\right) \cdot G(\boldsymbol{\theta})$$

- Post-fit distributions of PNN discriminant output

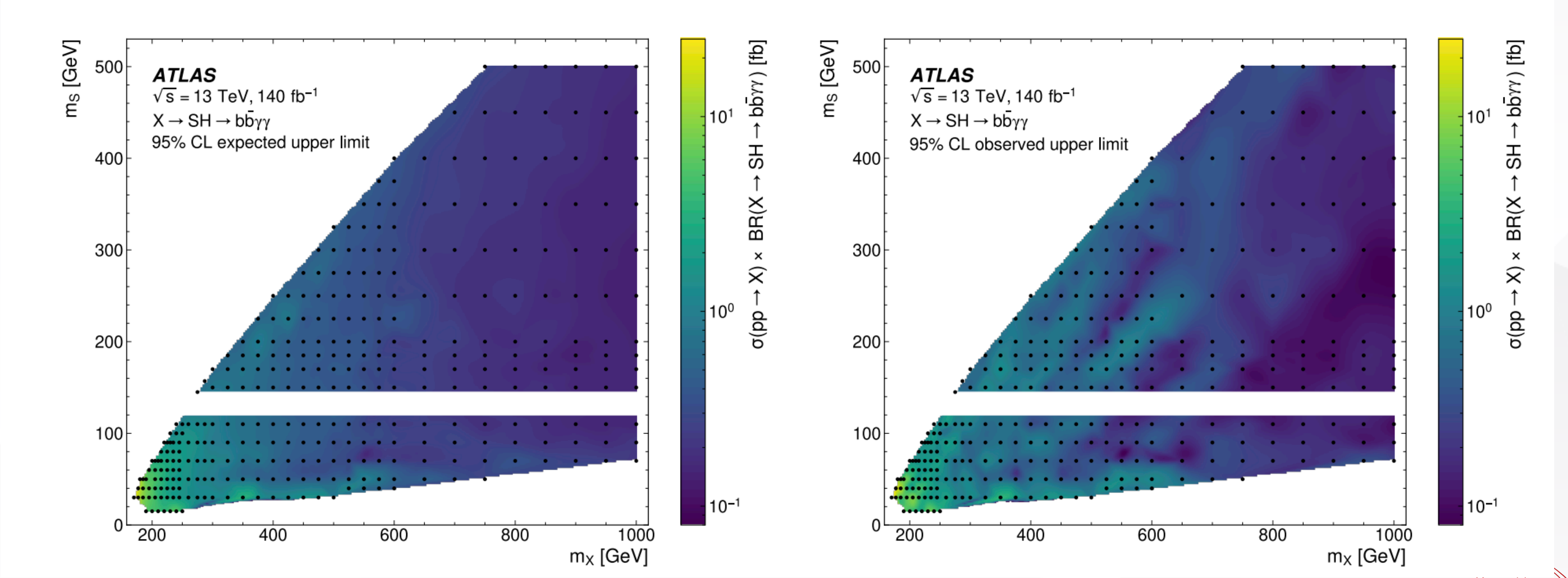


- Nuisance parameters are ranked by their impact to POI



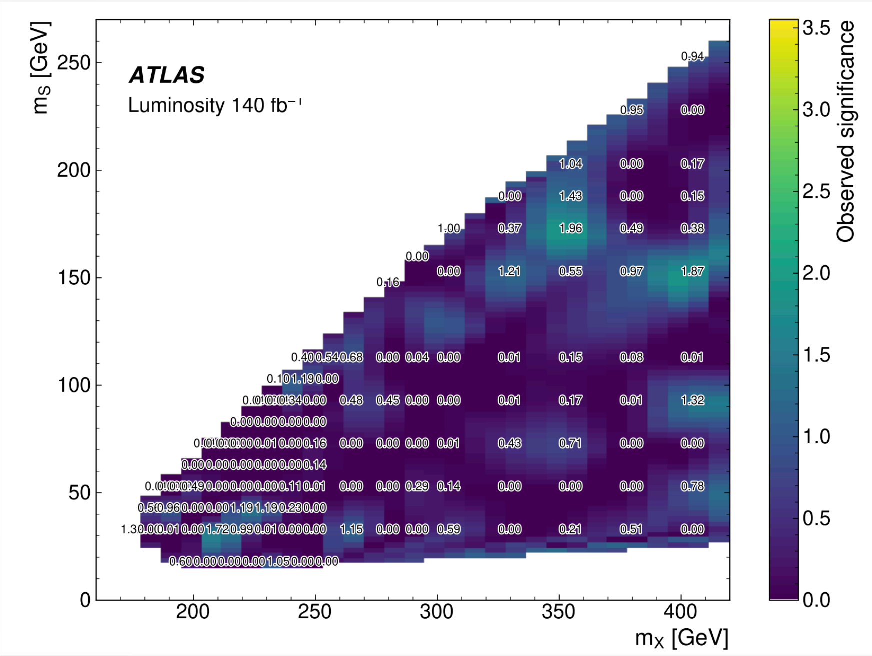
Results: Expected and Observed Upper Limits

