Analysis of non-resonant HH → $b\bar{b}\gamma\gamma$ using CMS Run3 data

Chuxue YANG* (Peking University)

Under supervision of Yajun Mao, Chen Zhou, Alexandra Carvalho Antunes De Oliveira

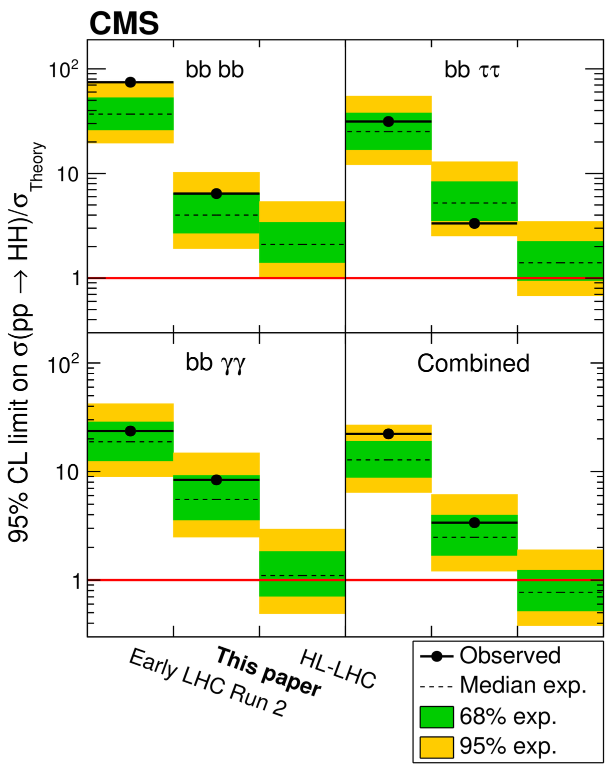
Introduction

Why Study HH?

- Tests the shape of the Higgs potential, still unprobed after Higgs discovery.
- Directly sensitive to the Higgs self-coupling, a key SM prediction.
- Probes BSM effects in the Higgs sector.

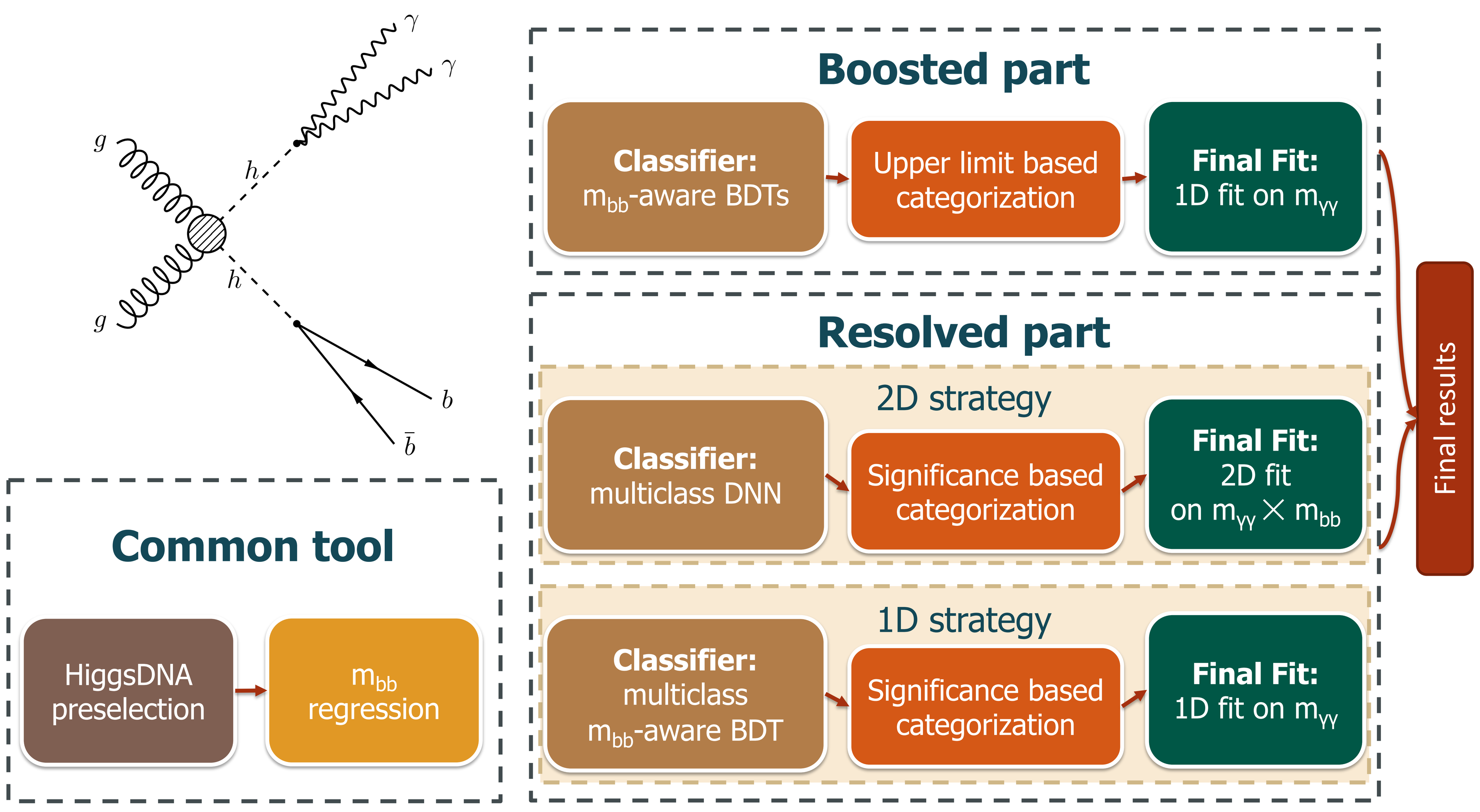
Why the $b\bar{b}\gamma\gamma$ final state?

