Maxime Gouzevitch

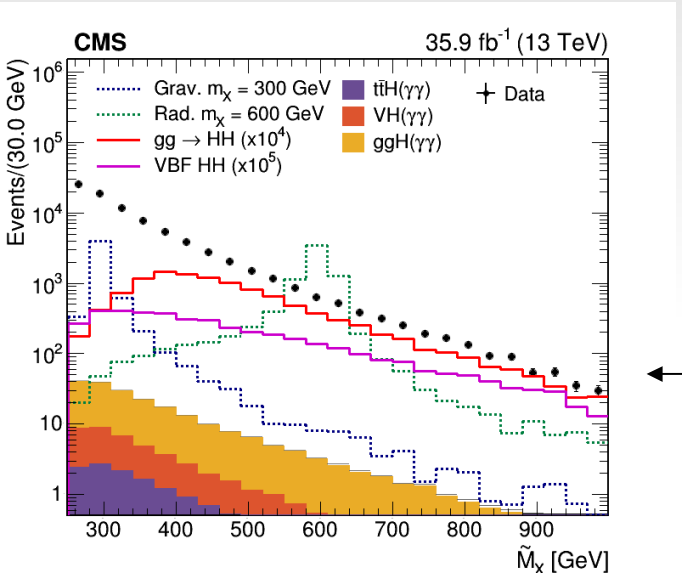
HH→ 2b2γ at CMS

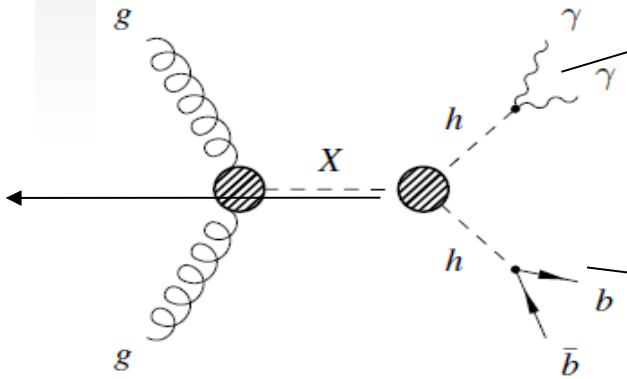
1) Introduction
2) Objects selection
3) Categorisation
4) Results

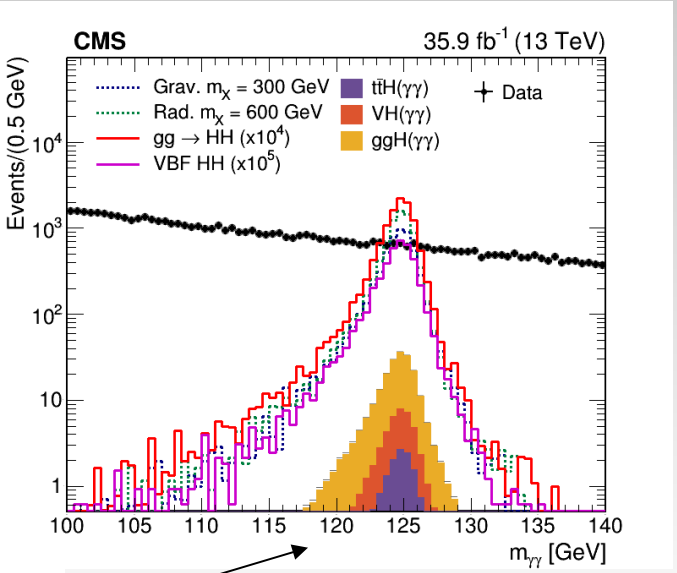
Introduction

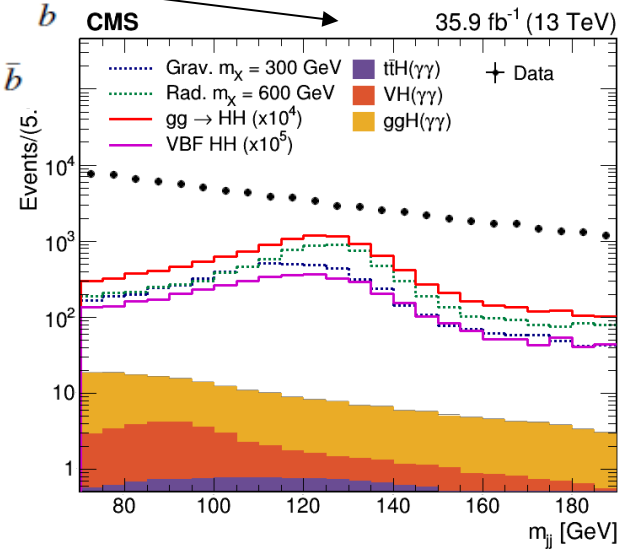
Introduction

- Fully reconstructible final state.
- Kinematically over-constrained analysis.



