Preparation to the Higgs self-coupling measurement using the HH → bbγγ channel

Maxime Fernoux

Supervisor : Elisabeth Petit

M2 Internship defence

Wednesday 23rd June 2021

Table of contents

- Introduction
  - Presentation of the context of the internship
  - Physics motivation and goals of HH research

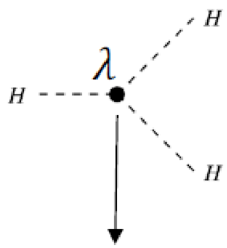
- Presentation of the analysis
  - Challenges of HH signal selection and background study
  - Preselection of the events and latest ATLAS analysis

- First look at ZH $\rightarrow b b \gamma\gamma$ channel
  - Explore the m_{bb} distribution
  - Modification of the preselection for ZH ?
  - Towards a 2D fit

- Conclusion

Introduction

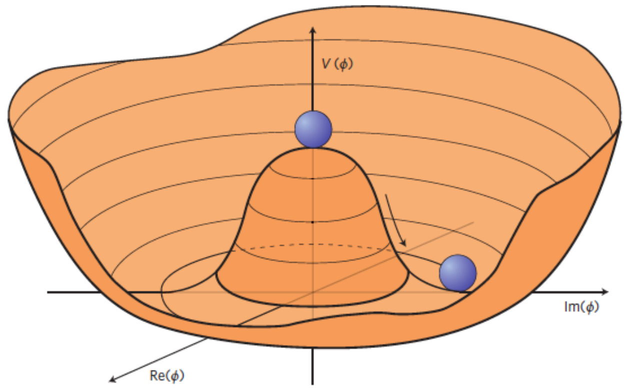
- Physics motivation to observe two Higgs bosons (HH)
  - Probe Higgs boson to self-coupling will help constrain Higgs potential "mexican hat" shape



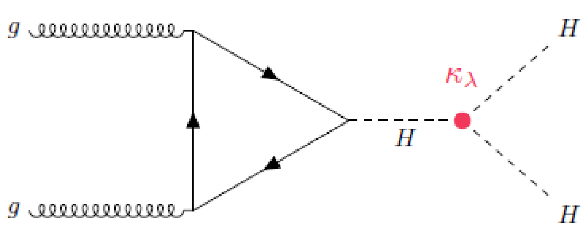
Direct access in HH pairs

$$\kappa_{\lambda} = \lambda_{HHH} / \lambda_{HHH\,SM}$$

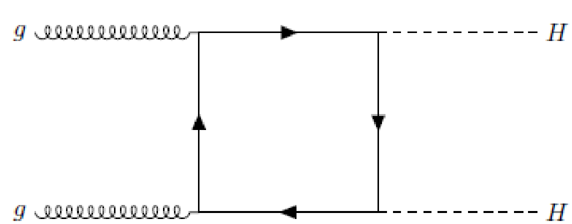
$$\begin{aligned} V(\phi^{\dagger}\phi) &= \mu^2\phi^{\dagger}\phi + \lambda(\phi^{\dagger}\phi)^2 \\ &\supset \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 \end{aligned}$$



  - Di-Higgs boson production and link with λ



(a) Trilinear Coupling



(b) Box Diagram

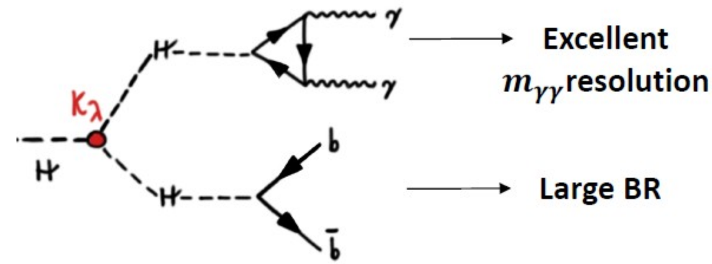
Main production mode : gluon-gluon fusion

σ really small : $\sigma_{HH}^{ggF} = 31{,}02\ fb$ to compare with single Higgs cross section : $\sigma_H = 59\ pb$

Analysis presentation (1)

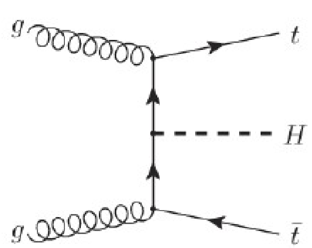
- bb$\gamma\gamma$ channel
  - Why this channel ?

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0005%

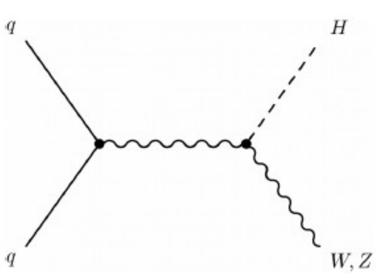


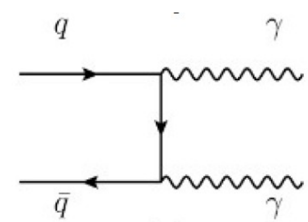
  - Different background processes
    - Single Higgs
    - Continuum : 2 photons + 2 jets

ttH

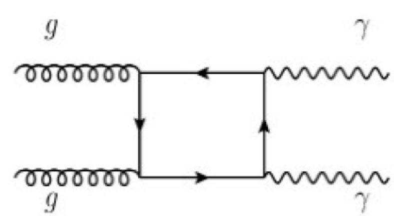


W/Z H

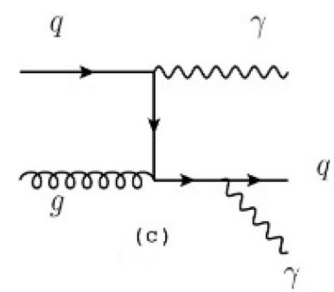




(a)



(b)



(c)

Analysis presentation (2)