- **H→$b\bar{b}$**: largest branching ratio among Higgs decays.
- **H→γγ**: clean final state with excellent mass resolution and trigger efficiency.
- This channel provides **a powerful balance of rate and resolution**, ideal for early Run 3 studies.



Analysis Overview

- Dataset: **61.9 fb⁻¹** in **2022-2023** @13.6TeV



Event Selection

Preselection

- Events are triggered using the dedicated **H→γγ HLT path**.
- **The photon selection** follows the H→γγ group preselection.
- **A loose dijet preselection** is applied on the H→bb dijet pair:
  - Specific enough to focus on the interesting phase space;
  - Generic enough to retain sufficient statistics for downstream MVA training.
- A tight lepton selection is used to implement **loose lepton vetoes**.
- An additional technique applied already at the preselection stage:

m_{bb} regression

- Improving the m_{bb} resolution enhances both event selection and signal extraction.
- Two regression methods are used:
  - **PNet-based regression** (centrally provided);
  - **Custom DNN regression** (trained on ggHH signal):
    - Trained to predict (gen − reco) / reco using jet- and event-level inputs.
- **10–20% m_{bb} resolution improvement**, varying with data-taking period.

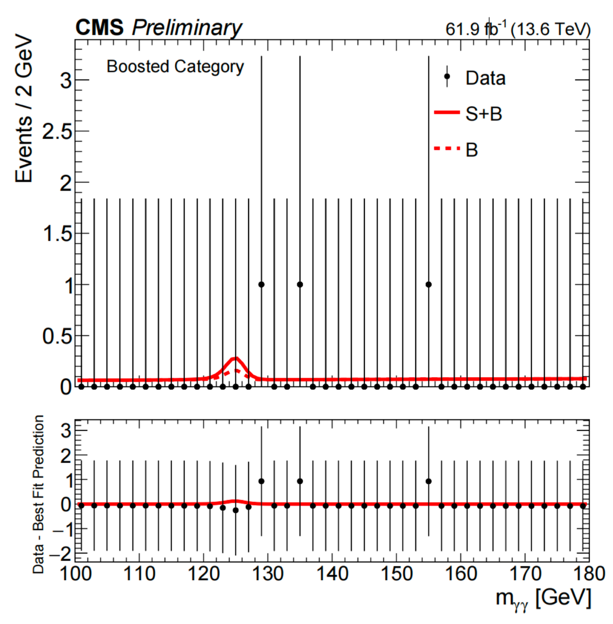
Event Categorization

Resolved Category – MVA-Based Event Selection

- The final selection relies on multiclass MVA discriminators to separate ggHH signal from backgrounds.
- Two complementary strategies are developed, differing in how m_{bb} is handled:
  - **1D Strategy**
    - m_{bb} is used as an input to the BDT, fit performed on **$m_{\gamma\gamma}$ only**;
    - Training optimized to suppress single-Higgs backgrounds resonant in $m_{\gamma\gamma}$;
    - Achieves ~10³ (10²) background rejection for QCD (top quark) at ~60% (50%) signal efficiency.
  - **2D Strategy**
    - DNN trained to be agnostic to the m_{bb} mass scale, m_{bb} is used explicitly in the 2D fit: **$m_{\gamma\gamma} \otimes m_{bb}$**;
    - Designed to suppress all backgrounds, including QCD and top quark;
    - Achieves ~10³ (10²) background rejection for QCD (top quark) at ~50% (55%) signal efficiency.
- Both strategies provide **strong signal vs. background separation.**
- Use different assumptions and statistical treatments, yet give consistent results.

