Prospects for HH measurements at the HL-LHC
Overview of CMS projection status

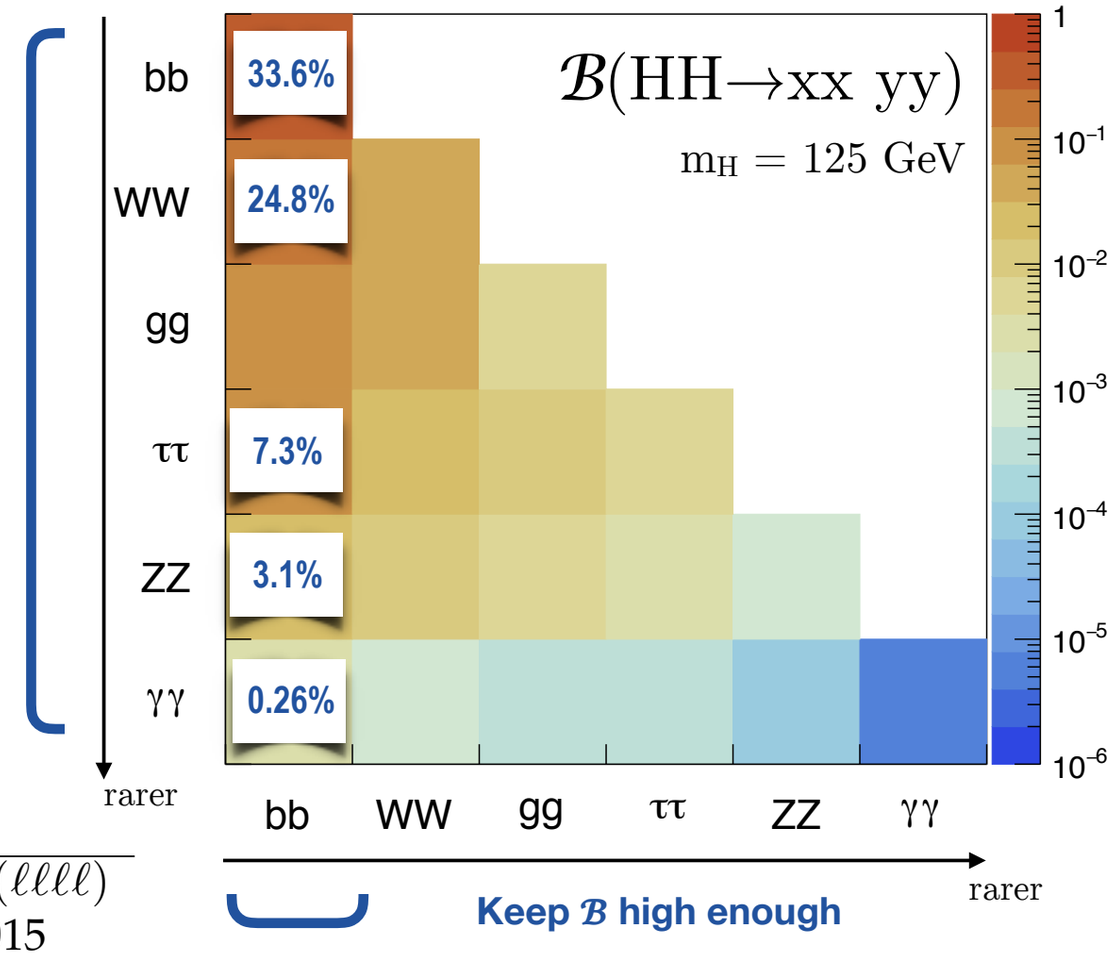
Luca Cadamuro for the CMS HH Yellow Report team

A. Benaglia, M. Bengala, O. Bondu, L. Borgonovi, S. Braibant, L. Cadamuro, A. Carvalho, C. Delaere, N. de Filippis, E. Fontanesi, M. Gallinaro, M. Gouzevitch, J. R. Komaragiri, D. Majumder, K. Mazumdar, F. Monti, G. Ortona, L. Panwar, N. Sahoo, R. Santo, G. Strong, M. Vidal, S. Wertz

Higgs @ HL/HE-LHC - Autumn WG2 Meeting, October 22nd, 2018

Overview

- **Goal:** Establish the HH CMS sensitivity at the HL-LHC
  - exercise of projecting Run II analysis done for ECFA (2.3 fb^{-1}) and detector Technical Design Report
  - extrapolation of Run 2 analyses does not allows to study new channels or new ideas to better exploit the HL-LHC potential
  - $\Longrightarrow$ develop full analysis using a parametrised CMS detector response with Delphes
- **How:** Five analyses and combination
  - bb$\gamma\gamma$
  - bbWW($2\ell 2\nu$)
  - bbbb
  - bb$\tau\tau$
  - bbZZ(4ℓ)



