Search for Higgs boson pairs production in the $b\bar{b}\gamma\gamma$ final state with ATLAS detector

Qiuping Shen

INPAC, Shanghai Jiao Tong University

The 8th China LHC Physics Workshop (CLHCP2022)

2022-11-24

Outline

- Why is $HH \rightarrow bb\gamma\gamma$?
- Object and event selection
- Data/MC comparison
- Signal and background parameterization
- Systematics uncertainties
- Results and summary

Why is $HH \rightarrow bb\gamma\gamma$?

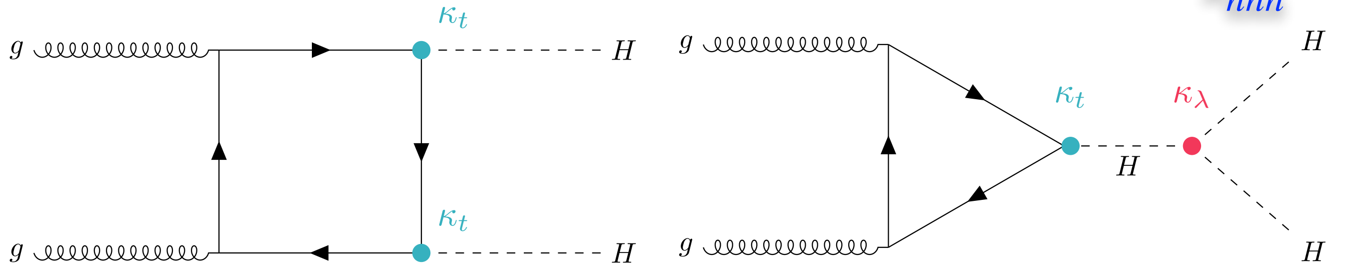
- HH production could directly access to the **trilinear Higgs self-coupling** (λ_{hhh}).

- Higgs to $b\bar{b}$: **Largest BR (59%)**
- Higgs to $\gamma\gamma$: **Small BR (0.2%)**
  - **Excellent photon trigger and resolution**
- **Main Backgrounds**:
  - **$\gamma\gamma$+jets and Single Higgs**
  - **Very clean final state**

- Full Run2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis
  - Glance, PRD

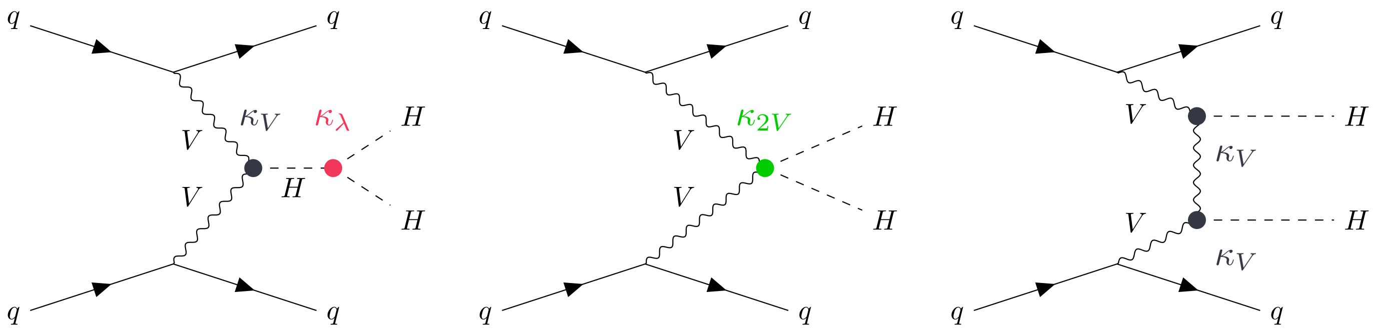
Gluon-Gluon Fusion (GGHH): LHCHWGHH

- σ_{NLO} = **31.02** [fb] @ 13TeV, mH=125.09 [GeV] ⇨ $\kappa_\lambda = \frac{\lambda_{hhh}}{\lambda_{hhh}^{SM}}$



Vector Boson Fusion (VBFHH):

- σ_{N3LO} = **1.723** [fb] @ 13TeV, mH=125.09 [GeV] ⇨ κ_λ, κ_{2V}

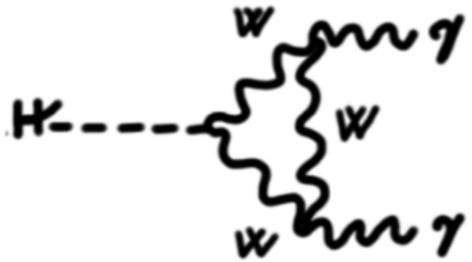


Object selection

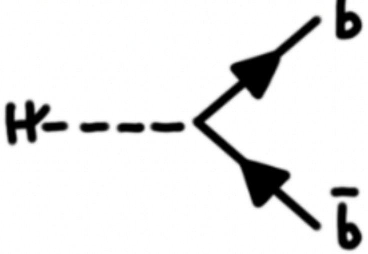
PhysRevD.106.052001

- A combination of **di-photon** and **single-photon** triggers are applied to maximize the efficiency.
  - Required **two loose or medium photons with (sub-) leading $pT > 35(25)$GeV**
  - Required **one loose photon with $pT > 120$ or 140 GeV**

- The **pre-selection** targeting the $HH \rightarrow b\bar{b}\gamma\gamma$ signature to define signal region of the analysis.