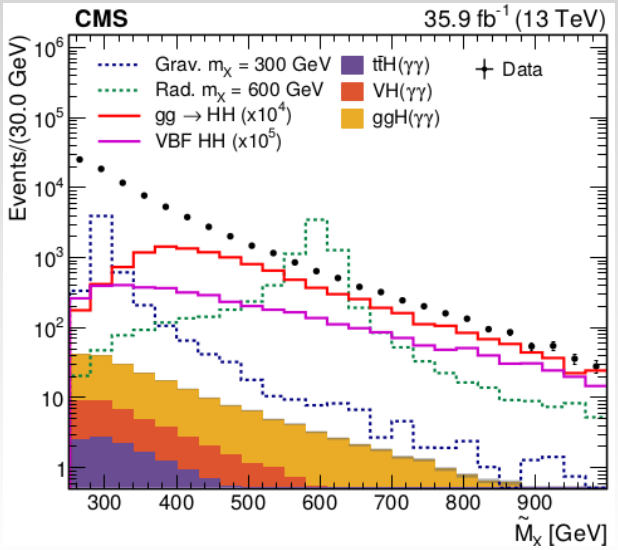


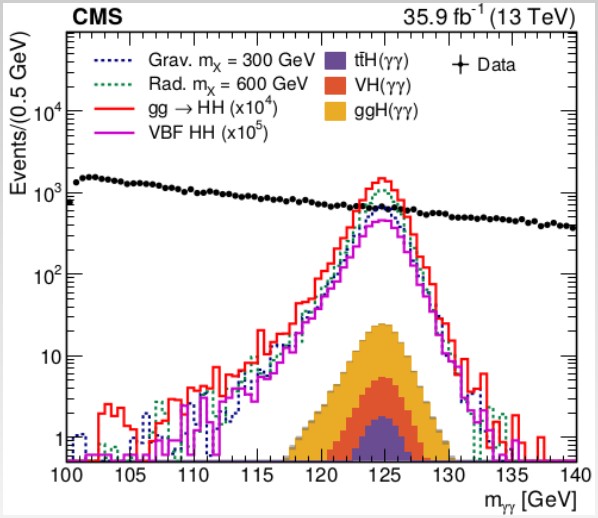


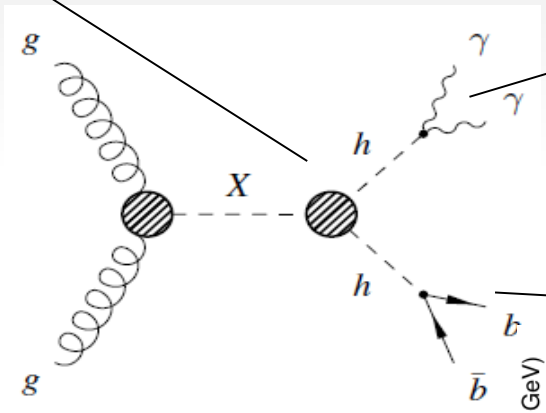


Main backgrounds:
- Non-resonant QCD = γγbb (>80%) + γjbb+jjbb (<20%).
- Resonant: SM H production – few events but positionned exactly under the diphoton peak.

Introduction

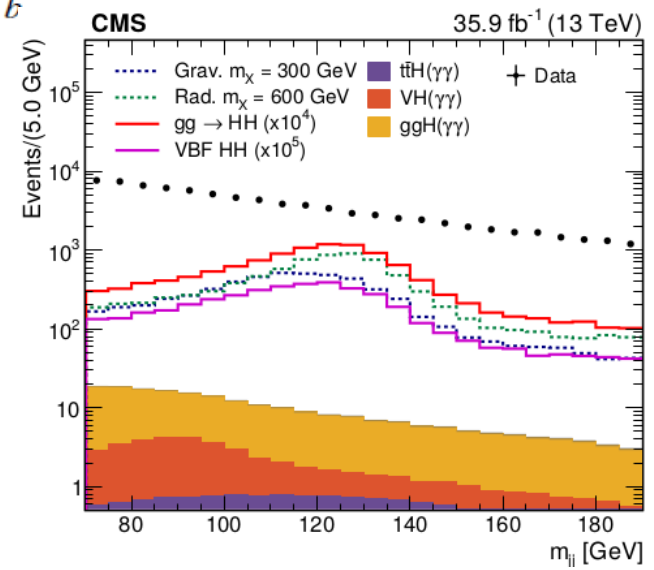


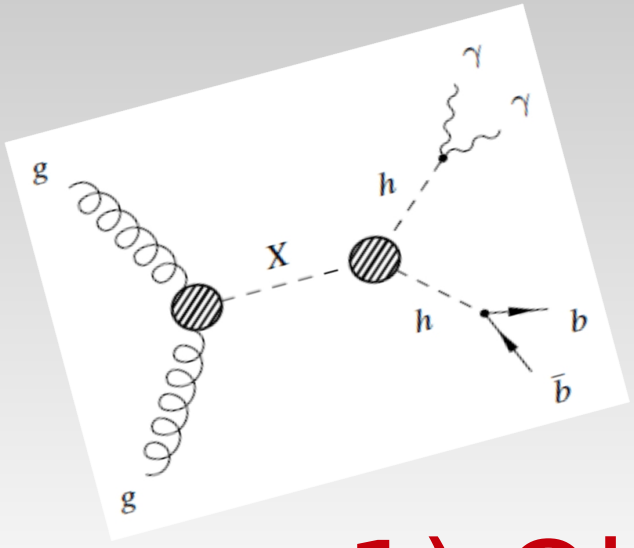


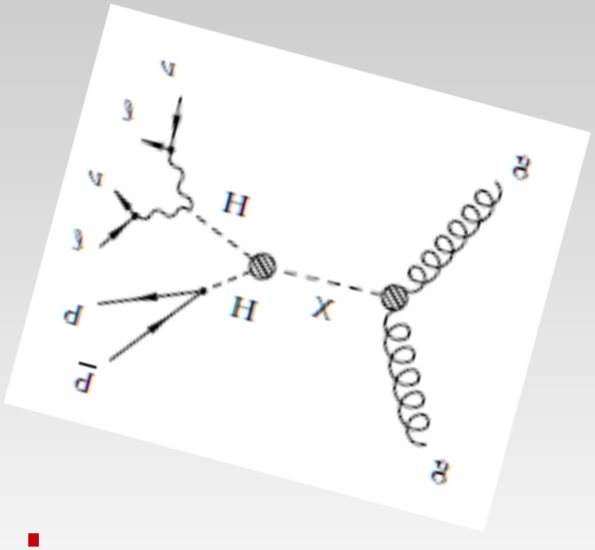


- Use likelihood $M_{\gamma\gamma}$ x M_{bb}.
- Categorize in $M_{\gamma\gamma bb}$ and MVA.