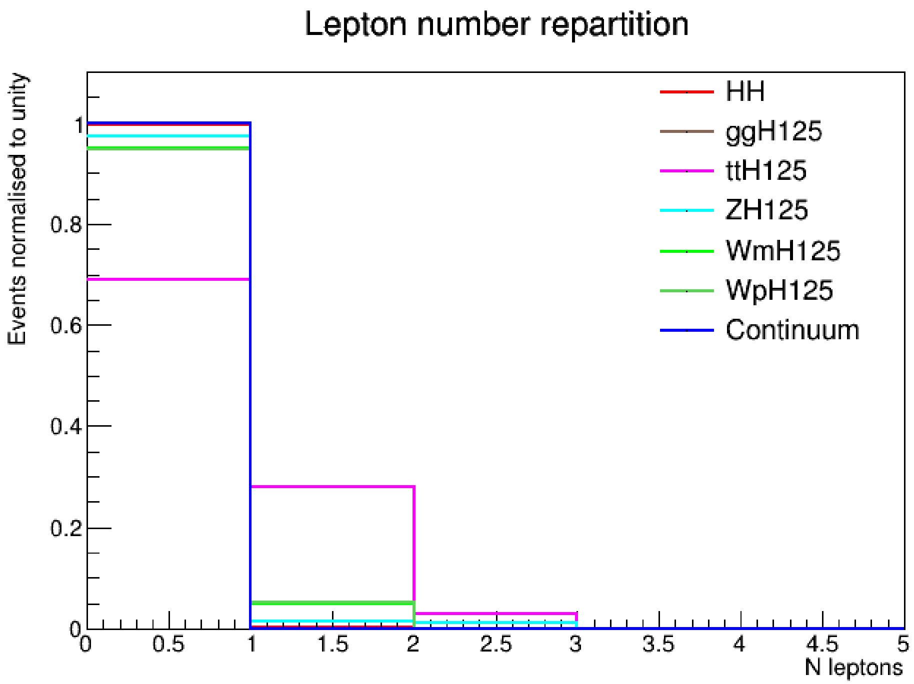
- A HH → bbγγ analysis with full LHC Run-2 dataset was released by ATLAS earlier in 2021: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-016/

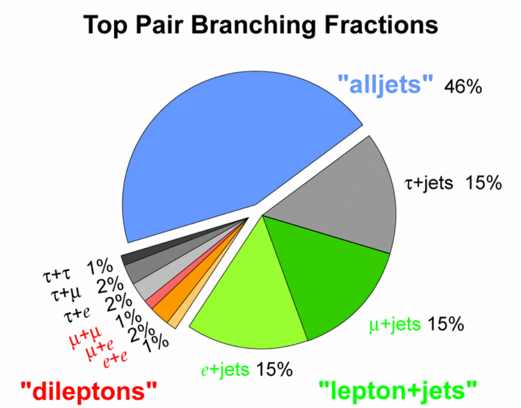
- Its preselection includes events :
  - with two photons, with a requirement on their transverse momentum pT : the pT of the leading (resp. sub-leading) photon must be larger than 35% (25%) of the invariant mass $m_{\gamma\gamma}$
  - without any leptons
  - with at least two jets
  - with less than 6 central jets
  - with exactly two b-tagged jets in order to separate this analysis with $HH \rightarrow b\bar{b}b\bar{b}$
    Usually, 77% b-tagging efficiency is used.

Analysis presentation (3)

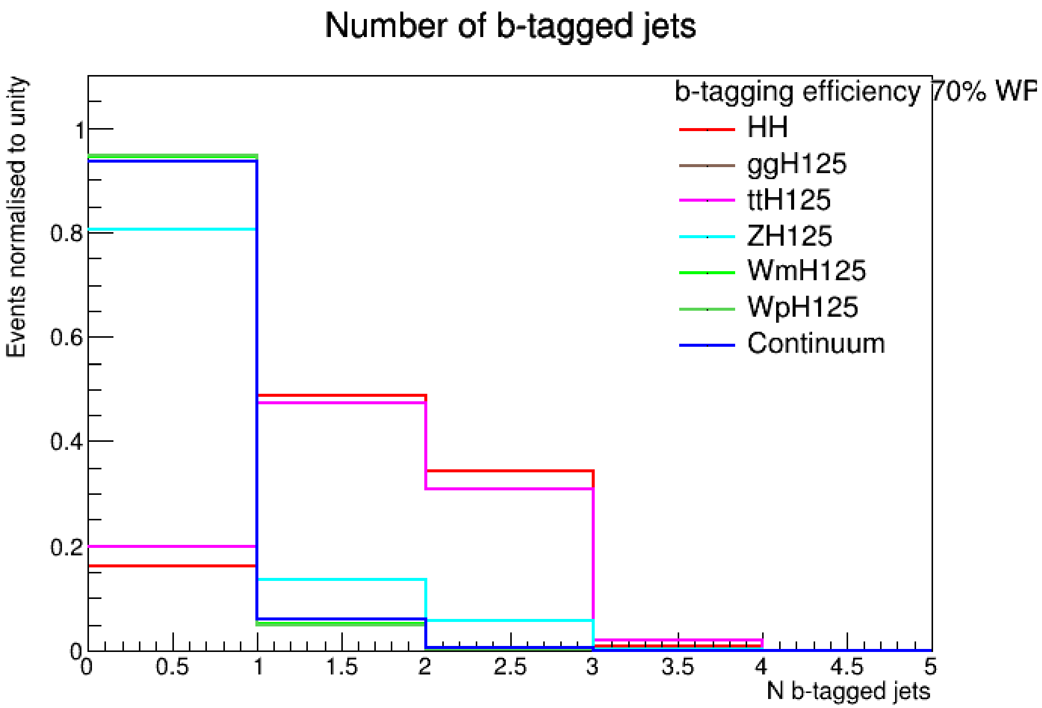
- Check the usefulness of the cuts

Lepton veto





Number of b-tagged jets requirement



Analysis presentation (4)