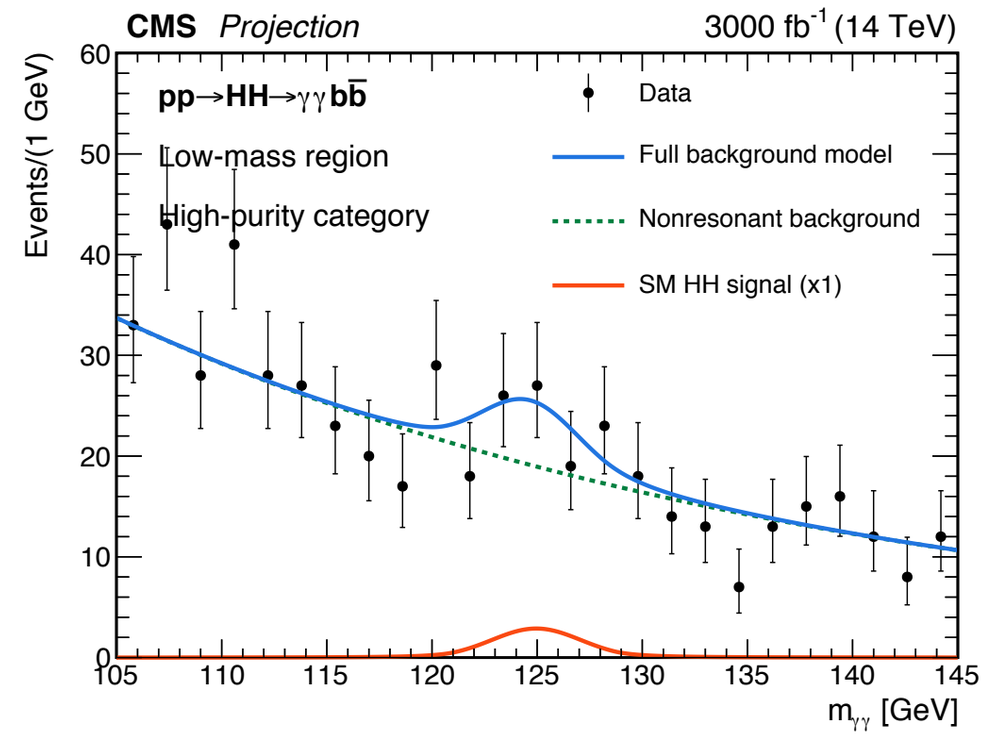
Channel	bbbb	bb$\tau\tau$	bbWW($\ell\nu\ell\nu$)	bb$\gamma\gamma$	bbZZ($\ell\ell\ell\ell$)
$\mathcal{B}$ [%]	33.6	7.3	1.7	0.26	0.015
Number of events	37000	8000	1830	290	17

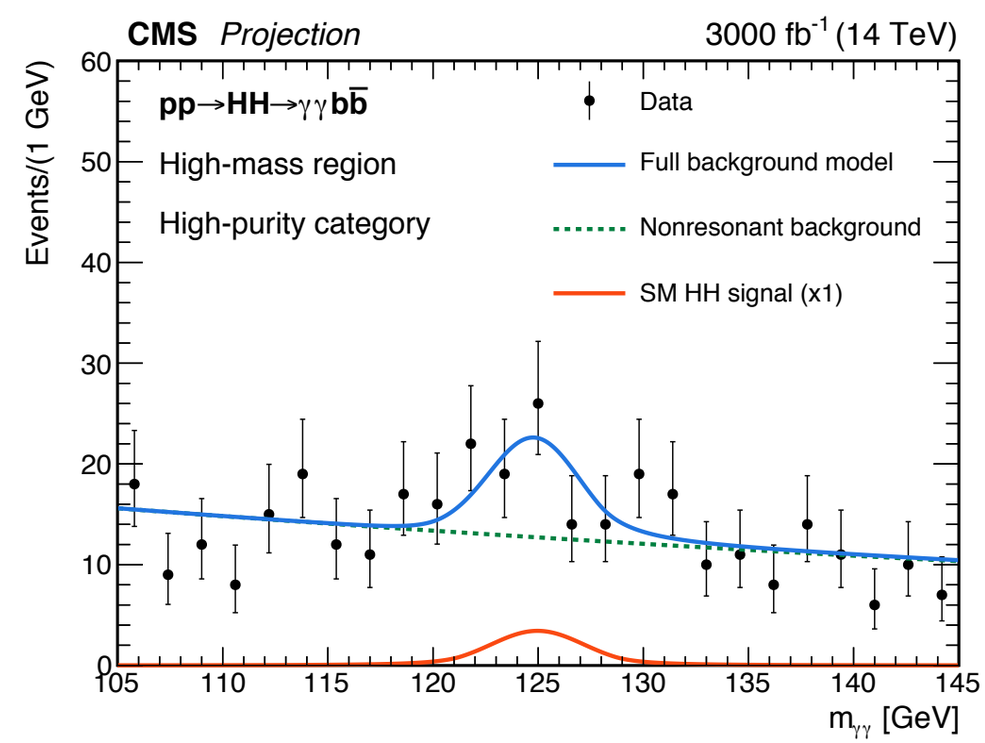
bb$\gamma\gamma$

- Analysis strategy
  - preselection of events containing $\gamma\gamma$ and bb pairs
  - rejection of ttH background with a dedicated BDT (75% reduction for 90% signal efficiency)
  - classification of events based on $M_X = m_{jj\gamma\gamma} - m_{\gamma\gamma} - m_{jj} + 250$ GeV into low and high mass categories
  - MVA event categorisation BDT that separates the background and the HH signal into low and high purity
  - ($m_{\gamma\gamma}$, m_{jj}) plane to look for a signal