Analysis	Region	Classification MVA	$\tilde{M}_X$
Nonresonant	High-mass	HPC: MVA > 0.97 MPC: 0.6 < MVA < 0.97	$\tilde{M}_X > 350$ GeV
	Low-mass	HPC: MVA > 0.985 MPC: 0.6 < MVA < 0.985	$\tilde{M}_X < 350$ GeV
Resonant	$m_X > 600$ GeV	HPC: MVA > 0.5 MPC: 0 < MVA < 0.5	Mass window
	$m_X < 600$ GeV	HPC: MVA > 0.96 MPC: 0.7 < MVA < 0.96	Mass window







1) Objects selection

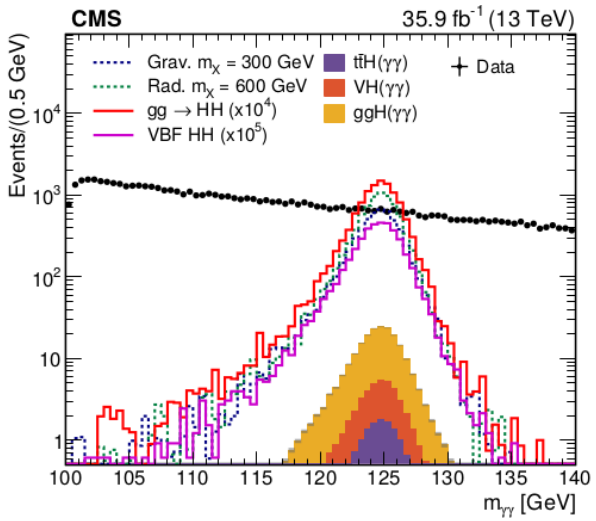
1.1) Photons selection

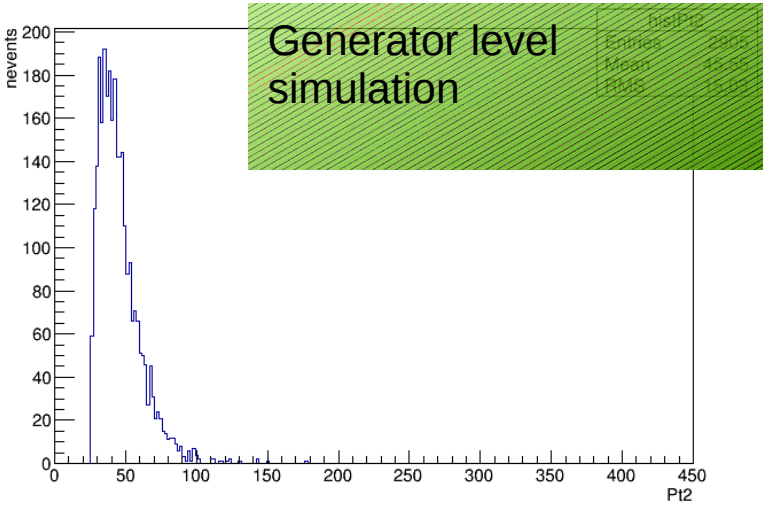
- Fully reconstructible final state.
- Similar to H → γγ but higher p_T of the objects.

Table 2: Summary of the baseline selection criteria.

Photons		Jets	
Variable	Selection	Variable	Selection
$p_T^{\gamma 1}$	$> m_{\gamma\gamma}/3$	p_T [GeV]	$> 25.$
$p_T^{\gamma 2}$	$> m_{\gamma\gamma}/4$	$\Delta R_{\gamma j}$	> 0.4
$\lvert\eta\rvert$	< 2.5	$\lvert\eta\rvert$	< 2.4

- Photons pT selection identical to H→γγ analysis. Optimized to keep $M_{\gamma\gamma}$ spectrum smooth.
- It is a bit more stringent than the trigger: $p_T^{\gamma 1} > 30\,\text{GeV}$ and $p_T^{\gamma 2} > 18\,\text{GeV}$. May be interesting to re-optimize in future for low M_X and κ_λ scan. Need to be careful to avoid too much fake photons.
- Photon id optimized using MVA definition. We realized that photon isolation for trailing photon can be worse than for leading one due to H→ bb presence for low M_X and κ_λ.





Second photon pT

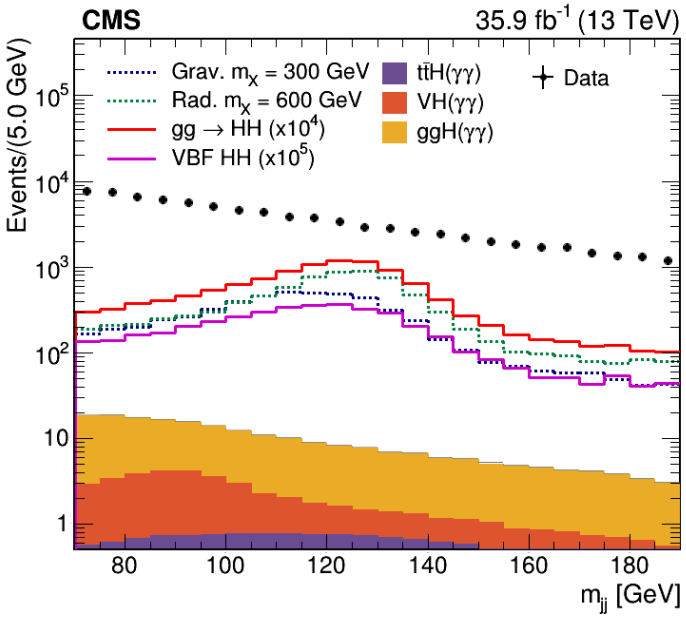
1.2) Jets selection

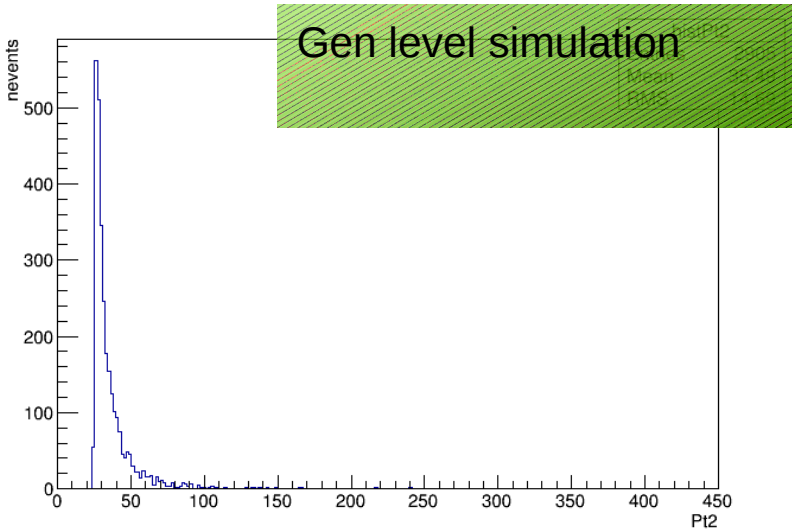
- Fully reconstructible final state.
- Similar to H → γγ but higher p_T of the objects.