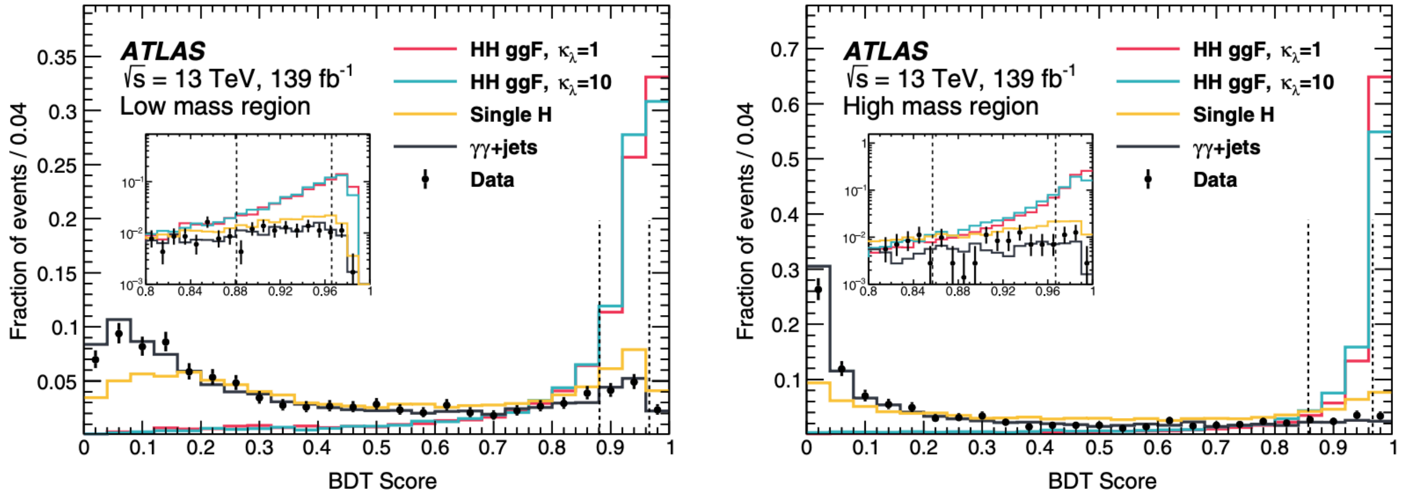
✓ Two **tight** and **isolated** photons
✓ (Sub-) leading $pT/m_{\gamma\gamma}$>0.35(0.25)
✓ Di-photon **invariant mass window** $105 < m_{\gamma\gamma} < 160$ GeV.



✓ **Exactly** two b-jets passing the 77% efficiency WP for DL1r.
✓ The b-jets candidates selected by **ranking them by b-tag scores** they pass and **tie breaking by pT.**

✓ **No leptons.**
  ✓ To suppress $t\bar{t}H$ background which decays electrons or muons
✓ **At least two jets**
✓ **Less than 6 central jets**
  ✓ To control $t\bar{t}H$ background which decays hadronically

- Common pre-selection applied
- Two regions defined by $m^*_{b\bar{b}\gamma\gamma}$ variable:
  - **High mass region**: $m^*_{b\bar{b}\gamma\gamma} \geq 350GeV$
    - Target **SM HH**
  - **Low mass region**: $m^*_{b\bar{b}\gamma\gamma} < 350GeV$
    - Target **BSM HH**