$bb\gamma\gamma$

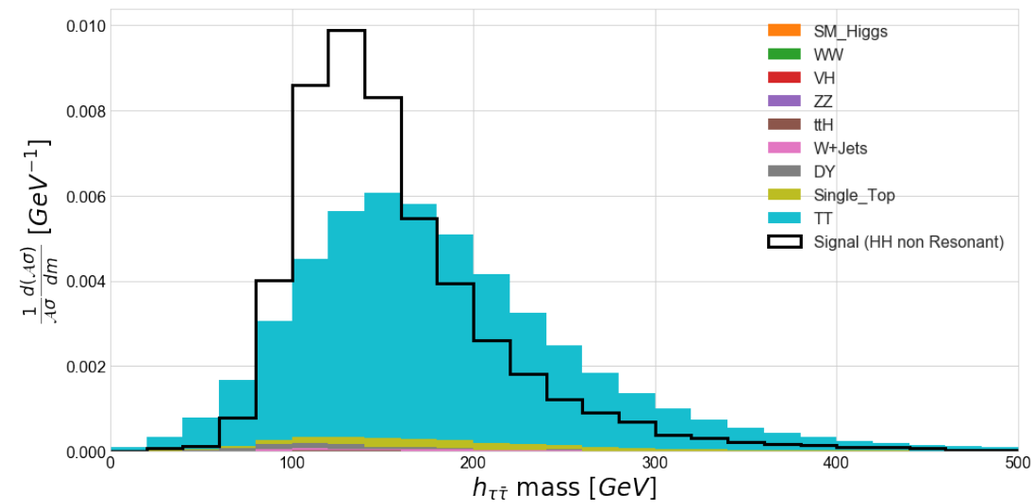
- Simultaneous fit on the $m_{\gamma\gamma}$ and m_{jj} distributions for the analysis categories
- Results: 1.8σ significance, 1.1 x σ_{HH}^{SM} limit
  - largely sensitive to the assumed $m_{\gamma\gamma}$ resolution
  - a few optimisations on the selections and MC generation are in the pipeline

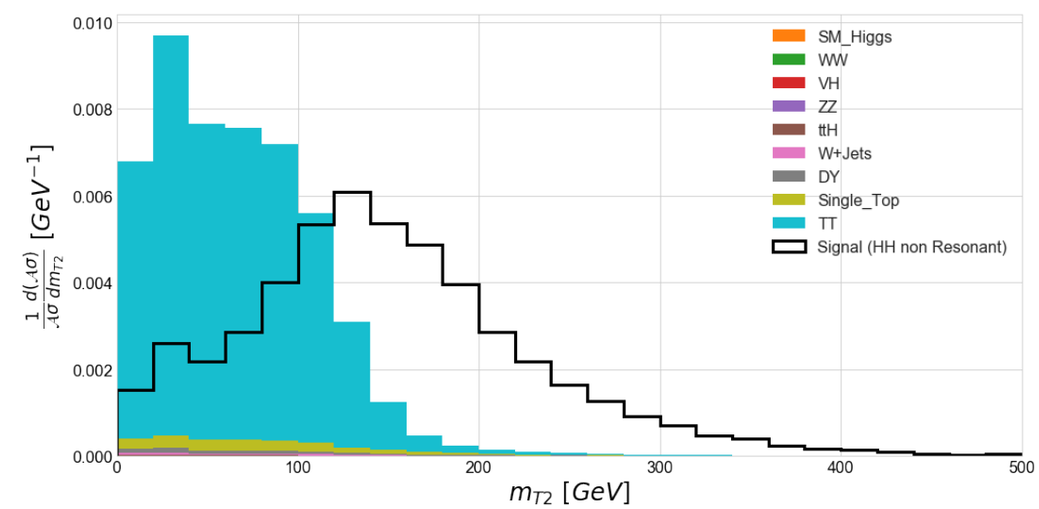




$bb\tau\tau$

- 3 $\tau\tau$ categories: $\mu\tau_h$, $e\tau_h$, $\tau_h\tau_h$
- Analysis strategy
  - preselect events based on μ, e, τ_h and b jet content
  - look for the presence of a signal using the output of a neural network
- State of the art techniques applied for the NN development
  - 27 basic + 21 reconstructed + 4 global features
  - deep learning techniques, with optimal data preprocessing, study of the activation functions, and data augmentation