Table 2: Summary of the baseline selection criteria.

Photons		Jets	
Variable	Selection	Variable	Selection
$p_T^{\gamma 1}$	$> m_{\gamma\gamma}/3$	p_T [GeV]	$> 25.$
$p_T^{\gamma 2}$	$> m_{\gamma\gamma}/4$	$\Delta R_{\gamma j}$	> 0.4
$\lvert\eta\rvert$	< 2.5	$\lvert\eta\rvert$	< 2.4

- Take jets with as low p_T as possible. B-tagging improves with lower p_T. Just need to be careful with PU jets. PF-low is particularly efficient there.
- Interesting question of jets pairing – correlated with classification:
Run II: mva classification → order jets by b-tag and take the first 2.

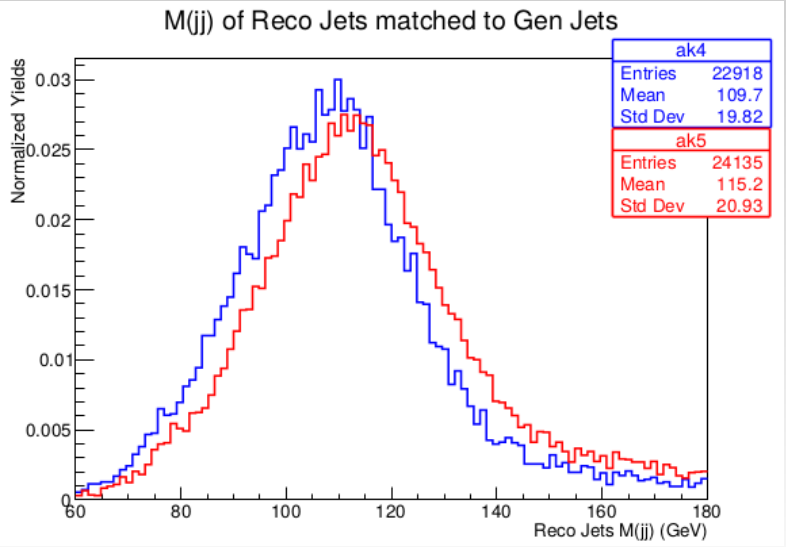




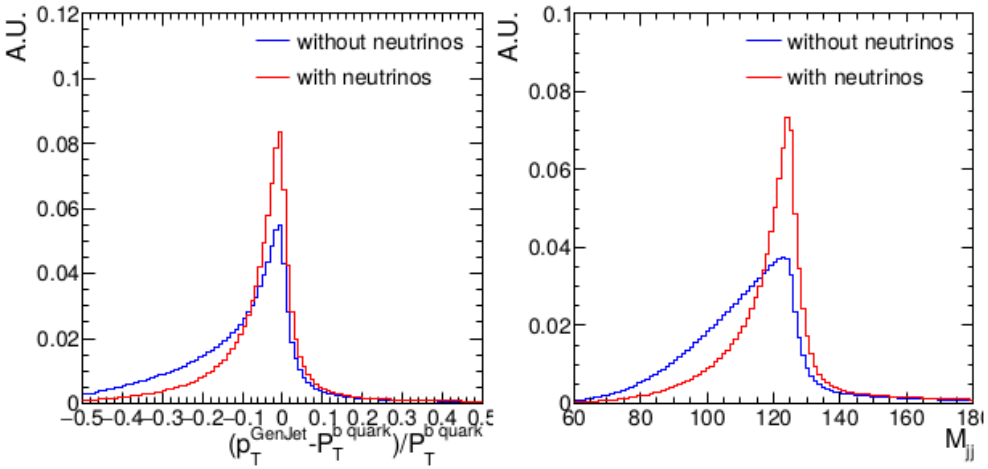
Second jet pT

1.3) Mbb reconstruction

- Mbb reconstruction negatively affected by moving from AK5 (Run I) to AK4 (Run II).



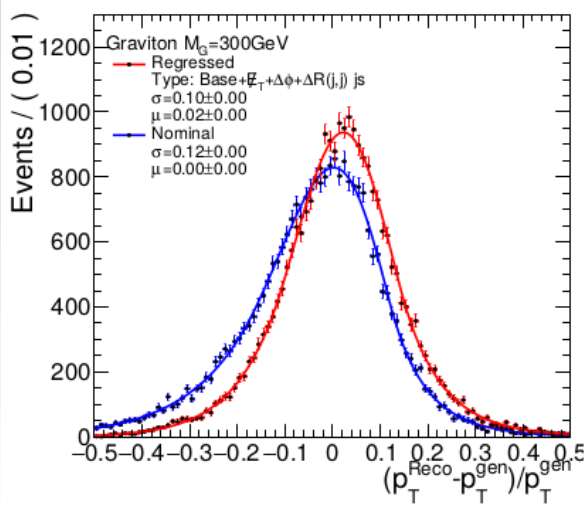
PhD thesis of Rafael Texeira de Lima
http://cds.cern.ch/record/2290218/?ln=de

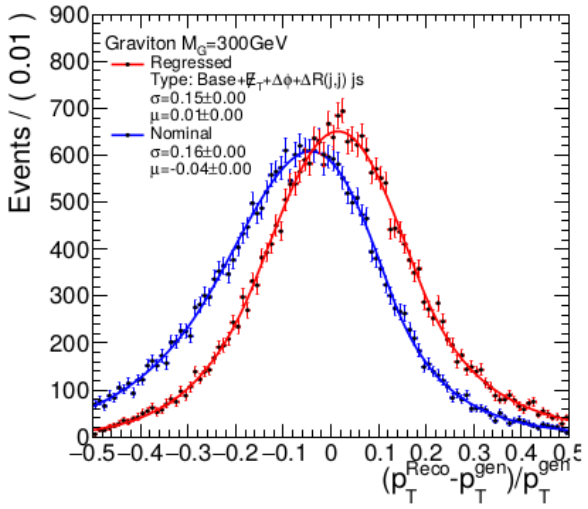


- After b-jets calibration based on JES correction, b-jets still miss energy due to presence of neutrinos in B mesons decays.