(a) Low mass region (b) High mass region

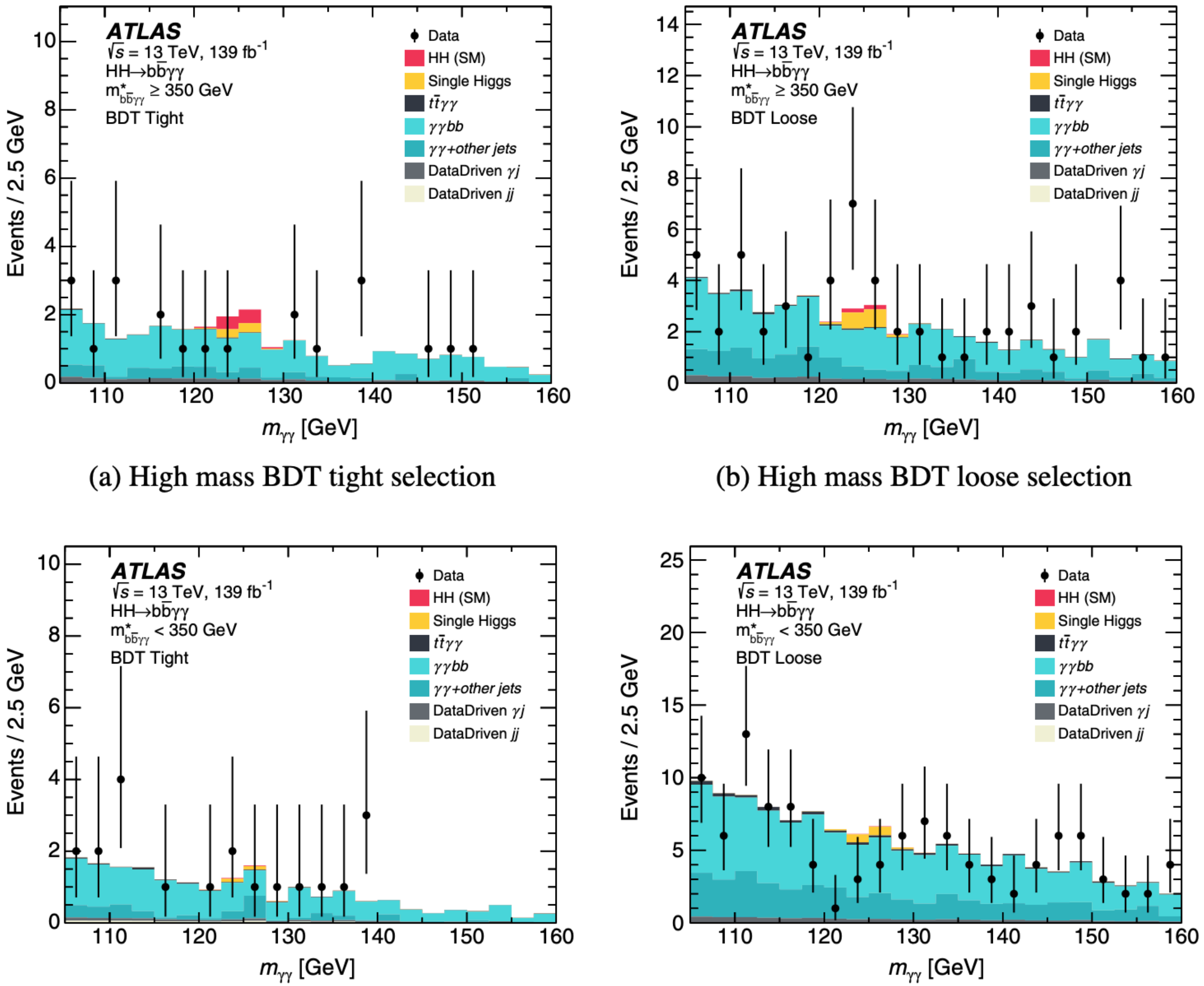
- **XGBoost** method used to discriminate
  - the **benchmark HH signals**
  - $\gamma\gamma, t\bar{t}\gamma\gamma$, **single Higgs**

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 $\in [0.881, 0.966]$

- Boundaries of the categories:
  - **Maximizing the combined counting significance** at **signal region** $m_{\gamma\gamma}$ **[120,130] GeV**

Data/MC comparison

- Distributions of $m_{\gamma\gamma}$ in all categories
- The **data-driven fractions of**
  - $\gamma\gamma, \gamma j$ **and di-jet background** are applied.
- Total background normalized to data sideband
- **Consistent within the margin of uncertaintie**