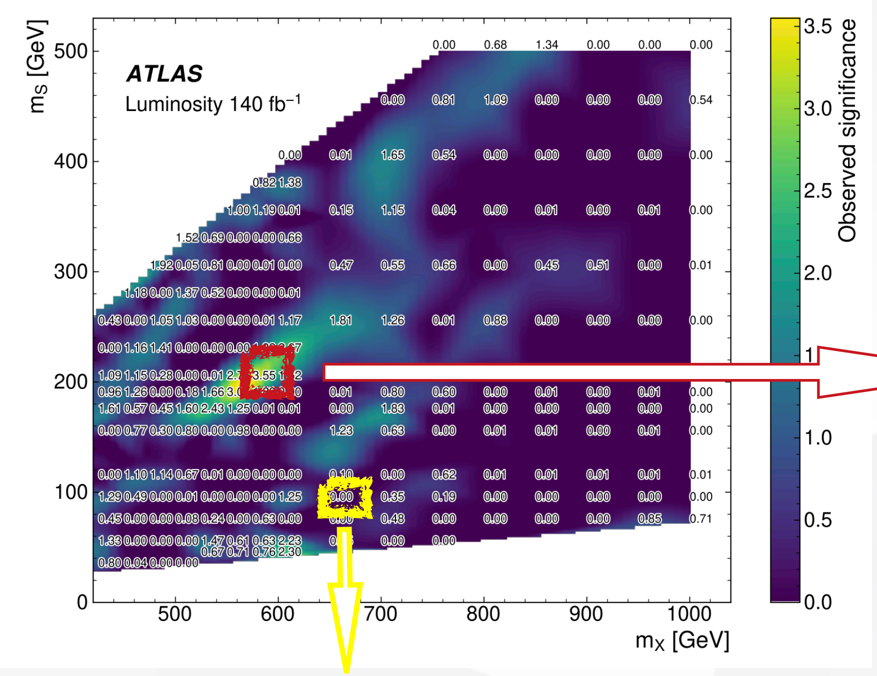
- Expected and observed 95% CL upper limits on the signal cross section times branching fraction for the $X \to SH$ signal, in the (m_X, m_S) plane. The band at $m_S = 125\ GeV$ is not shown as those points were probed in a previous analysis.
- The upper limits improve at higher masses, consistent with the fact that signals with higher mX become easier to differentiate from SM processes.
- In contrast, the upper limits are worsen at lower mS where the signal becomes predominantly boosted, while at low mX the sensitivity suffers from an increasing fraction of b-jets falling below the jet pT reconstruction threshold.

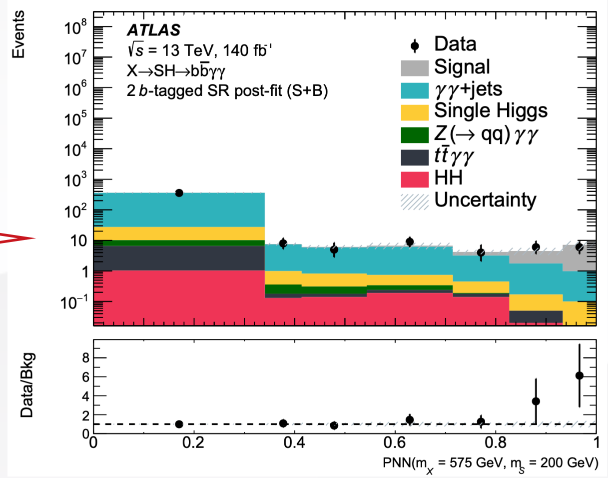


Results: Observed Local Significance

- For most mass points good agreement is observed between data and the SM background-only expectation
- A largest excess at $(m_X = 575\ GeV, m_S = 200\ GeV)$ with a local significance of 3.546σ
- The 'look-elsewhere effect' is taken into account using the asymptotic method, and the resulting global significance is calculated to be 2.0σ.