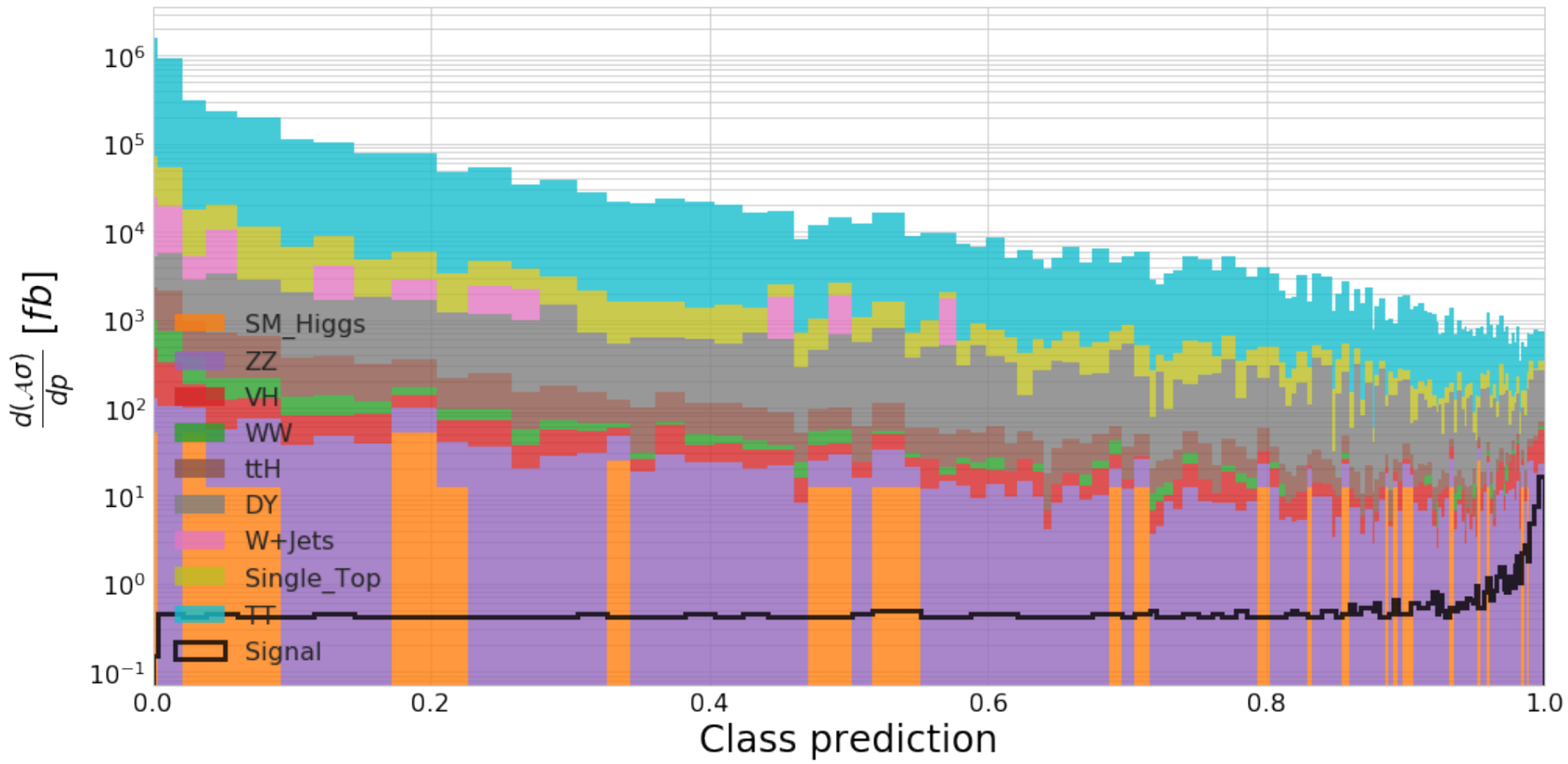




Examples of input features

$bb\tau\tau$

- Simultaneous fit of the NN output for the 3 decay channels
- Systematic uncertainties applied following the YR5 recommendations
- Results: 1.3σ significance, 1.5 x σ_{HH}^{SM} limit
  - without systematics: 1.5σ significance, 1.3 x σ_{HH}^{SM} limit

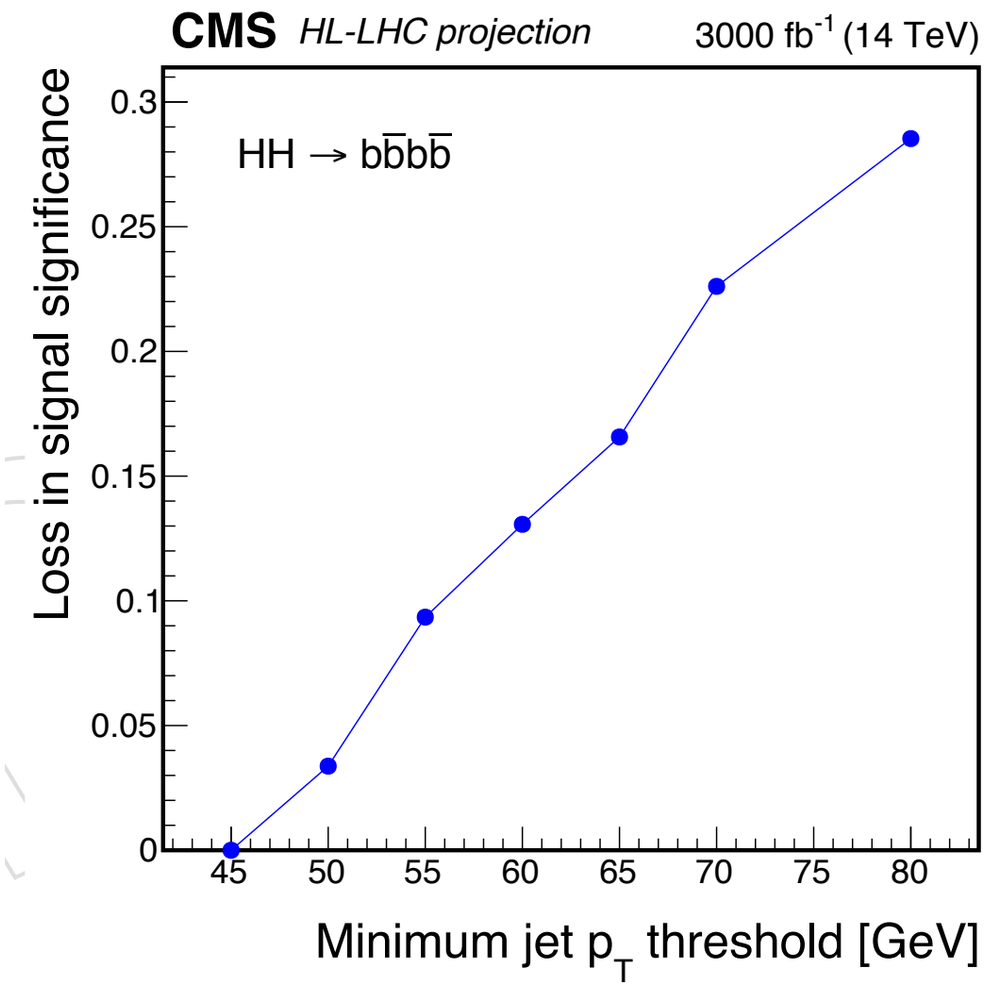


bbbb

- Analysis strategy:
  - preselect events with 4 jets of $p_T > 45$ GeV, $|\eta| < 3.5$, all jets b-tagged
  - build H_1 and H_2 as the jet pairs giving the minimal (H_1, H_2) mass difference
  - define the signal region in the (m_{H1}, m_{H2}) plane as