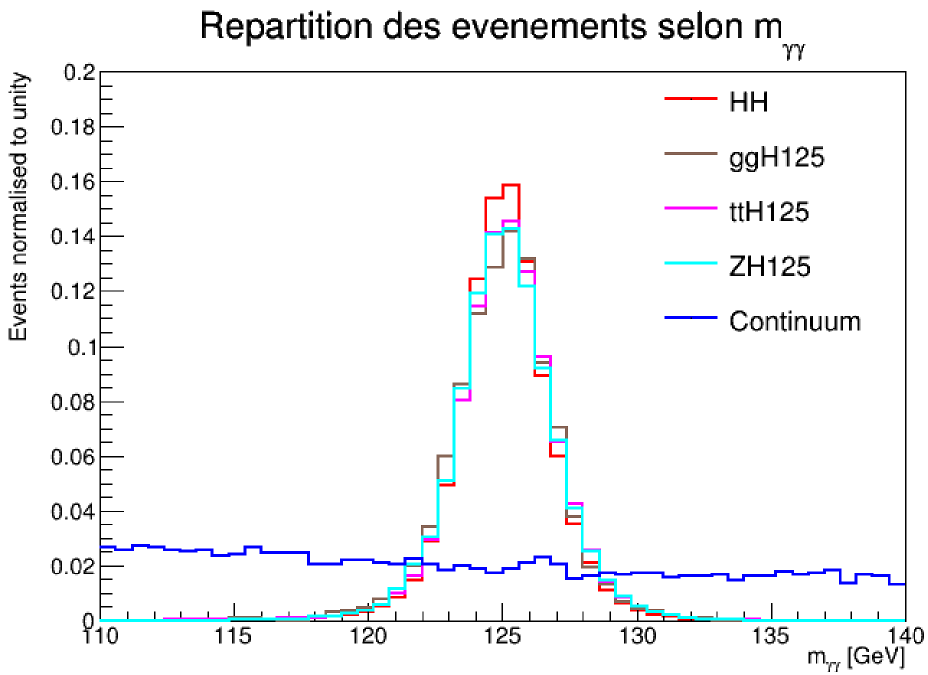
- Yields recap after LHC Run-2 dataset (139 fb^{-1})

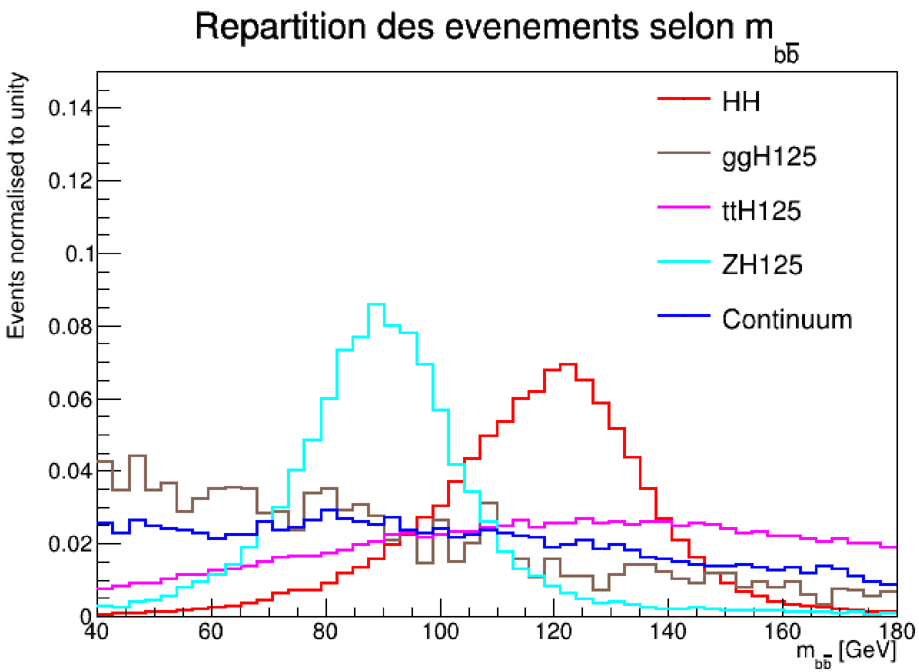
	HH	ZH	WH	ttH	ggH	continuum
Initial	11.37	239.7	432.1	159.8	15 310	7 203 397
2 photons	4.59	70.97	125.4	52.75	7840	918 518
N jets >= 2	3.87	35.36	59.35	51.78	1031	172 492
N leptons = 0 + N central jets < 6	3.72	34.01	55.69	22.42	1026	171 207
2 B-jet tagging 77 % WP	1.53	2.57	0.19	7.29	7.84	1376
Total efficiency (%)	13.45	1.07	0.04	4.56	0.05	0.02

- Signal / background ratio still low after preselection → the current analysis uses a boosted decision tree (BDT) selection

Analysis presentation (5)

- $m_{\gamma\gamma}$ and m_{bb} distribution





•Used by previous ATLAS analysis
•Good resolution

•HH signal is more isolated
•Bad resolution

- ZH is a background really similar to HH signal
  - it could be seen at the end of LHC Run-3 in 2025 thanks to a larger cross section
  - this would help validate the analysis

Branching ratios	
$H \to bb$	58%
$H \to \gamma\gamma$	0.228%
$Z \to b\bar{b}$	15%
$Z \to \gamma\gamma$	$O(10^{-5})$