Boosted Category – Targeting High-pT H→bb Events

- At high pT, AK4 jets fail to resolve close b quarks ⇒ Loss in signal efficiency.
  - AK8 jets recover the efficiency, motivating a separate boosted category.
- **Preselection** (from gen-level studies):
  - AK8 jet with:
    - ≥ 2 subjets;
    - $\tau_2/\tau_1 < 0.75$;
    - PNet(Xbb vs. QCD) > 0.4;
    - 30 < msoftdrop < 210 GeV.
  - Candidate: AK8 jet with highest PNet(Xbb vs. QCD).
- **Final selection** (by a two-step BDT approach):
  - BDT1: SM ggHH vs. VH;
  - BDT2: SM ggHH vs. all background:
    - using BDT1 score as input.
- → Achieves strong background rejection.
- **1 Signal region defined** by optimizing expected limit in cut-and-count.



Combination of Resolved & Boosted Categories

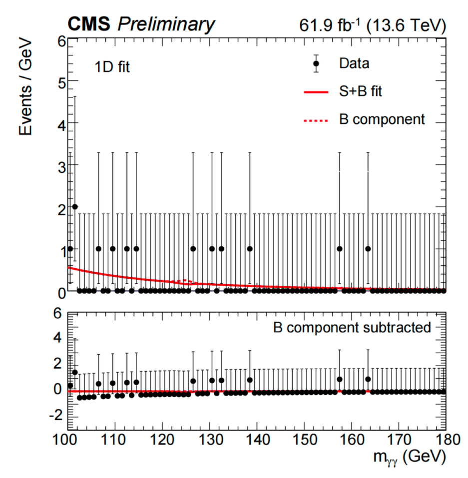
- Events are first assigned to the boosted category, which has high purity and low background;
- Remaining events are used to build the resolved category, maximizing sensitivity.

Fitting Strategy

We use two signal extraction strategies:

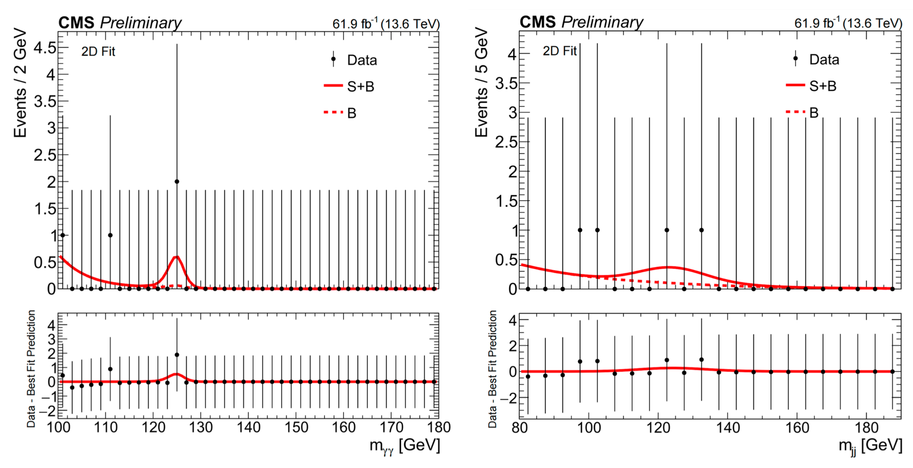
1D Fit (Resolved & Boosted Categories)

- Fit on **$m_{\gamma\gamma}$** only;
- Signal & resonant backgrounds: sum of Gaussians or DCB;
- Nonresonant background: discrete profiling over smooth analytic functions.