0σ observed at the point $(m_X = 650\ GeV, m_S = 90\ GeV)$ where CMS has a small excess

Global significance

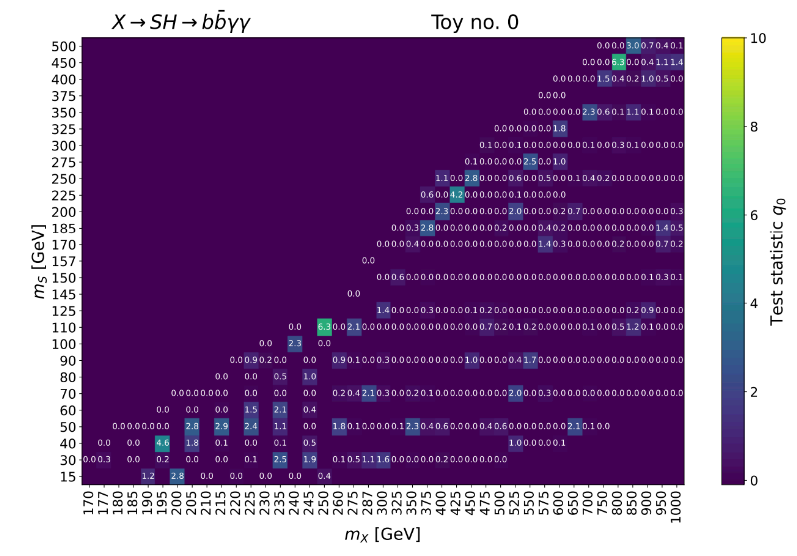
- The global significance (or, equivalently, the global p_0 value), quantifies the probability of observing a local deviation from the expected background anywhere within the search range under the background-only hypothesis:

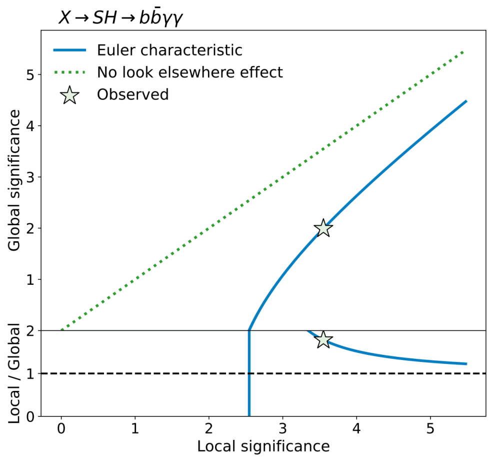
$$\text{Global } p_0 = \mathbb{P}(\max_{\theta \in \mathcal{M}} q(\theta) > q_{obs} | \mu = 0)$$

- For a 2-dimensional search space, the expected Euler characteristic $\phi(A_\mu)$ follows the equation:

$$\mathbb{E}[\phi(A_u)] = \mathbb{P}(\chi^2 > u) + e^{-u/2} \cdot (\mathcal{N}_1 + \sqrt{u} \cdot \mathcal{N}_2)$$

- 20 Background only toys are generated and a 2-dimensional map of test statistic q_0 are produced.
- Global significance is obtained after evaluating and fitting Euler characteristics.





Global significance	2.094 ± (−0.119, +0.165)
Maximum of the local p_0	0.000385
Global p_0	0.046358
Trial factor (global p_0 / local p_0)	240.678

Conclusion and prospects

- A search for a signal from a hypothetical scalar X is performed, considering the case where it decays into another hypothetical scalar S and a Higgs boson, which subsequently decay into pairs of b-quarks and photons, respectively.
- Two signal regions targeting resolved or boosted $S \rightarrow b\bar{b}$ decays are analysed using parameterised neural networks, which provide continuous sensitivity in the probed $(m_X,\ m_S)$ plane.
- No significant excess with respect to the Standard Model background is found. Therefore, 95% CL upper limits are set on $\sigma(X \rightarrow SH \rightarrow b\bar{b}\gamma\gamma)$ in the ranges $170\ GeV \leq m_X \leq 1000\ GeV$ and $15\ GeV \leq m_X \leq 500\ GeV$, expanding earlier LHC results to lower masses and providing higher sensitivity.
- The largest deviation from the background-only expectation occurs for $(m_X = 575\ GeV, m_S = 200\ GeV)$ with a local (global) significance of 3.5 (2.0) standard deviations.
- The paper has submitted to the arXiv. Run2+Run3 $X \rightarrow SH \rightarrow b\bar{b}\gamma\gamma$ analysis is ongoing, using easy jet produced n-tuples and looking at the Run3 limits using G2N.