(a) High mass BDT tight selection

(b) High mass BDT loose selection

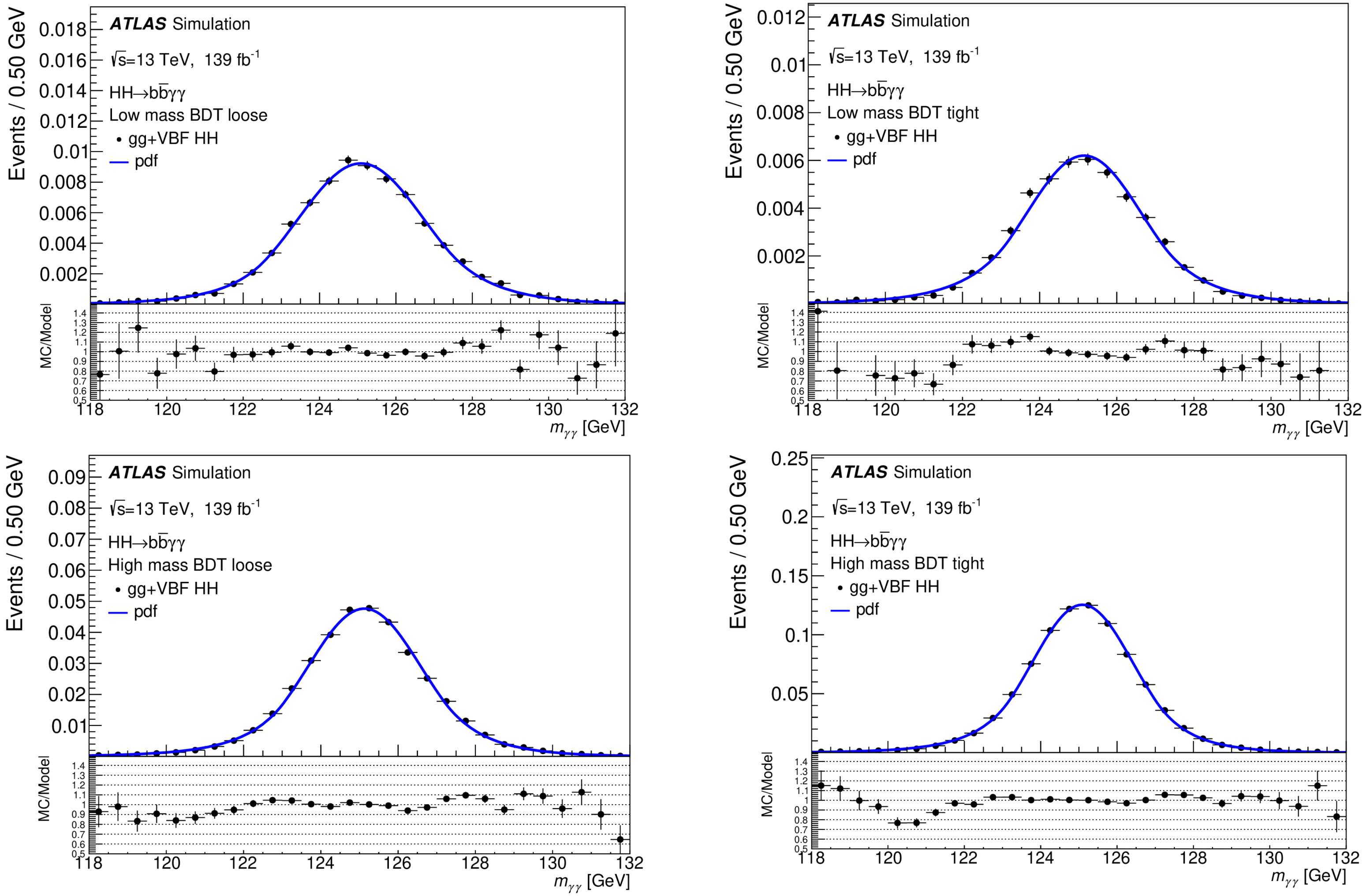
(c) Low mass BDT tight selection

(d) Low mass BDT loose selection

Signal parameterization

PhysRevD.106.052001

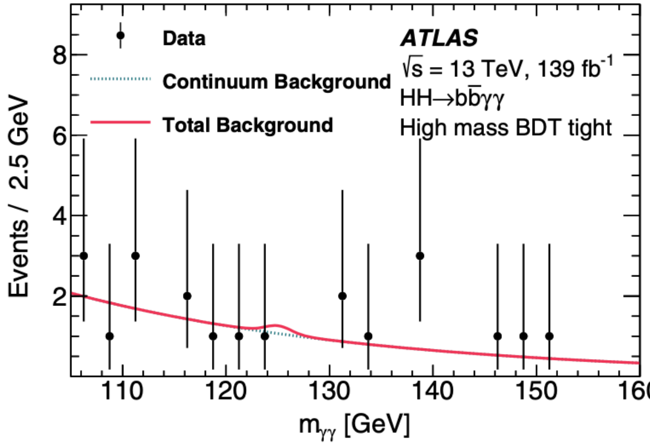
- The **HH signals** are modeled in the $m_{\gamma\gamma}$ spectrum by a **double-sided crystal ball** **(DSCB)** function.
  - The parameters of the DSCB obtained by fitting on the **SM ggF HH + VBF HH samples**.
  - The **same parameterized functions** are also used for modeling the **single Higgs process**



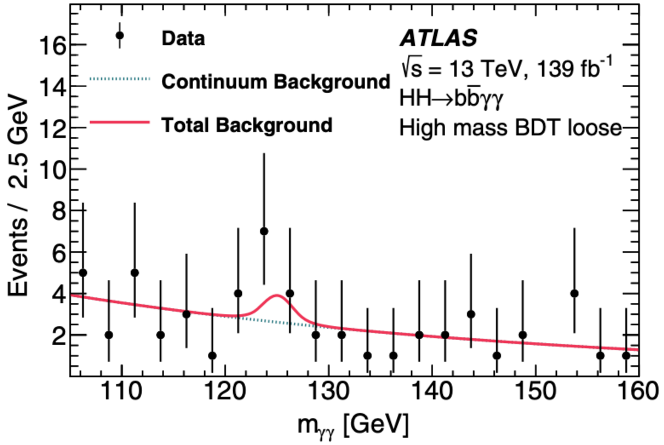
Background parameterization

PhysRevD.106.052001

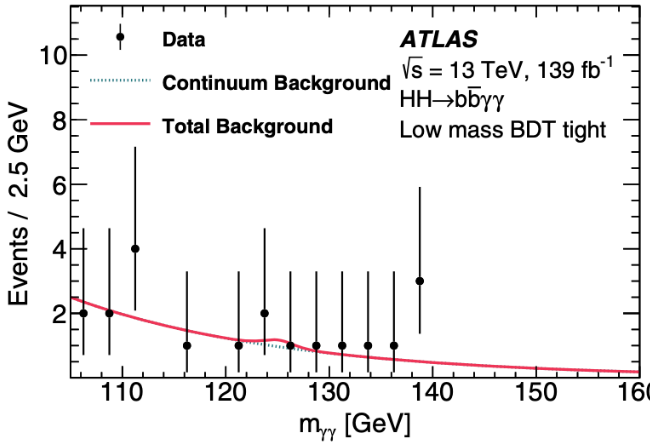
- The **shape of continuum background** is set by fitting
  - **smoothly falling analytic function to data sidebands.**
  - Functional form chosen by fitting MC background
  - The **exponential function: exp($a \cdot m_{\gamma\gamma}$)**
    - Smallest number of degrees of freedom
    - Yields a consistently small bias