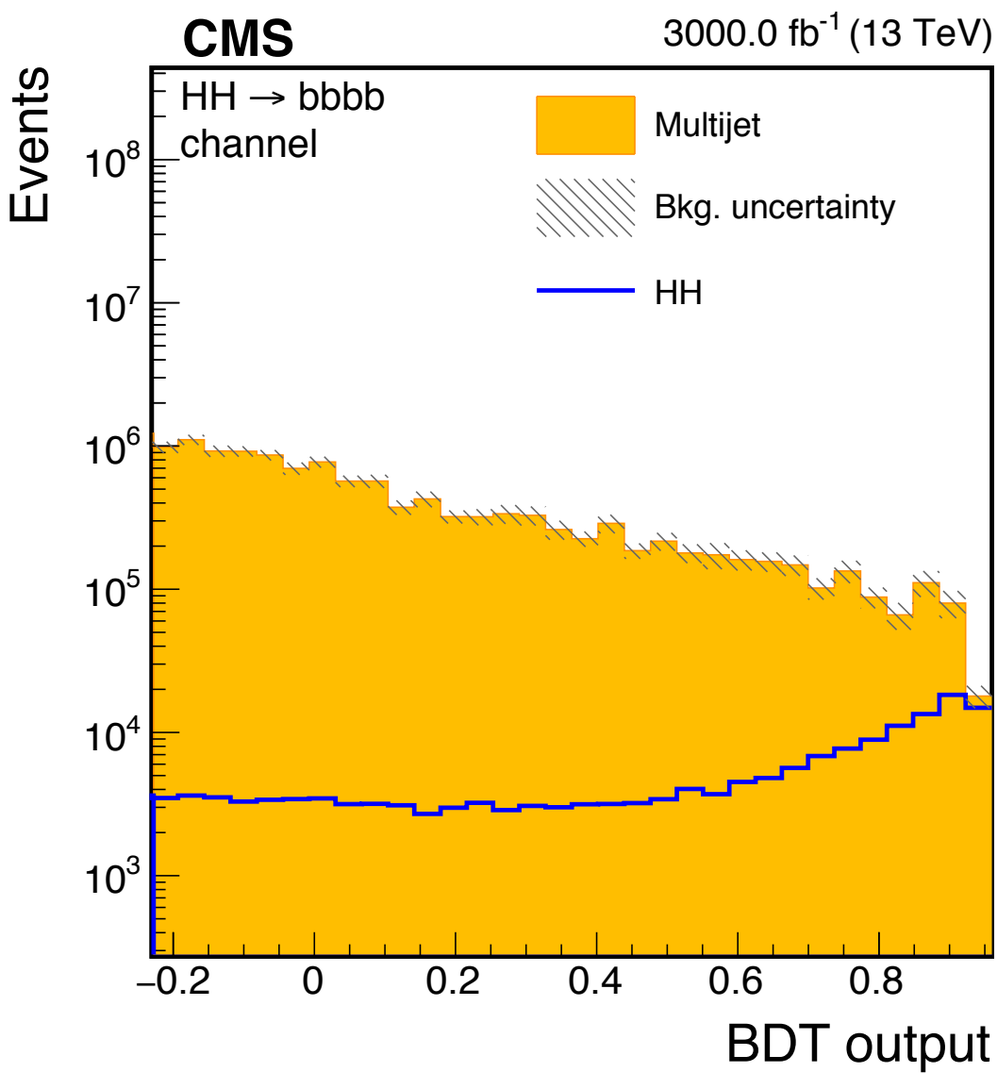
$$\sqrt{(m_{H_1} - 120\,\mathrm{GeV})^2 + (m_{H_2} - 120\,\mathrm{GeV})^2} < 40\,\mathrm{GeV}$$

    (note: improvements from b jet energy regression are not considered)
  - train a BDT to separate the background processes from the HH signal
  - use the BDT output as the fitted variable to look for a signal
- Usage of boosted categories also envisaged



Jet trigger threshold is a potential analysis limitation at the HL-LHC
The loss of significance vs minimal p_T threshold has been studied