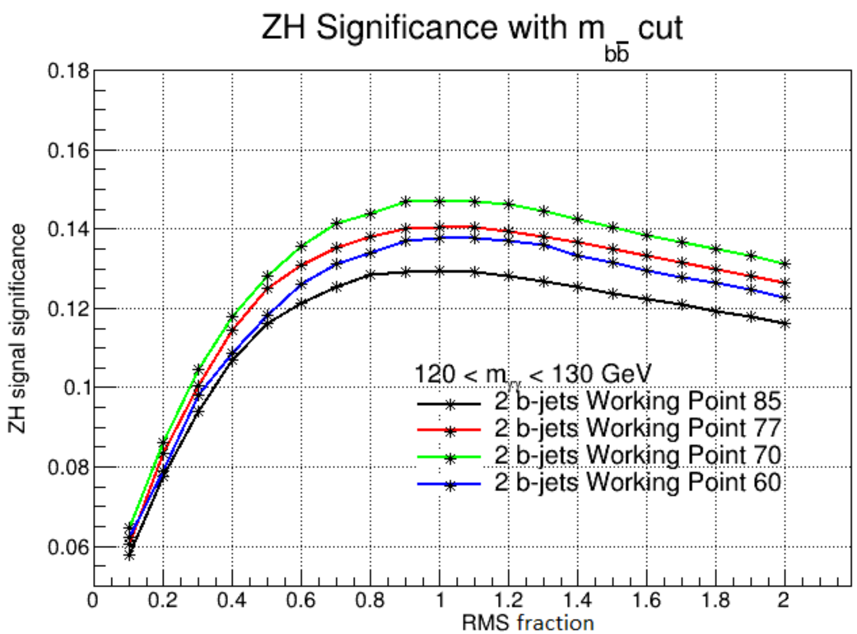
Prod. cross section	
$\sigma(HH)$	31 fb
$\sigma(ZH)$	880 fb
$\sigma(HH \to b\bar{b}\gamma\gamma)$	0.08 fb
$\sigma(ZH \to b\bar{b}\gamma\gamma)$	0.3 fb

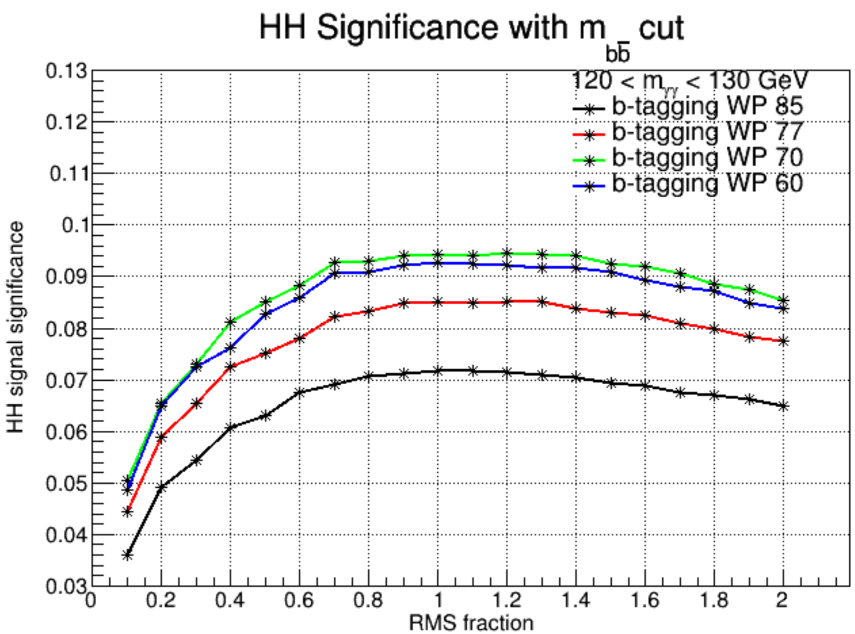
First look at ZH → bbγγ (1)

- Up to know, BDT is trained against all backgrounds including ZH but in the future, we would like two categories aiming for HH and ZH so that we can fit both signals simultaneously
- First simple approach : cuts around m_{bb} peak at m_H and m_Z
  - we select events in the interval $[m_{peak} – fraction*RMS ; m_{peak} +fraction*RMS]$

  and we study the effect of the cut on the signal strength
  - it is quantified by Asimov's formula for significance : $\sqrt{2\left((s+b)\ln(1+s/b)-s\right)}$

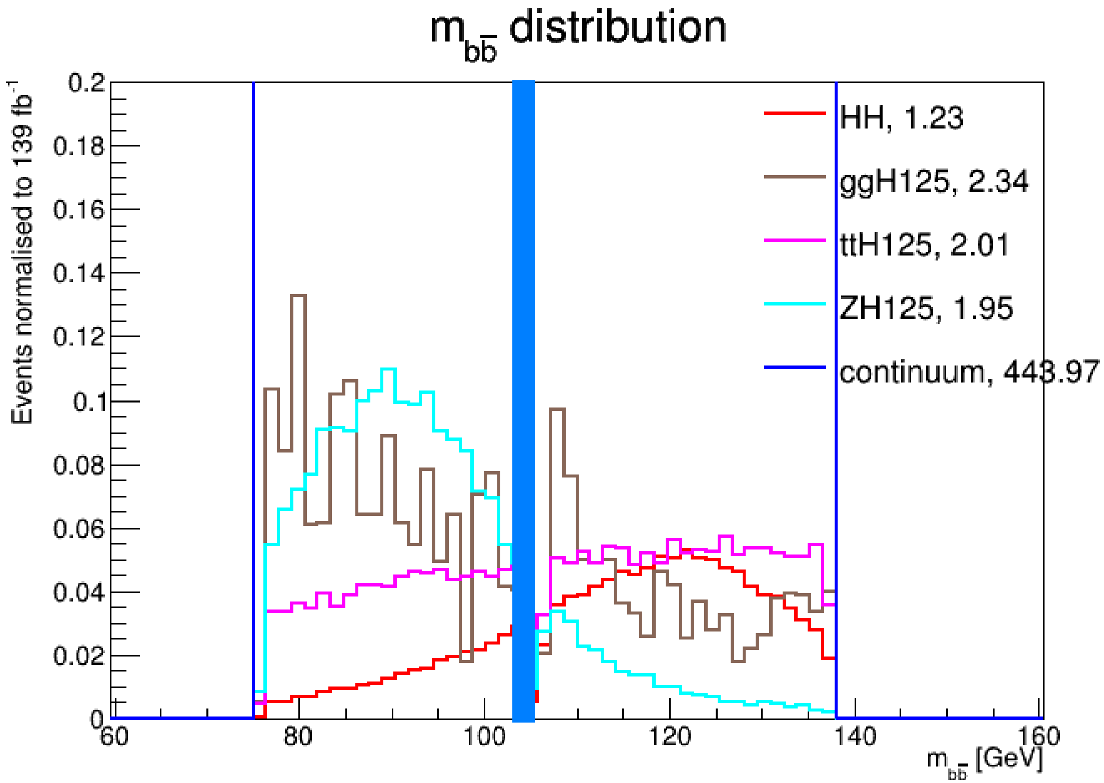
    where s (resp. b) is the number of signal (background) events