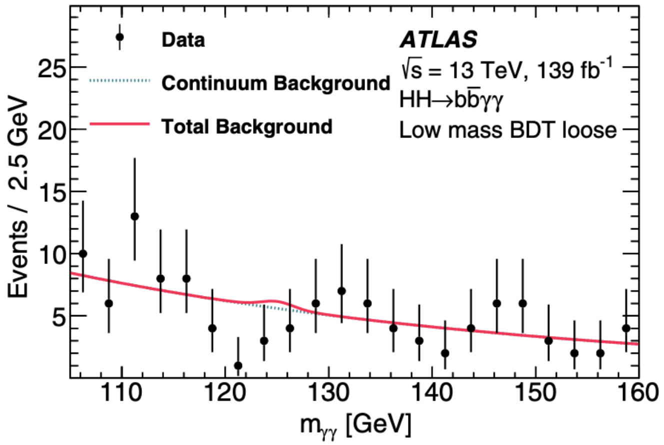
(a) High mass BDT tight



(b) High mass BDT loose



(c) Low mass BDT tight



(d) Low mass BDT loose

Systematic uncertainties

PhysRevD.106.052001

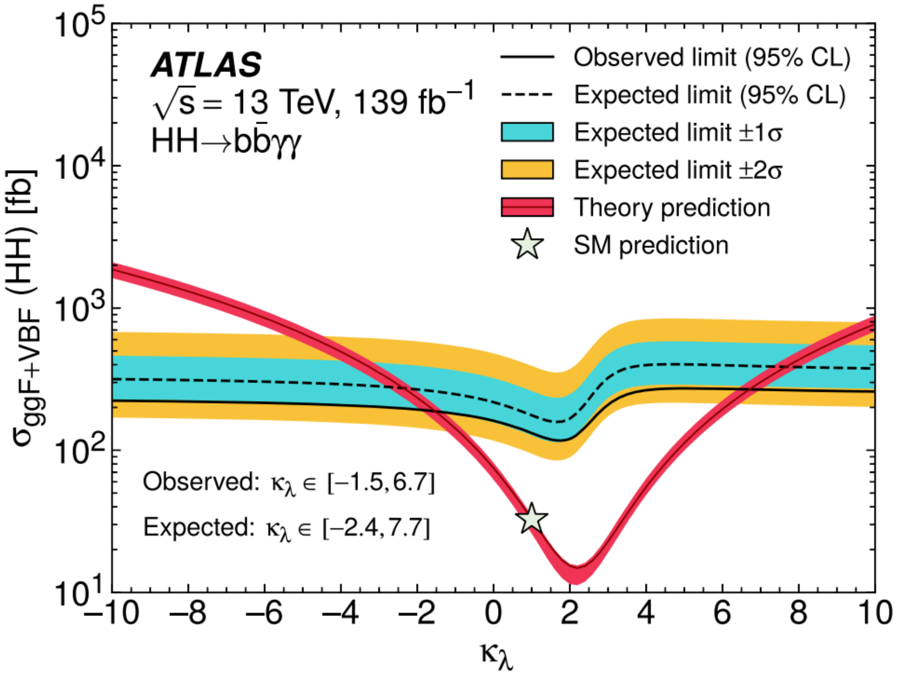
- **Sensitivity** of the analysis **limited** by the **statistical precision**.

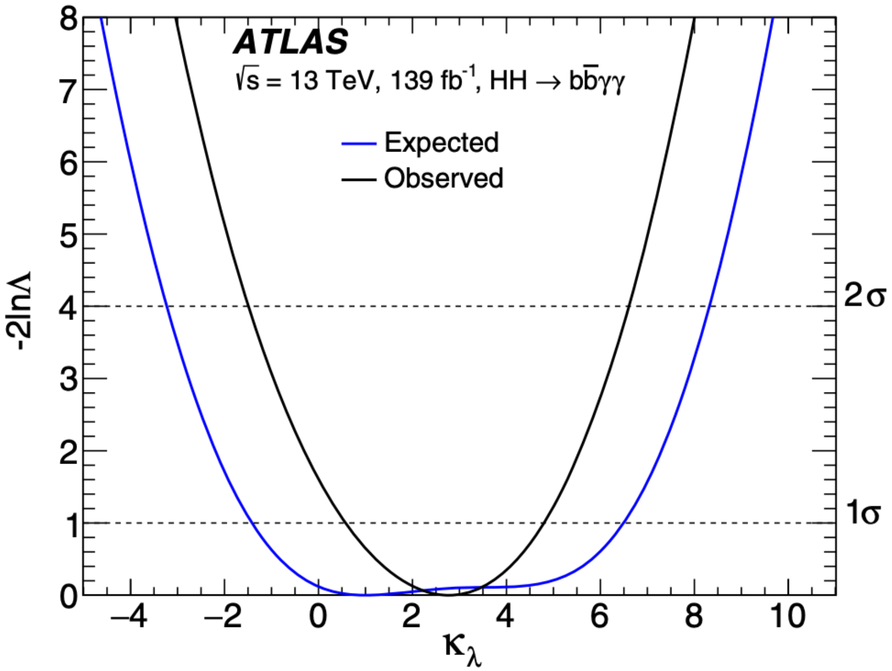
- The table showing the dominant systematic uncertainties
  - The **background functional form bias** assessed as an **additional uncertainty** in **total number of signal events** in each category. (**Spurious Signal**)
  - **The parton showering model**
  - **The photon energy resolution**