bbbb



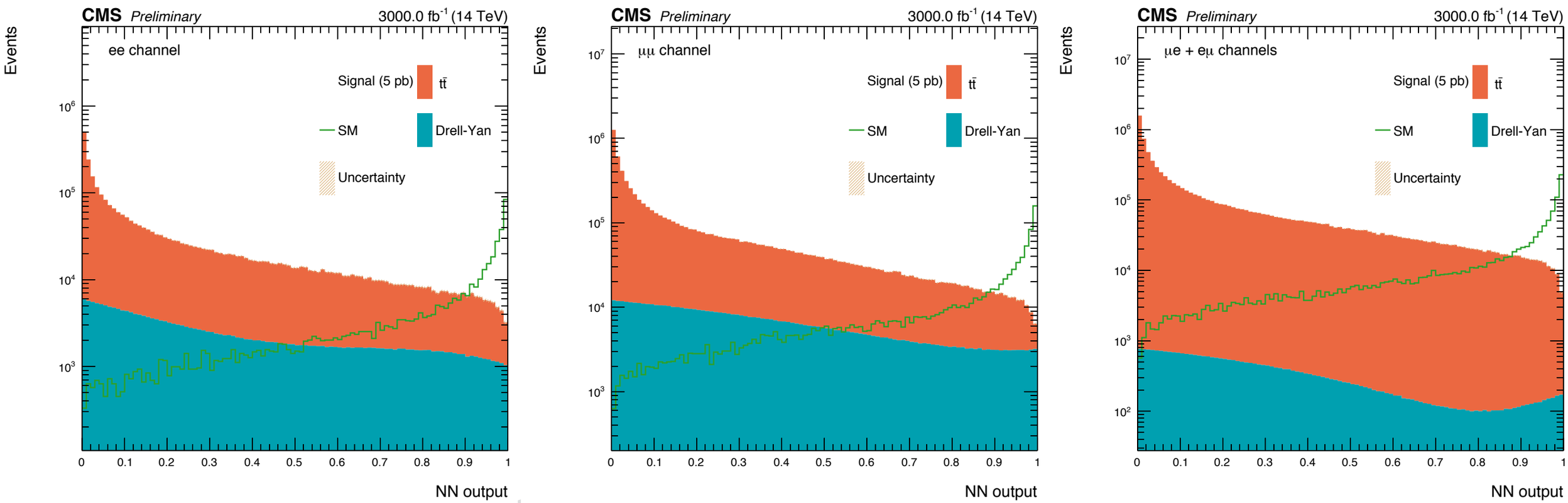
- High yields at low BDT score → backgrounds well constrained
  - almost no impact from syst. uncertainties
  - plan to consider variations of the most sensitive bins to evaluate impact of background uncertainty
    - but an aggressive assumption on the background estimation can be made for HL-LHC
- Results: 1.5σ significance, 1.3 × σ_{HH}^{SM} limit

bbWW ($2\ell\ 2\nu$)

- Using the fully leptonic WW final states: ee, μe, $\mu\mu$
- Select isolated leptons and two b jets in the event
- Large irreducible backgrounds: ttbar, DY
- Use a discriminant (neural network) based on kinematic properties to look for the signal
  - 9 input angular and mass variables
- Trigger and preselection: same thresholds as Run 2 applied
  - assumes lepton triggers will be used and 100% efficient in the analysis phase space
  - systematic uncertainties: following YR5 recommendations

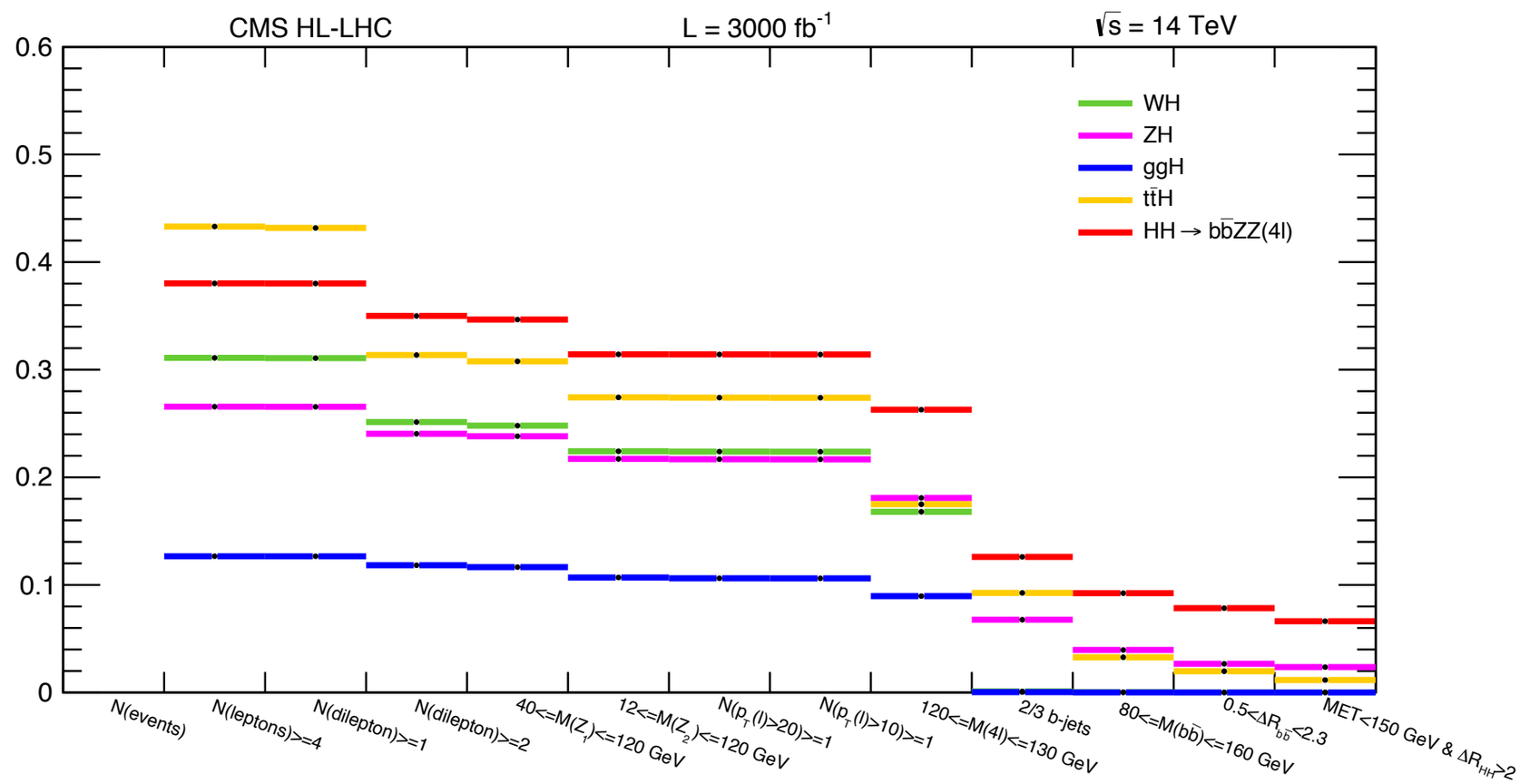
bbWW ($2\ell\ 2\nu$)