First look at ZH → bbγγ (2)

- The next step is to make a simultaneous cut around m_Z and m_H m_{bb} peaks and to quadratically sum the significance in the two zones if they don't overlap



  - The best summed significance is reached for 0.8*RMS and the cuts intervals are [76.2 ; 104.7] and [106.2 ; 137.6] GeV

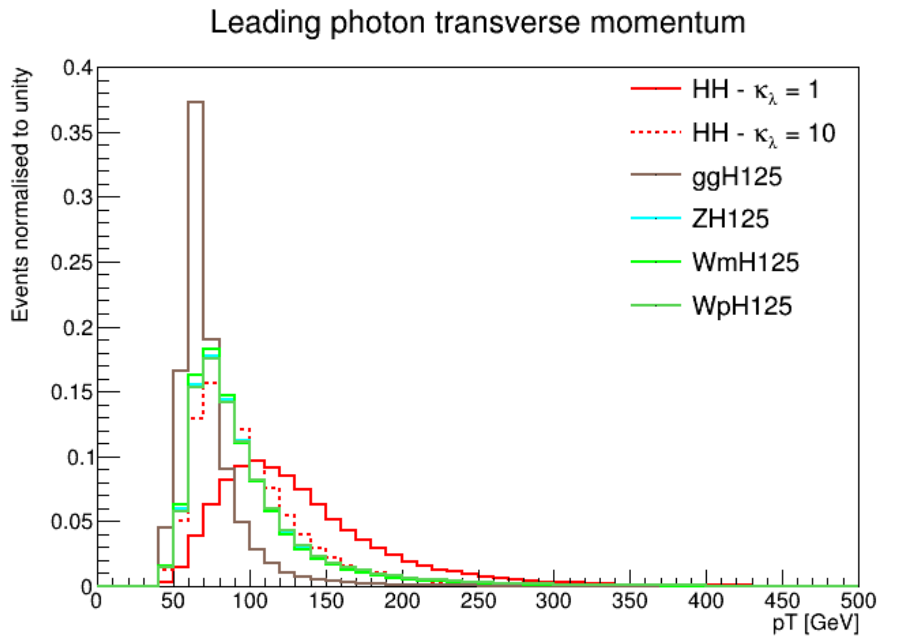
$m_{b\bar{b}}$ interval	m_Z peak zone	m_H peak zone	Combination
Z_{ZH}	0.139	0.028	0.142
Z_{HH}	0.024	0.083	0.087

First look at ZH → bbγγ (3)

- Improve preselection for ZH
  - Preselection is not optimized for ZH and high values of κ_λ

	HH	HH (κ_λ =10)	ZH	Z(bb) H(γγ)	Z(bb) H(γγ) / HH
Initial	11.37	246.3	239.7	36.24	3.19
2 photons	4.59	89.2	71.0	10.6	2.34
Selection	1.53	25.7	2.57	2.57	1.68
Total efficiency (%)	***13.45***	***10.43***	*1.07*	***7.09***	

- This may be due to pT selection and kinematic differences between ZH and high value κ_λ HH, and $\kappa_\lambda = 1$ HH

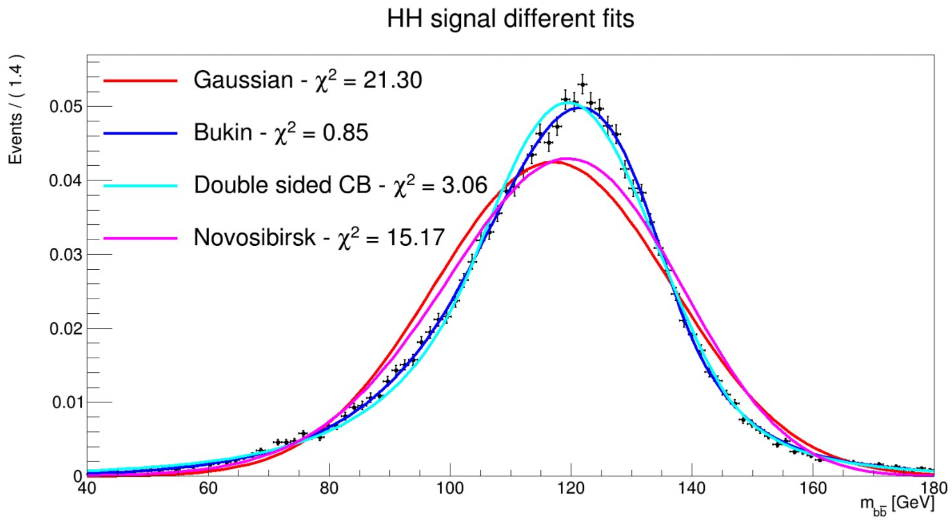


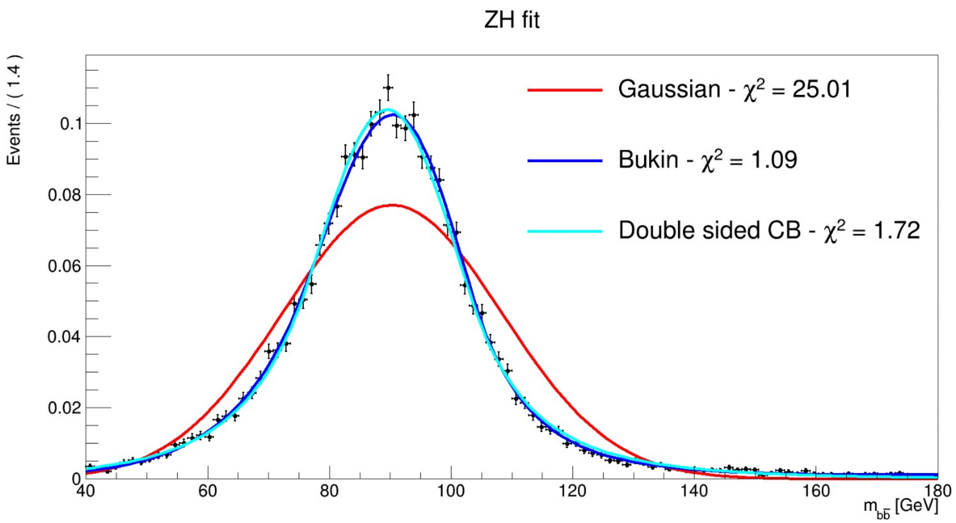
- Maybe modify photons pT trigger for Run-3 preselection ?