Relative impact of the systematic uncertainties [%]

Source	Type	Nonresonant analysis HH
Experimental		
Photon energy resolution	Norm. + Shape	0.4
Jet energy scale and resolution	Normalization	< 0.2
Flavor tagging	Normalization	< 0.2
Theoretical		
Factorization and renormalization scale	Normalization	0.3
Parton showering model	Norm. + Shape	0.6
Heavy-flavor content	Normalization	0.3
$\mathcal{B}(H \to \gamma\gamma, b\bar{b})$	Normalization	0.2
Spurious signal	Normalization	3.0

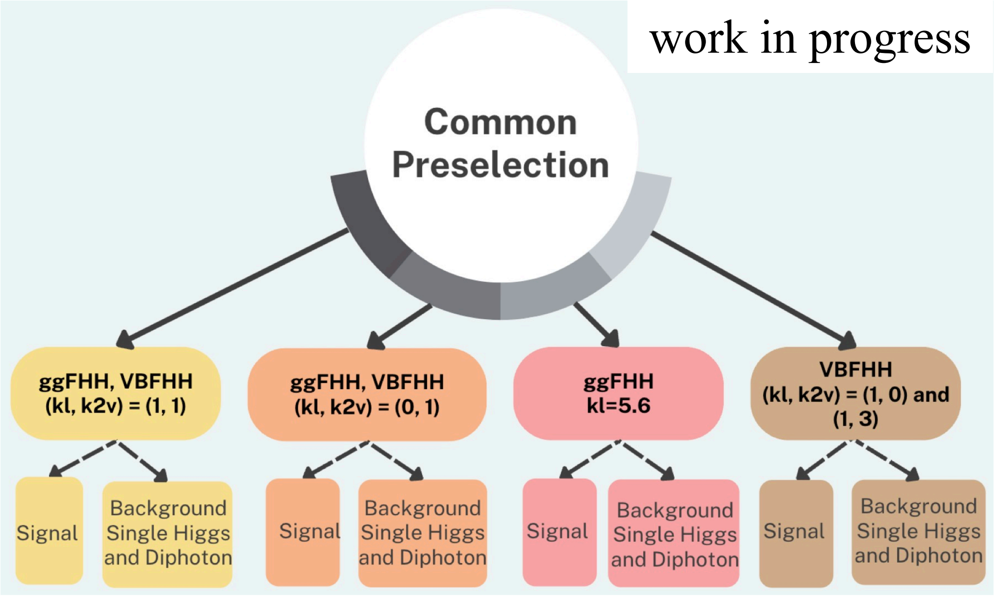
- Upper limits on HH production cross section calculated as a function of κ_λ.
  - The observed(expected) limit: 4.2(5.7) times cross section of the standard model
  - The **expected** constraints on κ_λ **at 95%CL: [-2.4, 7.7]**, whereas the **observed** constraints are **[-1.5, 6.7]**.
- The **best-fit value of κ_λ and its uncertainty** obtained by means of a **negative log-likelihood scan**
  - Best-fit value: $\kappa_\lambda = 2.8^{+2.0}_{-2.2}(^{+3.8}_{-4.3})$ for $1\sigma(2\sigma)$ confidence interval
  - Expected values: $\kappa_\lambda = 1.0^{+5.5}_{-2.4}(^{+7.3}_{-4.2})$ for $1\sigma(2\sigma)$ confidence interval