1.3) Mbb reconstruction

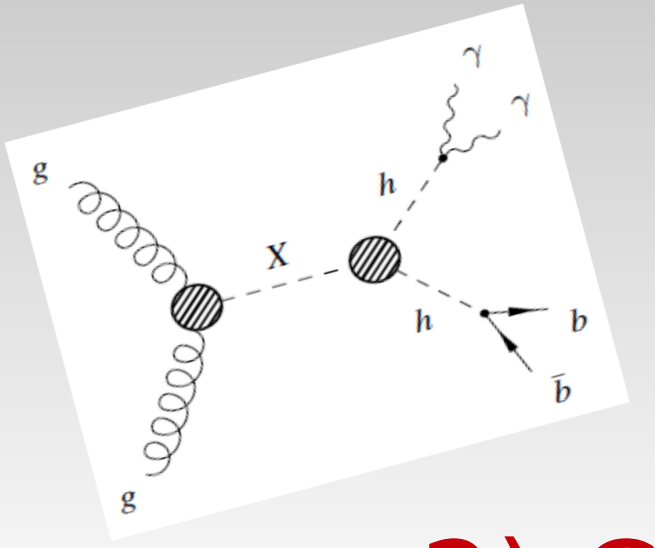
PhD thesis of Rafael Texeira de Lima
http://cds.cern.ch/record/2290218/?ln=de

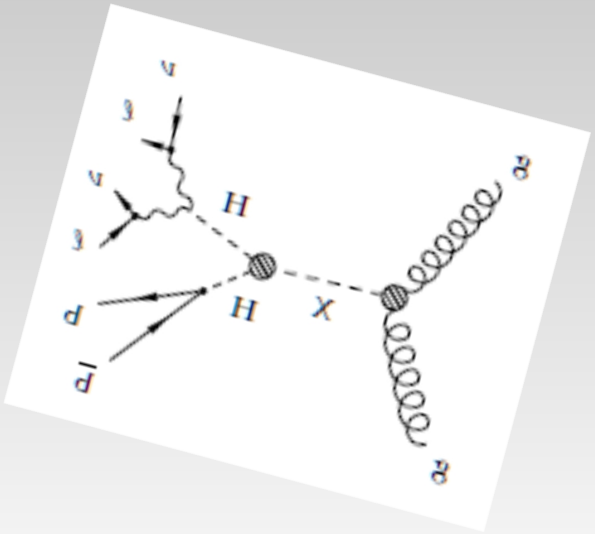




Input variables as in H → $b\bar{b}$ analysis regression:	Jet kinematics: η, p_T, m_T JEC, Neutral hadron energy fraction, Photon energy fraction SecVtxdL, SecVtxdeL, SecVtxPt, SecVtxM, SecVtxNtrk Soft Lepton: p_T, p_T (rel), ΔR Number of jet constituents, p_T (Lead Track), Number of verteces
Additional variables for our analysis:	E_T^{miss}, $\Delta\Phi(Jet, E_T^{miss})$, ΔR(Leading jet, Trailing jet)

→ Idea "b-jet energy regression": extra correction based on
- Track information connected to B-meson decay
- Jet shape information based on PFlow objects.
- Global event information: MET, jets correlation.

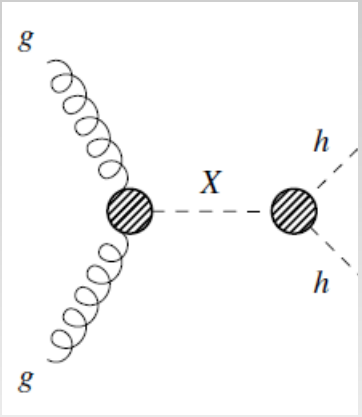




2) Categorisation

Analysis	Region	Classification MVA	$\tilde{M}_X$
Nonresonant	High-mass	HPC: MVA $>$ 0.97 MPC: 0.6 $<$ MVA $<$ 0.97	$\tilde{M}_X > 350$ GeV
	Low-mass	HPC: MVA $>$ 0.985 MPC: 0.6 $<$ MVA $<$ 0.985	$\tilde{M}_X < 350$ GeV
Resonant	$m_X > 600$ GeV	HPC: MVA $>$ 0.5 MPC: 0 $<$ MVA $<$ 0.5	Mass window
	$m_X < 600$ GeV	HPC: MVA $>$ 0.96 MPC: 0.7 $<$ MVA $<$ 0.96	Mass window