- Signal extracted from the NN output
  - simultaneous fit in the 3 categories
  - DY background modelling shown to impact the sensitivity by less than 5%
- Results: 0.4σ significance, 4.4 x σ_{HH}^{SM} limit



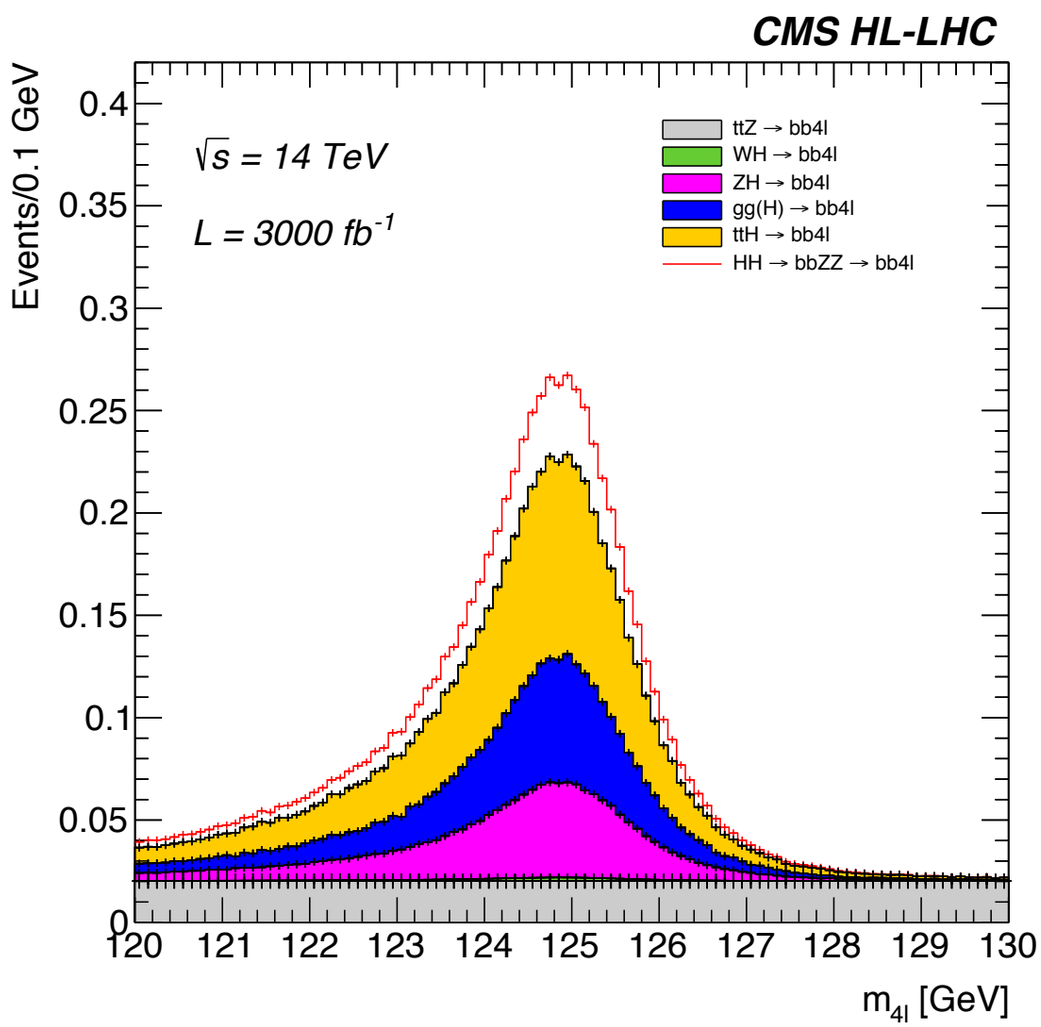
bbZZ (4ℓ)