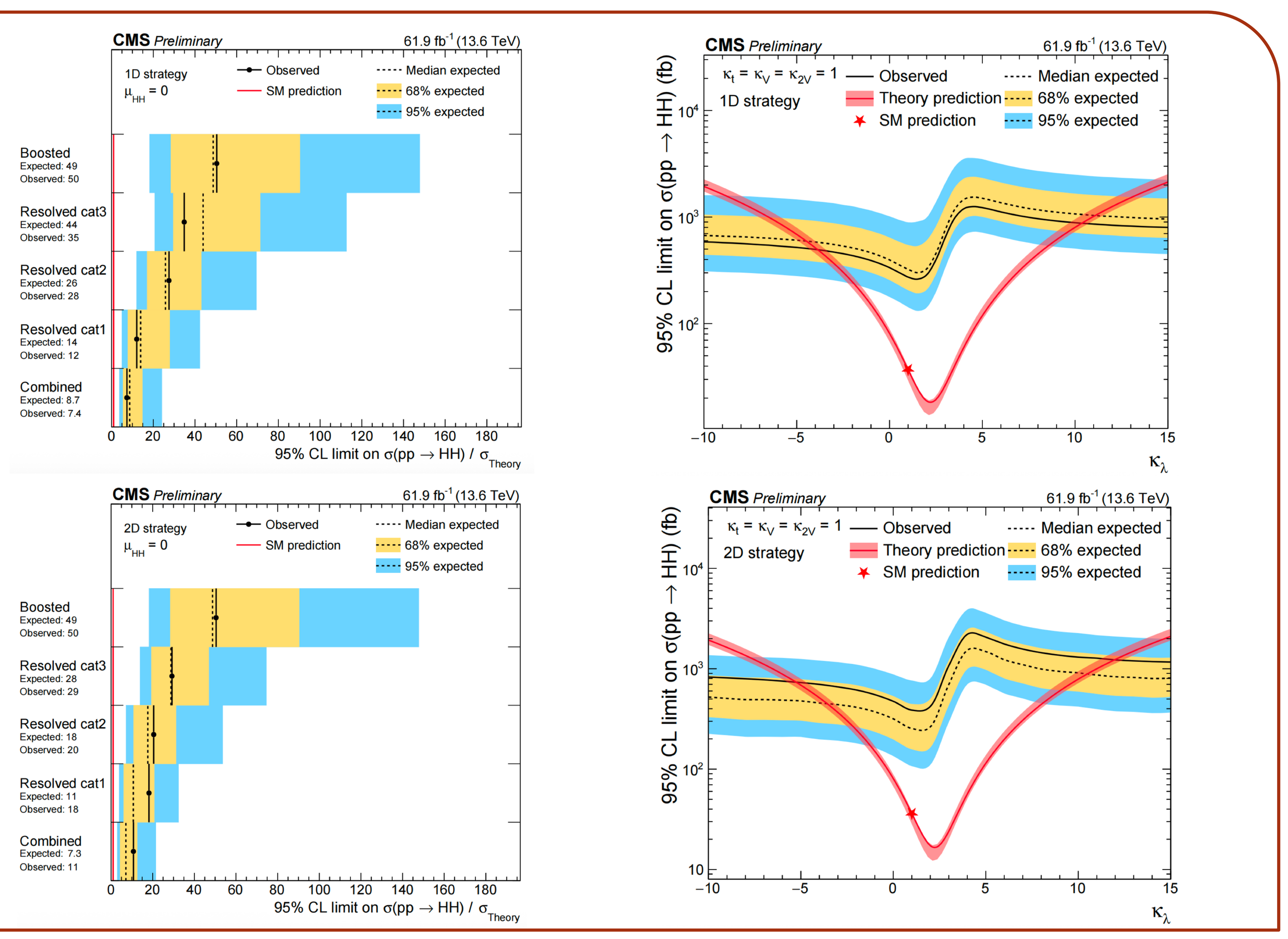
2D Fit (Resolved only)

- Simultaneous fit on **($m_{\gamma\gamma}$, m_{bb})**;
- Signal: sum of Gaussians;
- Resonant backgrounds: DCB for $m_{\gamma\gamma}$; DCB or a combination of falling and peaking functions for m_{bb}.
- Nonresonant background: discrete profiling as in 1D.



Results and Conclusion

- **Search for non-resonant HH** → **bbγγ** using **61.9 fb⁻¹** of Run 3 data (2022–2023, $\sqrt{s}$ = 13.6 TeV).
- Two complementary fit strategies are employed: 1D (mγγ) and 2D (mγγ ⊗ mbb).
  - Observed (expected) 95% CL limits: **7.4 (8.7)** × σ_{SM} (1D), **11.0 (7.3)** × σ_{SM} (2D).
  - Constraints on the Higgs self-coupling: **$-3.9 < \kappa_\lambda < 10.4$ (1D), $-5.0 < \kappa_\lambda < 12.0$ (2D)**.
  - Both strategies give **consistent and robust results**.
  - **The boosted events** improve sensitivity by ~5%.
- The analysis is **statistics-limited.**
- Dominant uncertainties come from **background normalization**, **signal theory**, and **heavy-flavor modeling** in ggH.
- **A bias uncertainty** is included in the 2D fit to cover $m_{\gamma\gamma}$–m_{bb} correlations.
- The analysis demonstrates the maturity of the CMS HH → bbγγ framework and its readiness for the full Run 3 dataset.



Reference: [1] The CMS Collaboration. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature* **607**, 60–68 (2022). https://doi.org/10.1038/s41586-022-04892-x

CMS-PAS-HIG-25-007

*chuxue.yang@cern.ch

CLHCP2025 The 11th China LHC Physics Conference
Oct. 29 – Nov. 3, 2025 · Xinxiang, Henan, China