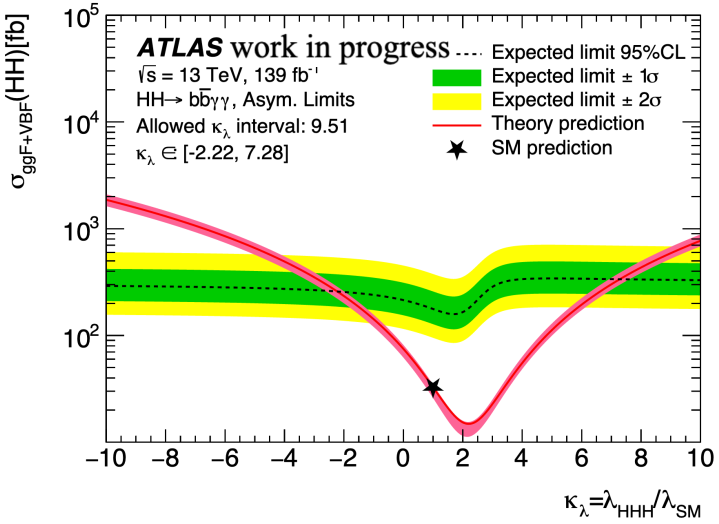
New strategy of event selection

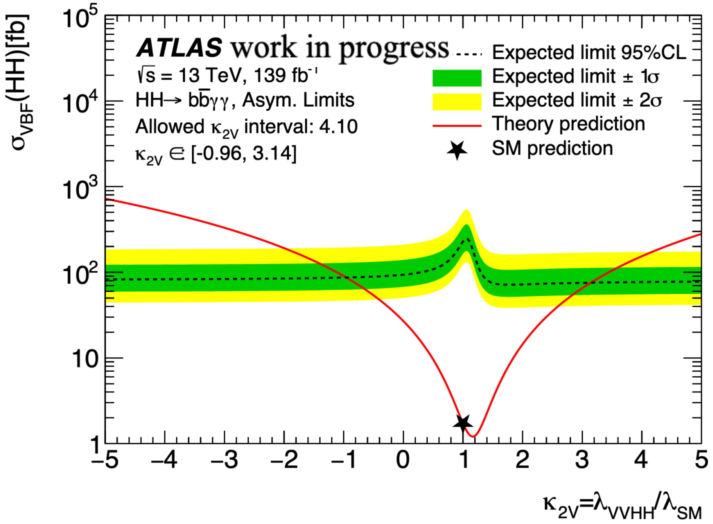
- **Ongoing Run2 legacy analysis**: a dedicated analysis for **ggFHH and VBFHH**
  - Target on **upper limit on HH**, and κ_λ **and** κ_{2V} **constraints**



- Current new strategy with XGBoost (Work in progress):
  - Following the common preselection
  - Two layers designed to separate signals
    - **Multi-class: target various κ_λ and κ_{2V} signal modes**
    - **Binary-class: target HH and backgrounds**
  - Preliminary statistical expected results (stats only):

	Results
SM HH Upper Limit	5.41
SM VBFHH Upper Limit	132.49
KL constraint (length)	[-2.22, 7.28] (9.51)
K2V constraint (length)	[-0.96, 3.14] (4.10)





- The method showing promising results
- **Still under development**

Summary and outlook

- The published paper (PhysRevD.106.052001) gives upper limit on HH cross section and κ_λ constraints.
  - **NJU and IHEP groups** contributed VBF-category, κ_λ-reweighting, background decomposition and data/MC studies.
  - **SJTU group** contributed VBF studies, sample production, κ_λ-reweighting and scan, limits scan.

- The **legacy analysis ongoing**:
  - Target on **upper limit on HH and VBFHH**, and κ_λ **and** κ_{2V} **constraints**
  - New strategy based on XGBoost applied to do the event selection
  - **Multi-class training strategy** and showing promising preliminary results.
- Let's looking forward to the more fruitful results from legacy analysis
- **Stay tuned for next Run3 analysis!!!**

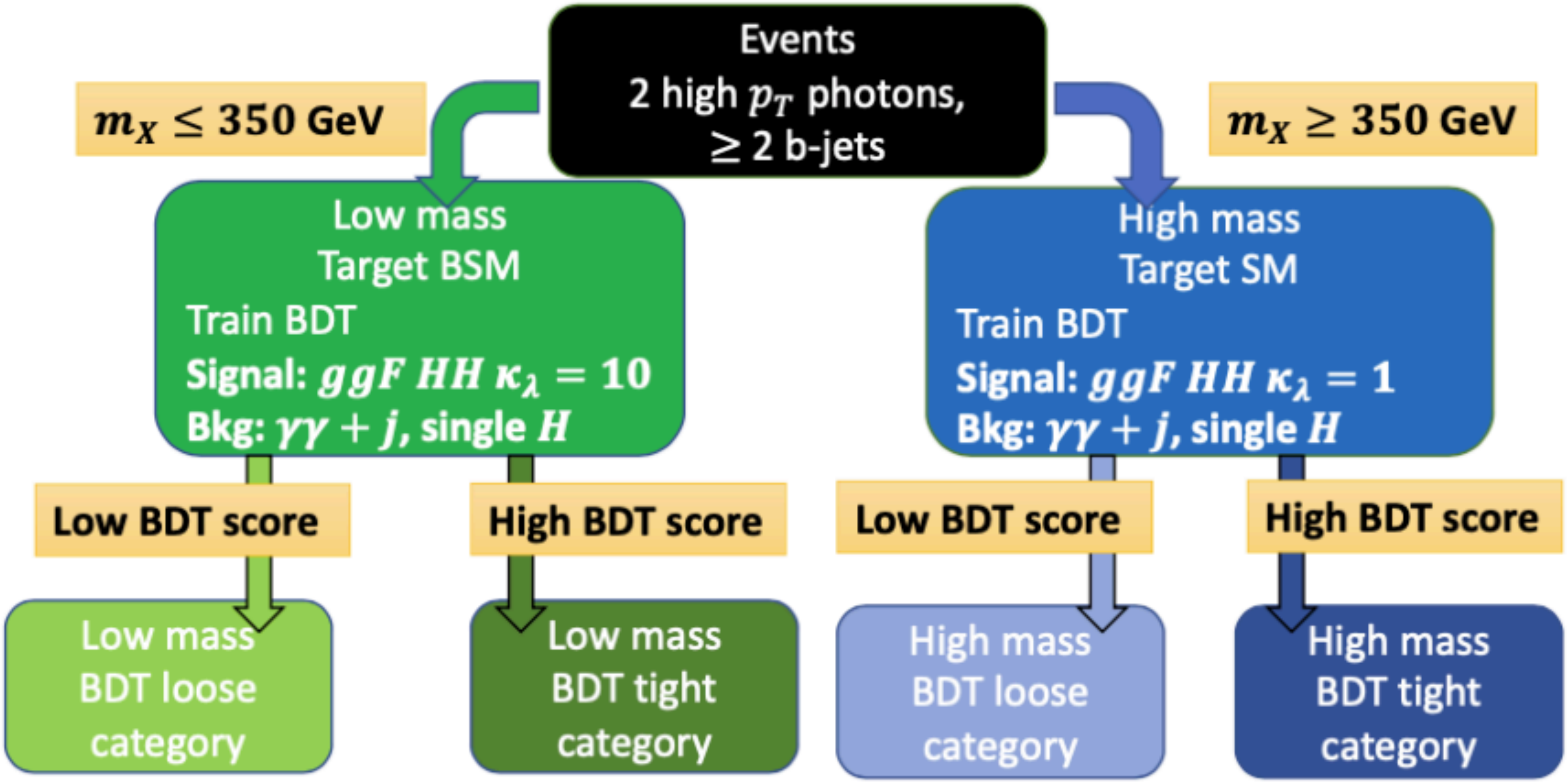
Thanks for your attention!