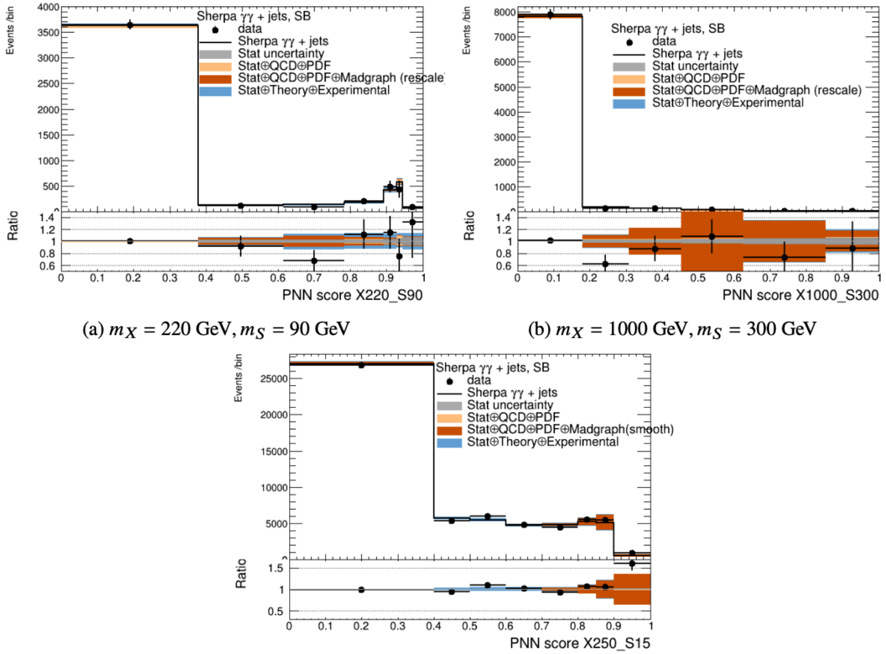
Thanks for your attention

饮 水 思 源　　爱 国 荣 校

Backup

Systematic uncertainties

- The full systematic uncertainties also allow additional checks on the validity of PNN – sideband data to background estimate in SB regions well covered by uncertainty estimates.
- These checks show that the difference data to background in general is smaller than the difference data to alternative MC(+)QCD(+)Madgraph(+)PDF



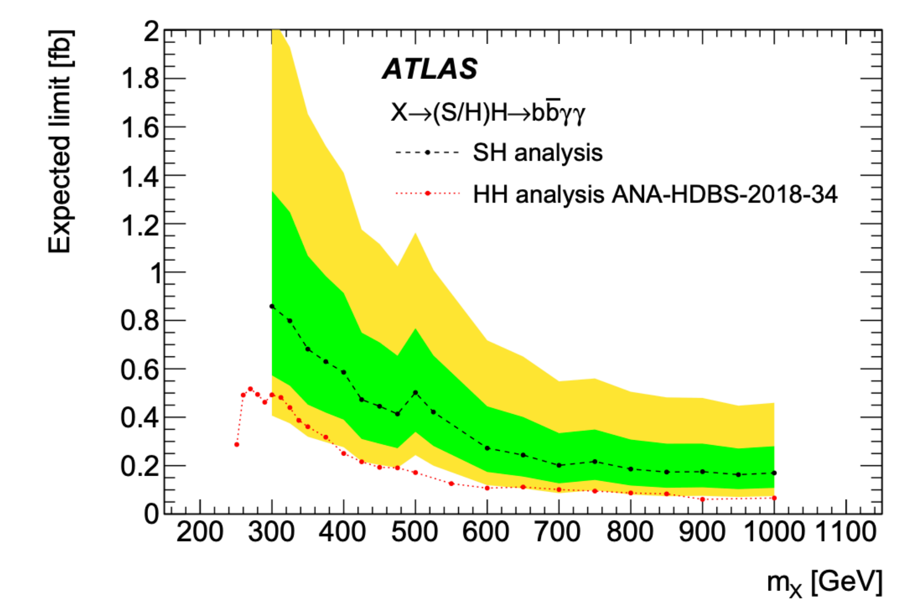
(a) m_X = 220 GeV, m_S = 90 GeV (b) m_X = 1000 GeV, m_S = 300 GeV