2.1) MVA: LHE or LO level

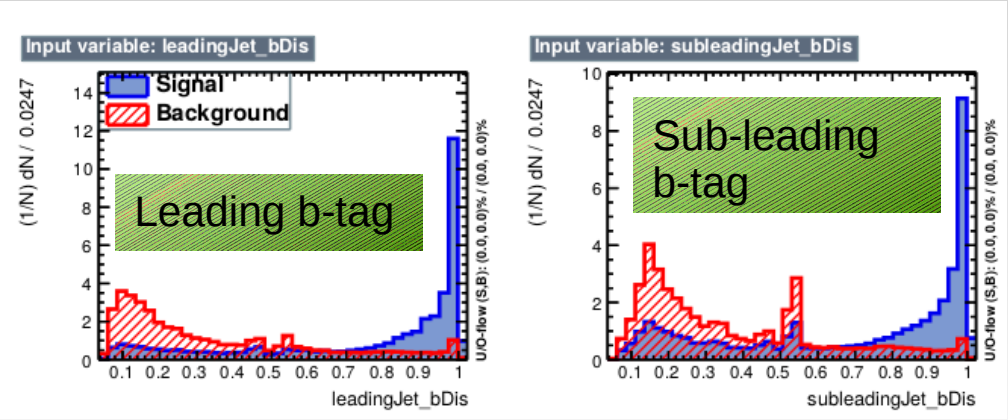


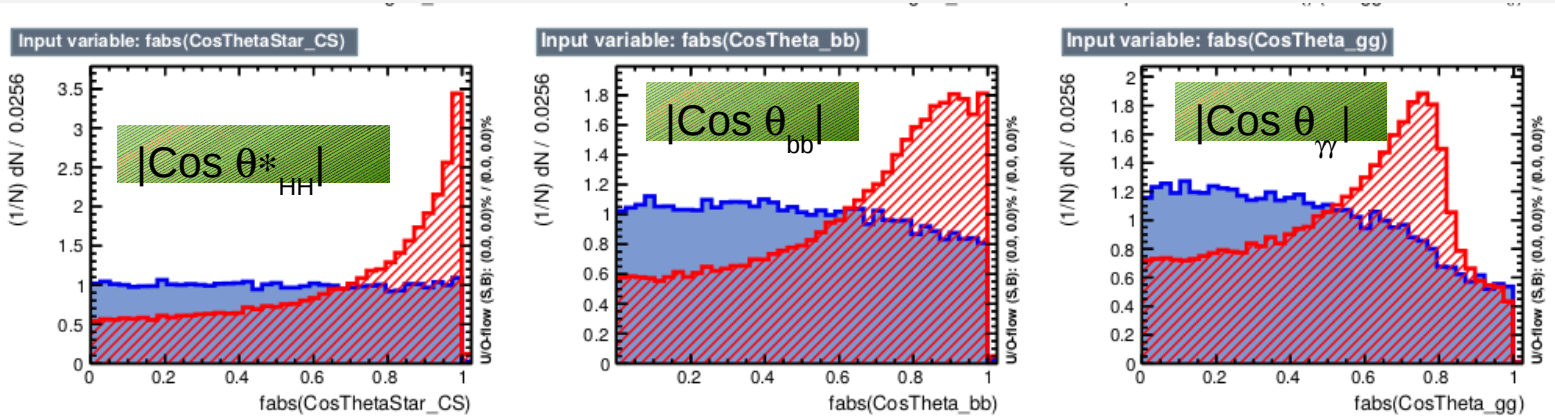
- The HH final state at LO before H decay can be described by 3 relevant degrees of freedom: m_{HH}, cos θ* and $p_{Z,HH}$.
  - m_{HH} = sqrt(shat) encode the ME properties and interplay between different operators.
  - cos θ* describe the helicity properties of the ME. For non-resonant production is mainly s-wave dominated. For resonant production X depends on X spin.
  - $p_{Z,HH}$ is encoded y the PDF properties for gg → HH production. Correlated with m_{HH} through PDFs.