First look at ZH → bbγγ (4)

- So far, the signal is extracted from fits of the $m_{\gamma\gamma}$ distribution but this doesn't allow to separate well HH from single Higgs background so we would like to make a $m_{\gamma\gamma}$ x m_{bb} fit

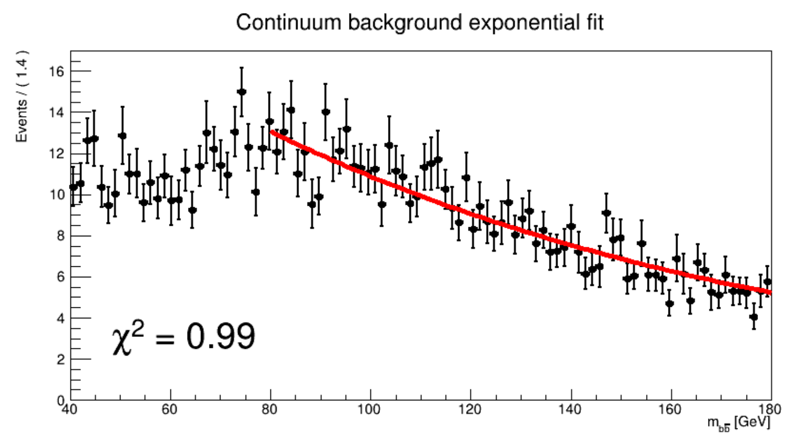
- We first need analytical functions of m_{bb} distribution
  - Fit HH and ZH with functions with peaks – Bukin works best

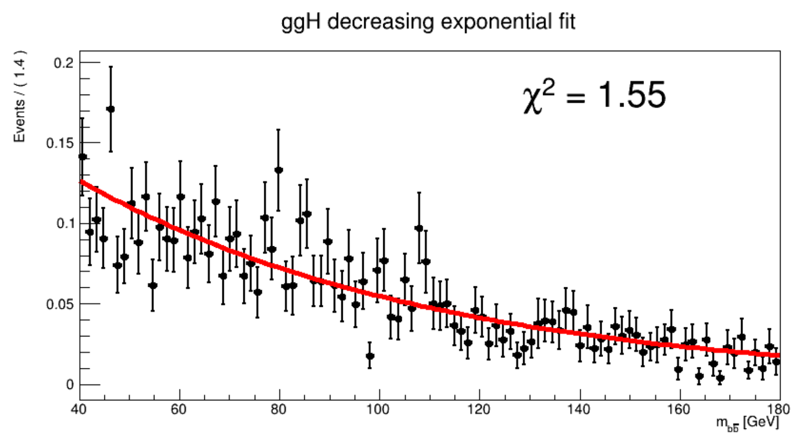


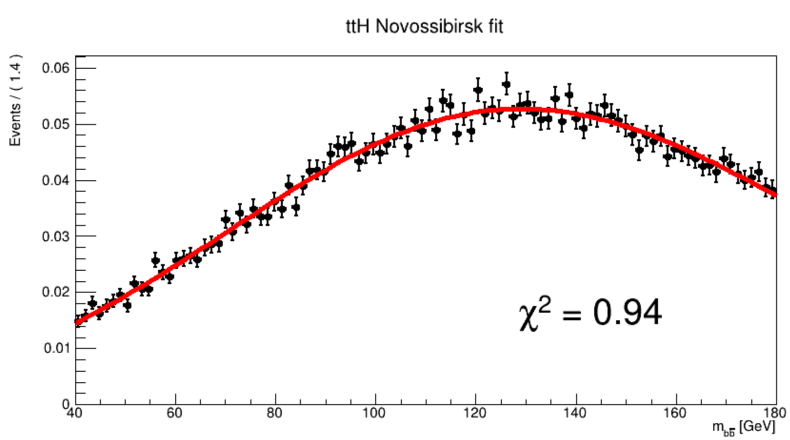


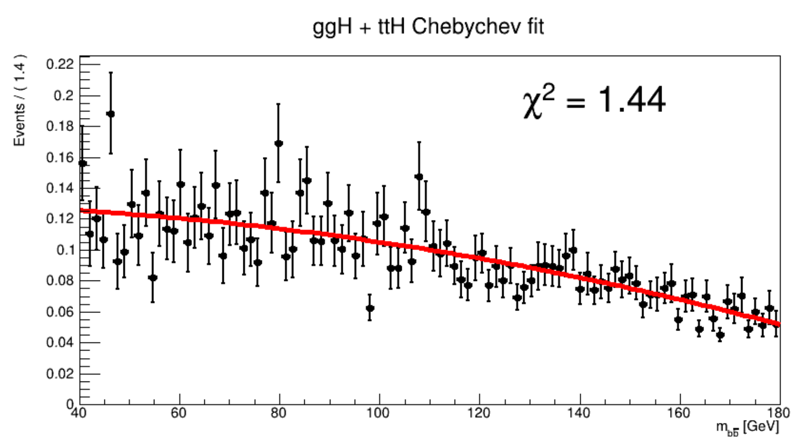
First look at ZH → bbγγ (5)