Email Address:
qiuping.shen@cern.ch
shen@apc.in2p3.fr

2022-11-24

Backup Slides

Event selection strategy



Data and MC samples

- Monte Carlo samples

Process	Generator	PDF set	Showering	Tune
Nonresonant ggF HH	POWHEG BOXv2+FT [41,42,43]	PDFLHC [44]	PYTHIA8.2 [67]	A14 [68]
Nonresonant VBF HH	MadGraph5_aMC@NLO [47]	NNPDF3.0NLO [69]	PYTHIA8.2	A14
Resonant ggF HH	MadGraph5_aMC@NLO	NNPDF2.3LO	HERWIG7.1.3 [53,54]	H7.1—Default [70]
ggF H	NNLOPS [71–73] [74,75]	PDFLHC	PYTHIA8.2	AZNLO [76]
VBF H	POWHEG BOXv2 [41,72,77–83]	PDFLHC	PYTHIA8.2	AZNLO
WH	POWHEG BOXv2	PDFLHC	PYTHIA8.2	AZNLO
$qq \to ZH$	POWHEG BOXv2	PDFLHC	PYTHIA8.2	AZNLO
$gg \to ZH$	POWHEG BOXv2	PDFLHC	PYTHIA8.2	AZNLO
$t\bar{t}H$	POWHEG BOXv2 [78–80,83,84]	NNPDF3.0NLO	PYTHIA8.2	A14
bbH	POWHEG BOXv2	NNPDF3.0NLO	PYTHIA8.2	A14
tHq	MadGraph5_aMC@NLO	NNPDF3.0NLO	PYTHIA8.2	A14
tHW	MadGraph5_aMC@NLO	NNPDF3.0NLO	PYTHIA8.2	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	PYTHIA8.2	...

BDT Variables

- Variables used in the BDT

TABLE II. Variables used in the BDT for the nonresonant analysis. All vectors in the event are rotated so that the leading photon ϕ is equal to zero, while their relative azimuthal angular differences are kept unchanged.

Variable	Definition
Photon-related kinematic variables	
$p_T/m_{\gamma\gamma}$	Transverse momentum of each of the two photons divided by the diphoton invariant mass $m_{\gamma\gamma}$
η and ϕ	Pseudorapidity and azimuthal angle of the leading and subleading photon
Jet-related kinematic variables	
b-tag status	Tightest fixed b-tag working point (60%, 70%, or 77%) that the jet passes
p_T, η and ϕ	Transverse momentum, pseudorapidity and azimuthal angle of the two jets with the highest b-tagging score
${p_T}^{b\bar{b}}$, $\eta_{b\bar{b}}$ and $\phi_{b\bar{b}}$	Transverse momentum, pseudorapidity and azimuthal angle of the b-tagged jets system
$m_{b\bar{b}}$	Invariant mass of the two jets with the highest b-tagging score
H_T	Scalar sum of the p_T of the jets in the event
Single topness	For the definition, see Eq. (1)
Missing transverse momentum variables	
E_T^{miss} and ϕ^{miss}	Missing transverse momentum and its azimuthal angle

Data and MC sample

- **Data:**
  - This analysis relies on the **full Run2 dataset** => **139fb^{-1}**

More details about the signal and background in backup slides

- **MC samples**:

- **Signals**:
  - **ggf HH** samples at NLO
    - Nominal samples produced by Powheg + Pythia8
    - With $\kappa_\lambda = 1$ **(SM)** and $\kappa_\lambda = 10$
  - **vbf HH** samples at LO
    - Nominal samples using MadGraph + Pythia8
    - **SM samples** + **12 samples** with BSM values for the coupling modifiers κ_λ, κ_{2V} and κ_V

- **Backgrounds**:
  - **Single Higgs** samples including all the production modes $ggH,\ VBFH,\ WH,\ qq \to ZH,\ gg \to ZH,\ t\bar{t}H, tHjb,\ tWH,\ b\bar{b}H$
  - **Sherpa2 $\gamma\gamma$+jets** MC sample
  - $t\bar{t}\gamma\gamma$ MC samples based on aMG@NLO + Pythia8
  - The continuum background modeling is data-driven