(c) m_X = 250 GeV, m_S = 15 GeV

Figure 54: Data-to-simulation comparison for the PNN score distributions for (a) m_X = 220 GeV, m_S = 90 GeV, (b) m_X = 1000 GeV, m_S = 530 GeVand (c) m_X = 250 GeV, m_S = 15 GeV in the $m_{\gamma\gamma}$ side-bands, with the statistical, theoretical and experimental uncertainties included.

Comparison with HH and CMS

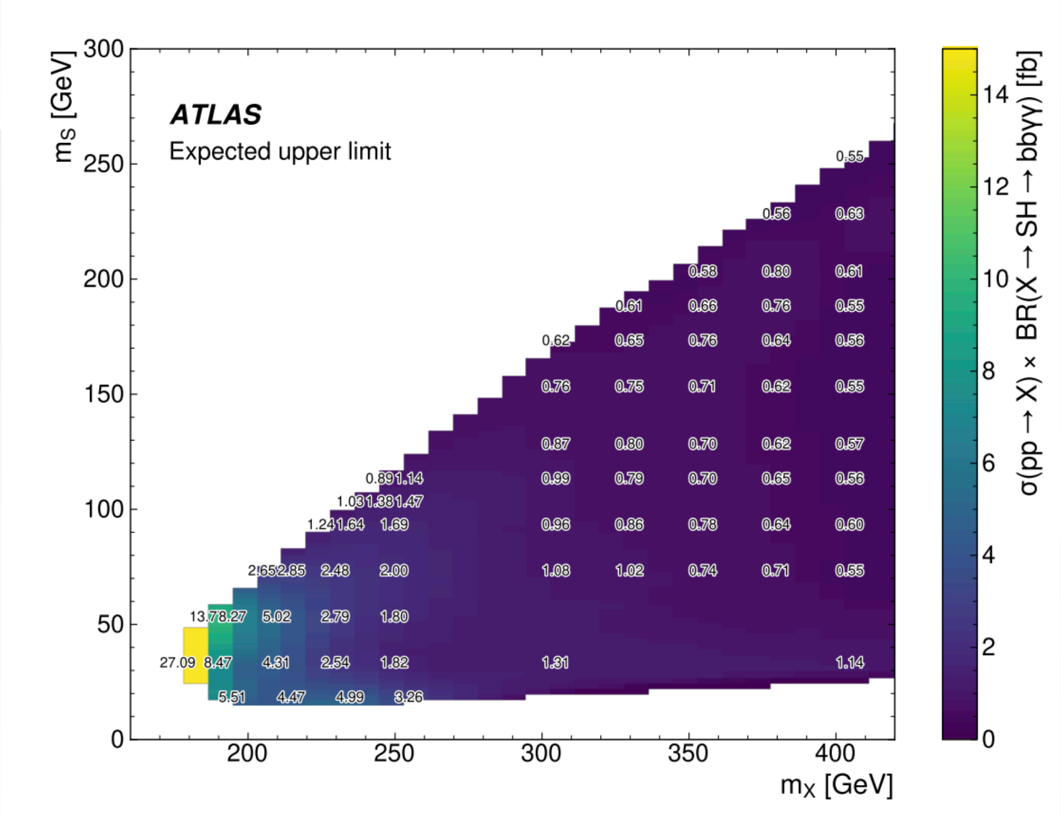
- The expected limit for mS = 125 GeV is compared to the one of the resonant HH search in Ref.