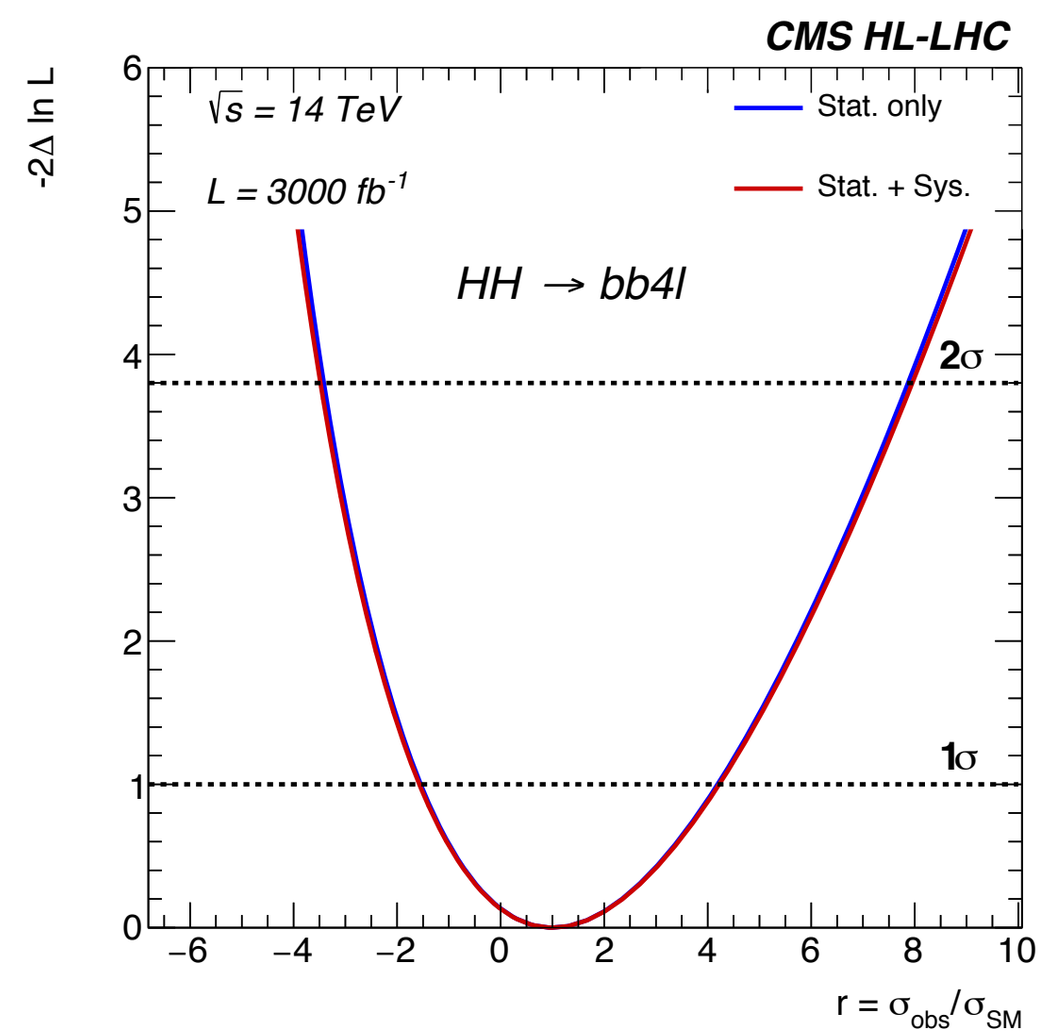
- Very rare but clean final state, yet unexplored at the LHC
- Powerful H→4ℓ signature ⟹ single H processes as dominant backgrounds
- Consider the $4e$, 4μ, $2e2\mu$ final states
- Loose preselection of leptons and jets, leptons combined to build the Z and Z* candidates
- Select events with $m_{4\ell}$ compatible with m_H
- Large acceptance because of low p_T of leptons
  - mostly limited by b tagging and p_T threshold
  - large acceptance for anomalous λ_{HHH} (small m_{HH})



bbZZ (4ℓ)

- Counting experiment with events around m_H
- YR 5 systematics included, but the analysis sensitivity is totally stat. dominated
- Results: 0.4σ significance, 6.7 × σ_{HH}^{SM} limit





Status and next steps

- SM analysis for the five decay channels ready
- Working on the combination of the 5 final states
  - close to 3σ from a naïve sum in quadrature, a few improvements in the pipeline
- Extend the set of results
  - anomalous λ_{HHH} couplings, exclusion of the $\lambda_{HHH} = 0$ hypothesis
  - likelihood scan
  - for bbbb only in boosted topologies, constraints on contact interactions (since they impact the high m_{HH} regime)
- Editing status:
  - text with analysis and result description not yet in the YR draft
  - SM results close to be finalised

Channel	Significance $[\sigma]$	95% CL limit on $\mu = \sigma_{\mathrm{HH}}/\sigma_{\mathrm{HH}}^{\mathrm{SM}}$
bbbb	1.5	1.3
bb$\tau\tau$	1.3	1.5
bbWW($\ell\nu\ell\nu$)	0.4	4.4
bb$\gamma\gamma$	1.8	1.1
bbZZ($\ell\ell\ell\ell$)	0.4	6.7

Combination with ATLAS

- ATLAS and CMS have differences in their results
  - different hierarchies of sensitivities: discussed between the authors and understood in terms of object performance, analysis strategy, and assumptions on the projection
  - CMS explores 5 channels, ATLAS projects 3 channels, similar overall sensitivity
  - $\Longrightarrow$ complementary approach giving a broad overview of the HL-LHC reach
- Analyses affected by systematic uncertainties that are mostly uncorrelated between the experiments (bkg estimation method, object performance, etc..)
  - correlations between experiments are expected to be negligible
  - $\Longrightarrow$ propose a “simple” combination at the likelihood level (sum of $\log\mathscr{L}$) to project reliably the HL-LHC reach for HH
  - same framework for both the projected combination of the SM HH signal sensitivity and the projected λ_{HHH} measurement

Conclusions

- The CMS Collaboration has projected the sensitivity to HH production using 5 decay channels (bbbb. bbWW, bbZZ, bb$\tau\tau$, bb$\gamma\gamma$)
  - full analyses based on Delphes CMS upgraded detector simulation
  - innovative techniques (e.g. machine learning) and rare channels (bbZZ(4ℓ)) studied for the first time to assess the possibilities at the HL-LHC
  - result in the five decay channels ready, combination in preparation
  - overall sensitivity close to 3σ with a few optimisations in the pipeline
  - extension of the results to λ_{HHH} measurements in progress
- Draft of the CMS projection section still to be inserted in the YR
- Propose a likelihood based approach for the combination with ATLAS
  - expect naïvely to reach a sensitivity beyond 4 sigma for the SM HH signal