- Fits of backgrounds m_{bb} distribution



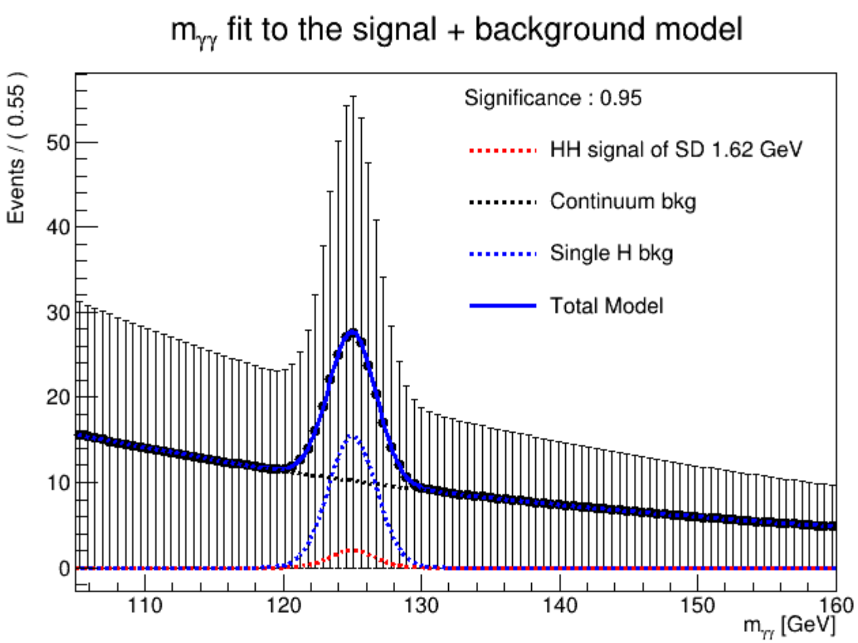


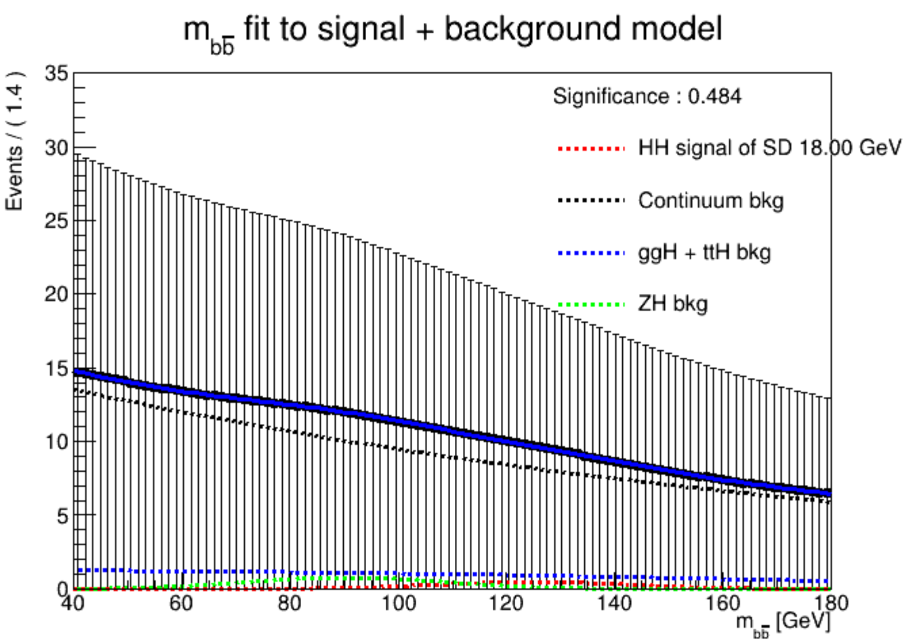




First look at ZH → bbγγ (6)

- Use of toys to compute significance of the signal with the fits
  - A toy is a simulated data set based on a base model, which is constructed with the previous fits here
  - The toy data set is compared with the "background only" hypothesis to compute the p-value and the significance