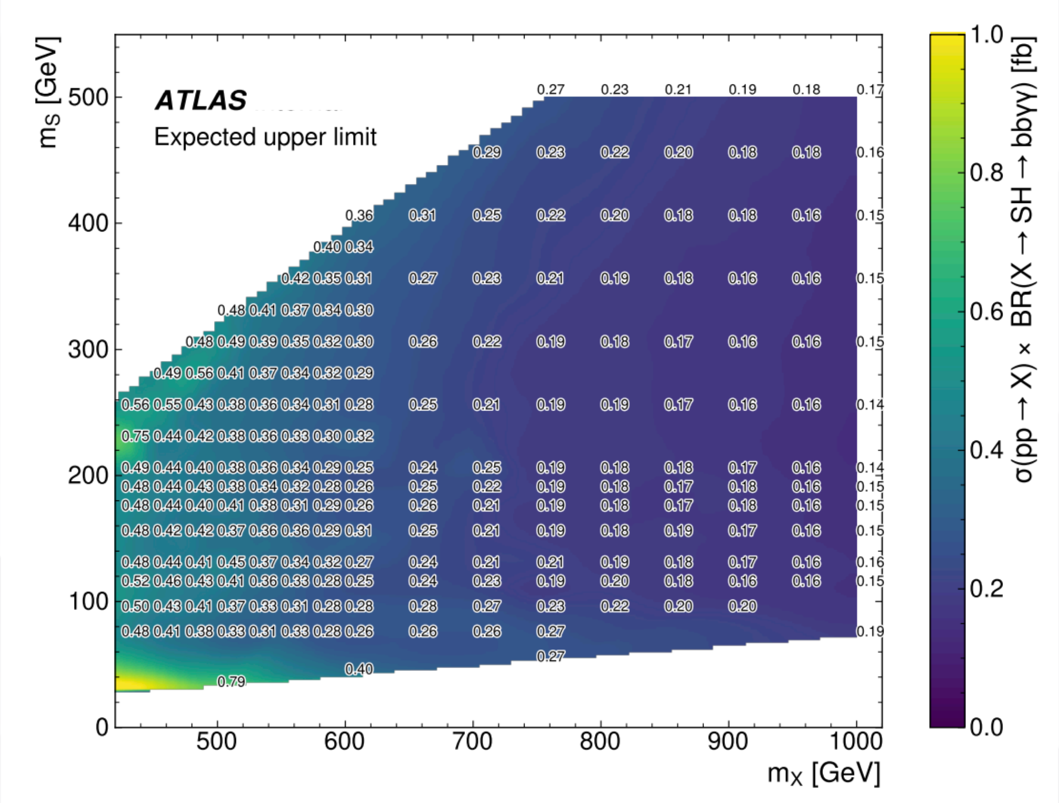


- CMS has reported an excess on $(m_X = 650\ GeV, m_S = 90\ GeV)$ with a local significance 3.8. No excess is found in our analysis.