2.2) MVA

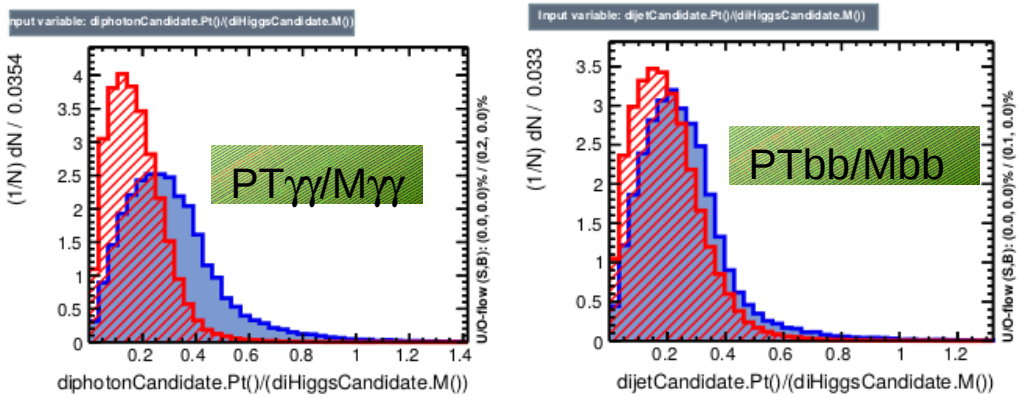
Plots for illustration only

- Weakly correlated variables.
- Jets ordered by btag

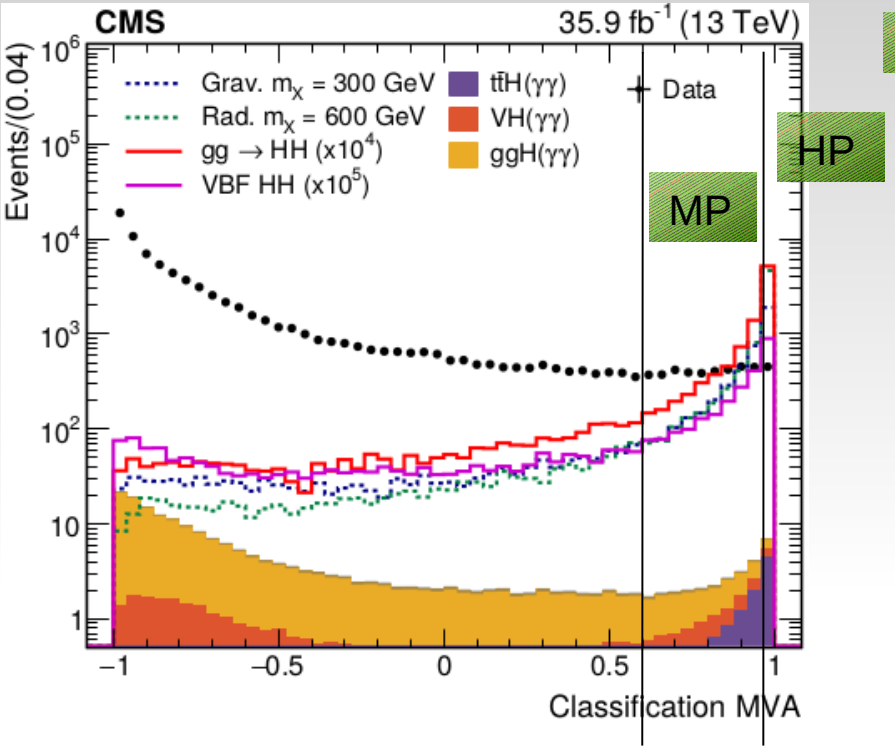


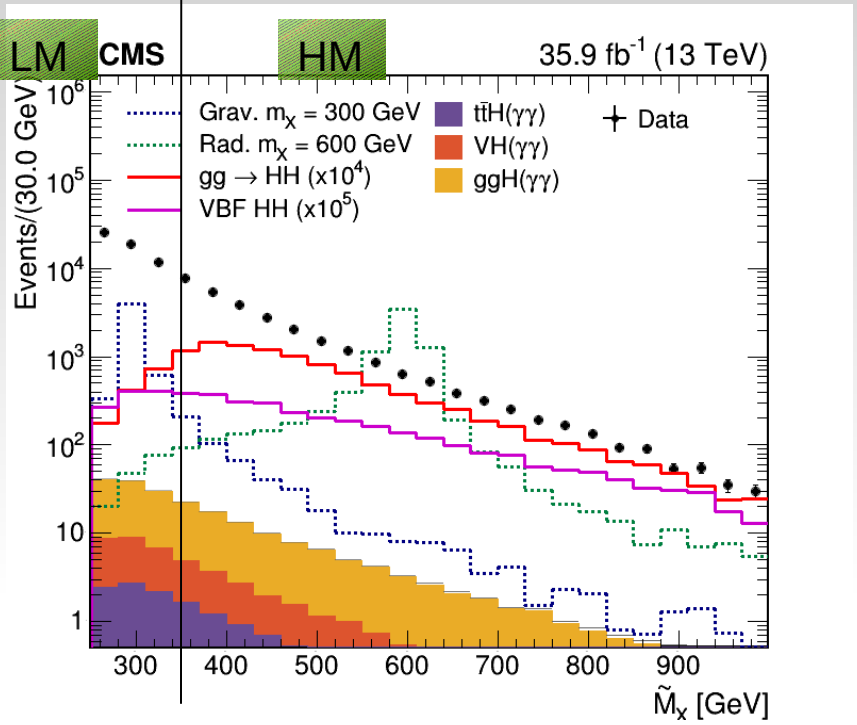


Helicity angles



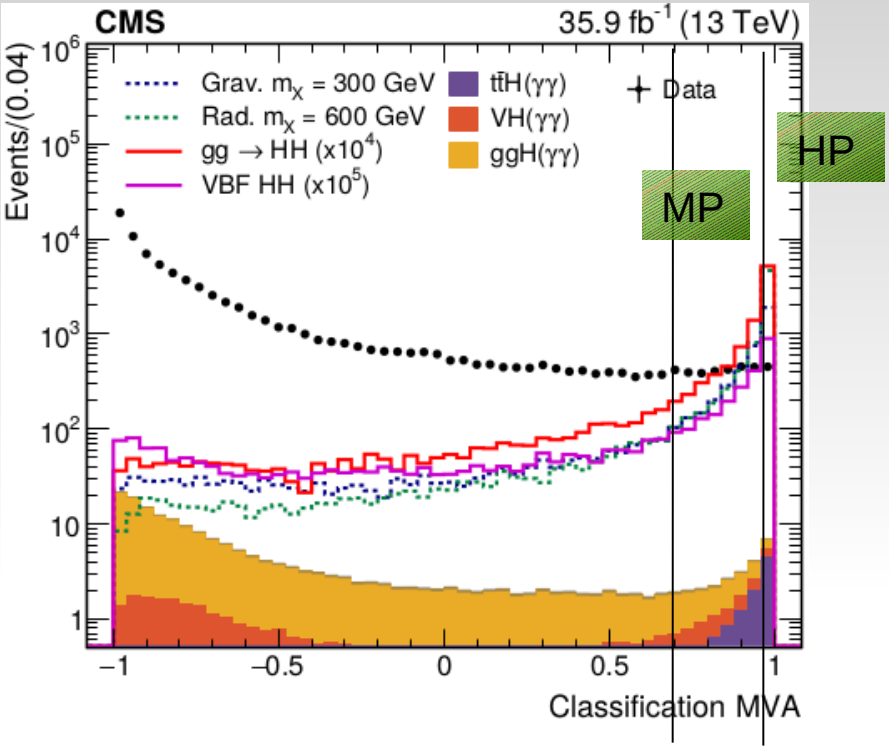
2.4) Non-resonant categorisation

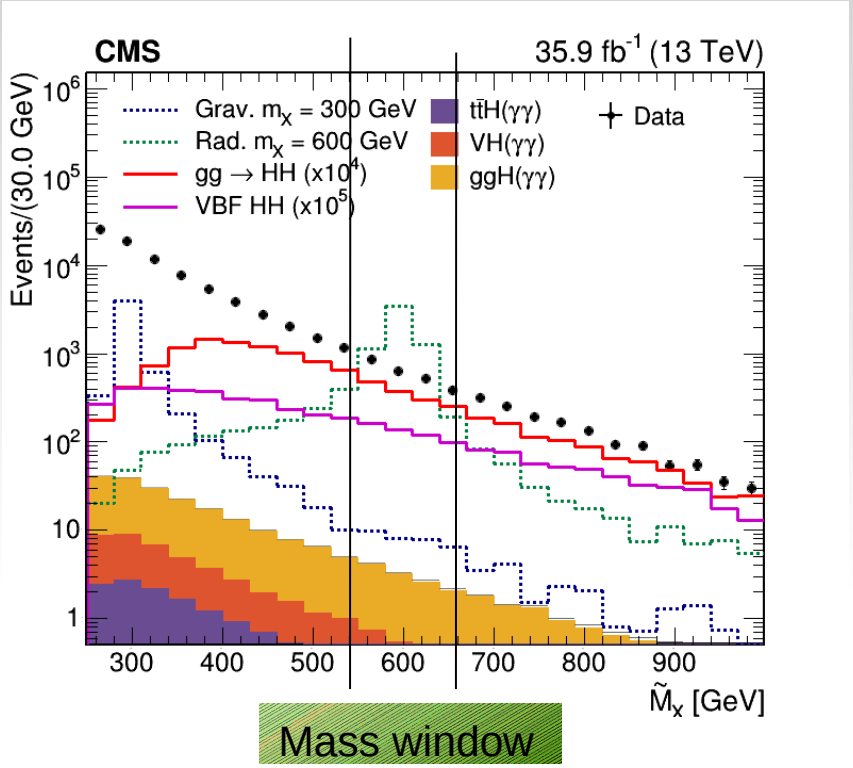




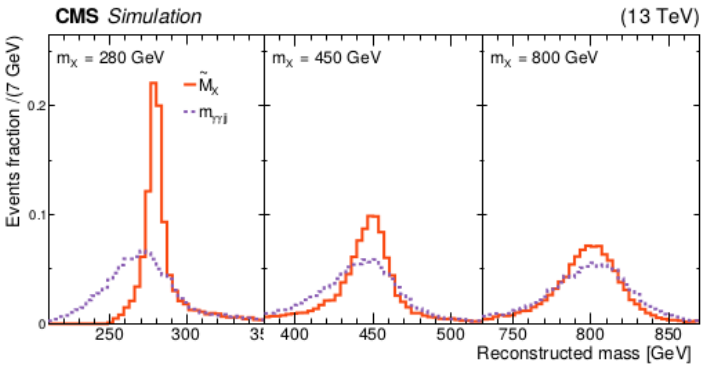
Analysis	Region	Classification MVA	$\tilde{M}_X$
Nonresonant	High-mass	HPC: MVA $>$ 0.97 MPC: 0.6 $<$ MVA $<$ 0.97	$\tilde{M}_X > 350$ GeV
	Low-mass	HPC: MVA $>$ 0.985 MPC: 0.6 $<$ MVA $<$ 0.985	$\tilde{M}_X < 350$ GeV