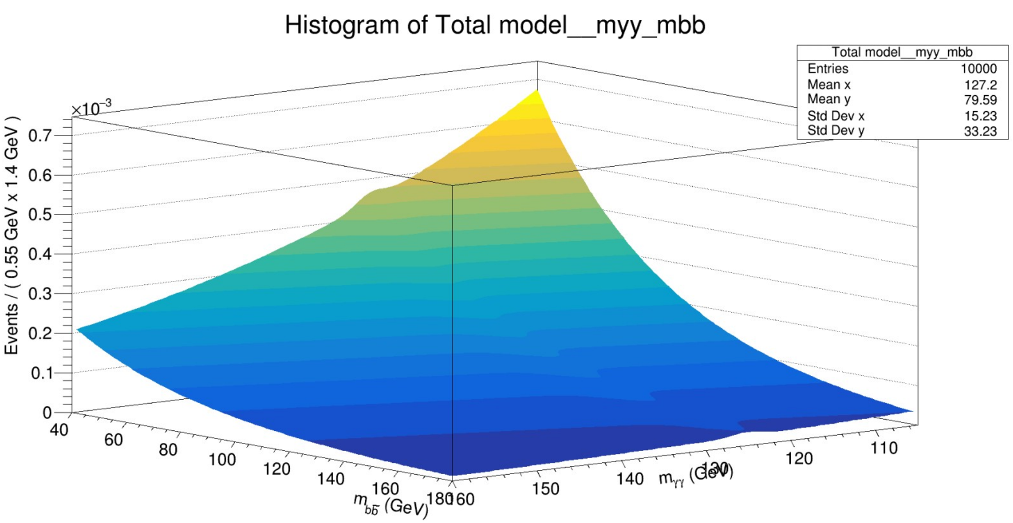


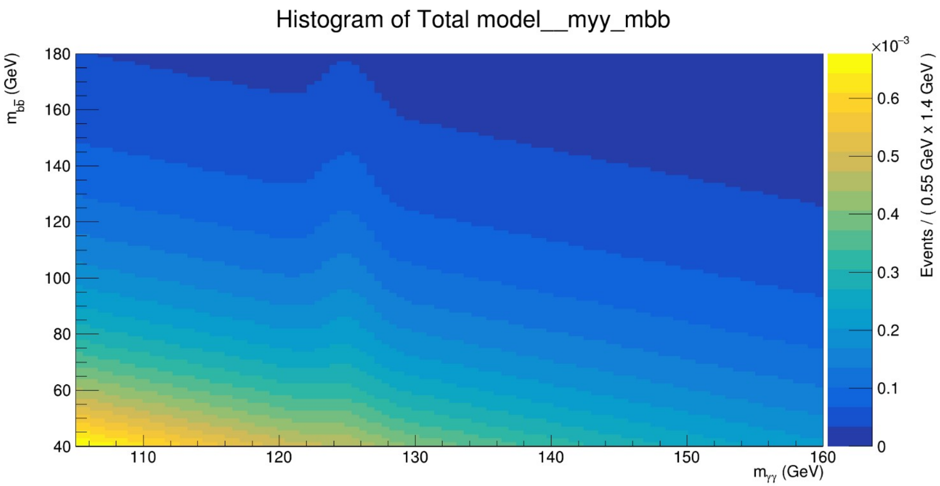
Here HH and single H background are multiplied by 10 to improve the visibility

  - As expected, the true significance values are better for a fit with $m_{\gamma\gamma}$ (0.129 σ) than with m_{bb} (0.051 σ)

First look at ZH → bbγγ (7)

- Towards a 2D fit of $m_{\gamma\gamma}$ and m_{bb}
  - First step : make a 2D toy data set





  - Second step : make a 2D fit of the Monte Carlo samples
  - Final step : use the 2D fit to make a direct 2D toy dataset and compute the significance to compare this method to the other ones

Conclusion

- Di-Higgs boson signal in the bbγγ channel is a challenge to observe and we don't expect it until 2030's. However, we may be able to see $Z \rightarrow bb$ $H \rightarrow \gamma\gamma$ at the end of Run-3 in 2025

- The first part of the internship has consisted in understanding the analysis preselection of events and we might need to modify it to look for ZH

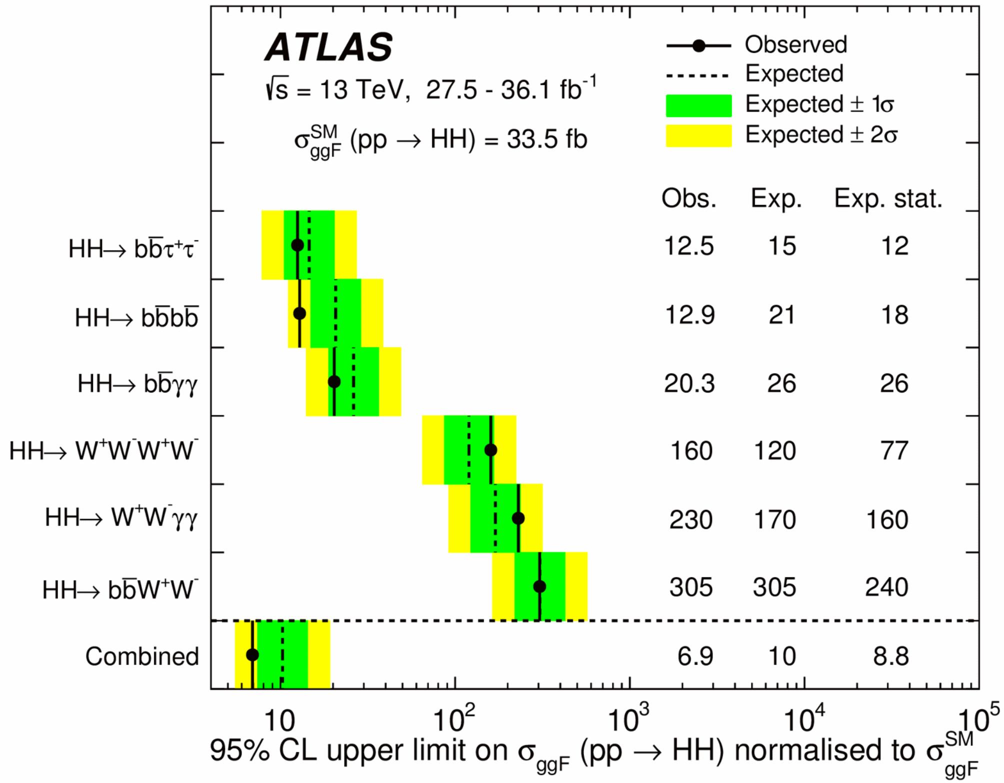
- The next step was to compare two processes to extract the signal : cuts around signal zone in m_{bb} distribution and fits along m_{yy} and m_{bb}

- The final idea is to make 2D fits and to compare the results with the previous method

Thank you for your attention !

Back up

- Interest of the bbγγ channel and comparison to the other ones



bbbb	Largest BR ☺ Large multijet and tt bkg ☹
bbττ	Sizeable BR ☺ Relatively small bkg ☺
bbγγ	Small BR ☹ Good diphoton resolution ☺ Relatively small bkg ☺
bbVV (→ lνlν)	Sizeable BR ☺ Large bkg ☹
bbZZ (→ 4l)	Very small BR ☹ Very small bkg ☺

Back up

- Changes of production cross section for a higher centre of mass energy

		HH	ZH
cross-sections [pb]	13 TeV	3,11E-02	8,84E-01
	14 TeV	3,67E-02	9,86E-01
	ratio	1,18	1,12

- Yields expected with LHC Run-3 dataset of 300 fb^{-1} (with the same selection)

	HH	ZH	Z(bb) H(γγ)	WH	ttH	ggH	continuum
Initial	24.54	517.4	78.22	932.5	344.9	33 042	15 546 900
Selection	3.30	5.55	5.55	0.41	15.73	16.92	2970
Total efficiency (%)	13.45	1.07	7.09	0.04	4.56	0.05	0.02