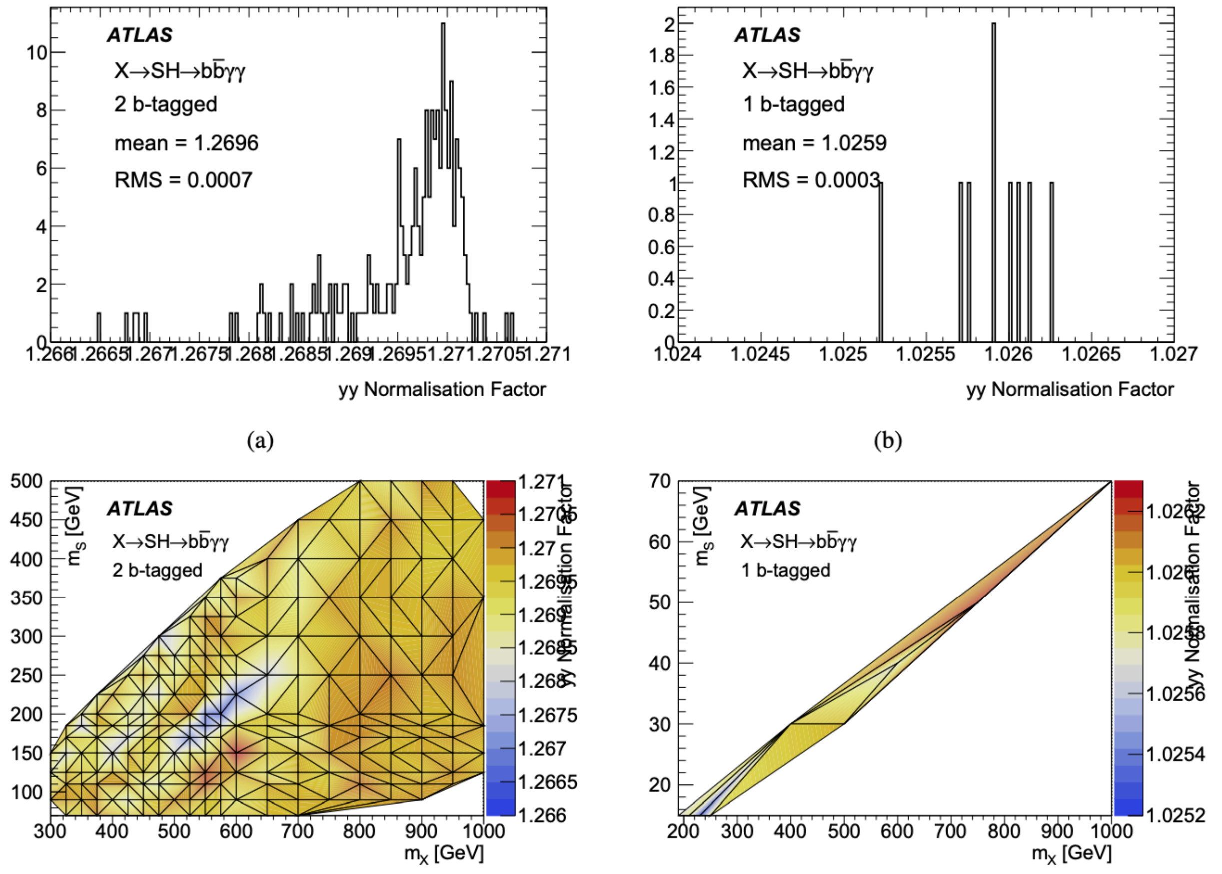
Signal template	mX	mS	p0	Z0	observed	-2	-1	expected	1	2
Interpolated	650	90	0.5	0.00	0.172	0.113	0.163	0.253	0.414	0.673
MC	650	90	0.5	0.00	0.142	0.094	0.135	0.211	0.345	0.559

Expected limits





Normalization factors



(a) (b) (c) (d)

Figure 66: Distribution of the Normalisation Factors for each point of the (a) 2 b-tagged and (b) 1 b-tagged categories. Normalisation Factors as a function of m_S and m_X of the (c) 2 b-tagged and (d) 1 b-tagged categories.

Number of events

Number of events for the different process categories obtained from a background-only fit to data in the signal regions and sidebands.

Background	2 b-tagged region Sideband	Signal region	Signal-like bin	1 b-tagged region Sideband	Signal region	Signal-like bin
Non-res. $\gamma\gamma$	1480 ± 37	372 ± 16	1.64 ± 0.37	13450 ± 110	3392 ± 53	2.45 ± 0.43
Single Higgs	0.46 ± 0.11	19.9 ± 5.3	0.04 ± 0.01	2.3 ± 1.1	92 ± 44	0.21 ± 0.10
ggF+$b\bar{b}H$	0.14 ± 0.11	6.5 ± 5.2	0.01 ± 0.01	1.5 ± 1.1	56 ± 43	0.11 ± 0.09
$t\bar{t}H$	0.21 ± 0.01	7.91 ± 0.77	0.01 ± 0.01	0.31 ± 0.01	11.4 ± 1.1	0.03 ± 0.01
ZH	0.08 ± 0.01	3.56 ± 0.30	0.02 ± 0.01	0.17 ± 0.01	7.35 ± 0.60	0.02 ± 0.01
Other	0.03 ± 0.01	1.94 ± 0.70	< 0.005	0.40 ± 0.23	17 ± 10	0.05 ± 0.03
Double Higgs	0.03 ± 0.01	1.65 ± 0.25	< 0.005	0.03 ± 0.01	1.79 ± 0.27	0.01 ± 0.01
Total	1480 ± 37	394 ± 16	1.67 ± 0.37	13450 ± 110	3486 ± 48	2.67 ± 0.45
Signal (m_X, m_S)						
(250, 100) GeV	0.38 ± 0.04	8.3 ± 1.2	1.43 ± 0.21			
(1000, 70) GeV				0.97 ± 0.10	33.3 ± 5.8	23.9 ± 4.2
Data	1479	395	0	13450	3491	4