Resonant	$m_X > 600$ GeV	HPC: MVA $>$ 0.5 MPC: 0 $<$ MVA $<$ 0.5	Mass window
	$m_X < 600$ GeV	HPC: MVA $>$ 0.96 MPC: 0.7 $<$ MVA $<$ 0.96	Mass window

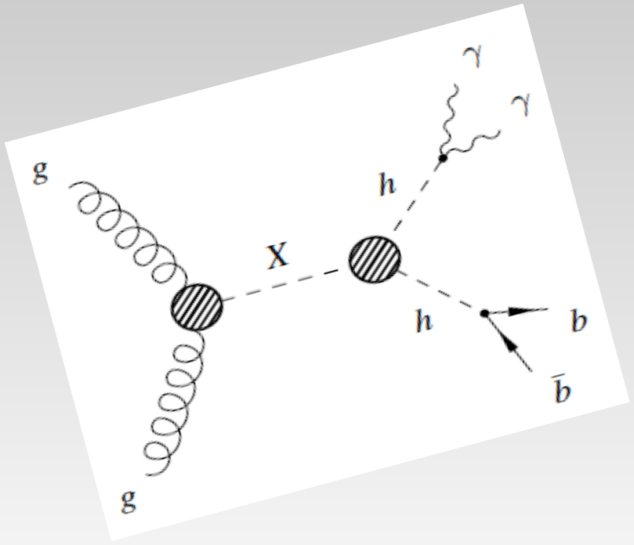
2.5) Resonant categorisation

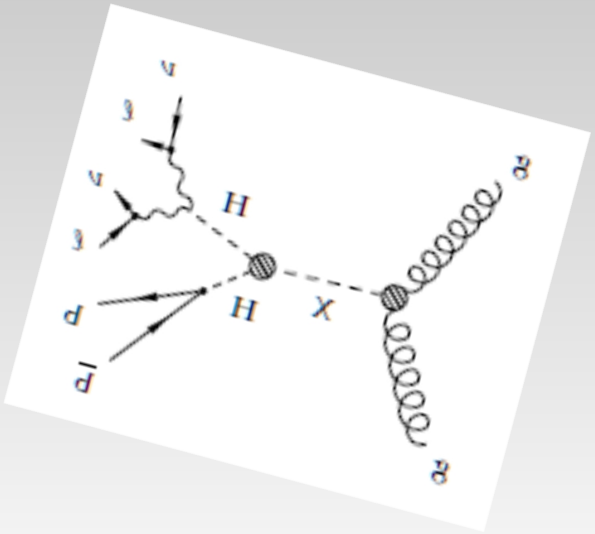




- Use $\tilde{M}_X = m_{\gamma\gamma jj} - (m_{jj} - m_H) - (m_{\gamma\gamma} - m_H)$ variable, mainly important for resonant mass.
- For AK4 jets, M_{tilda}, it appears to have better resolution than $M_{\gamma\gamma bb}$ or kinematic fit.



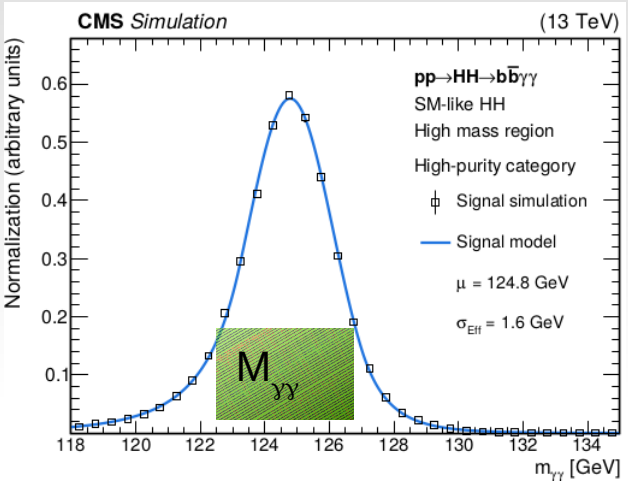


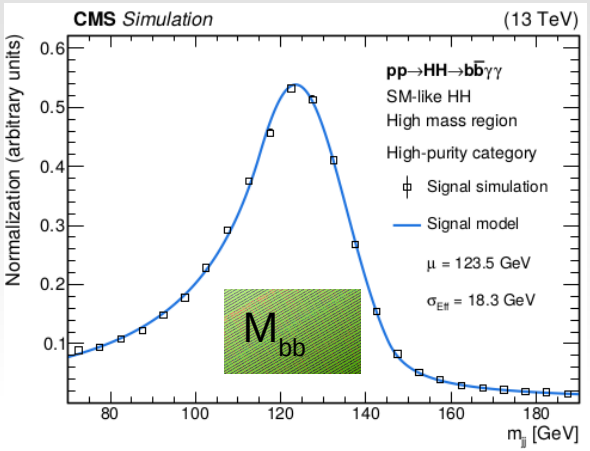


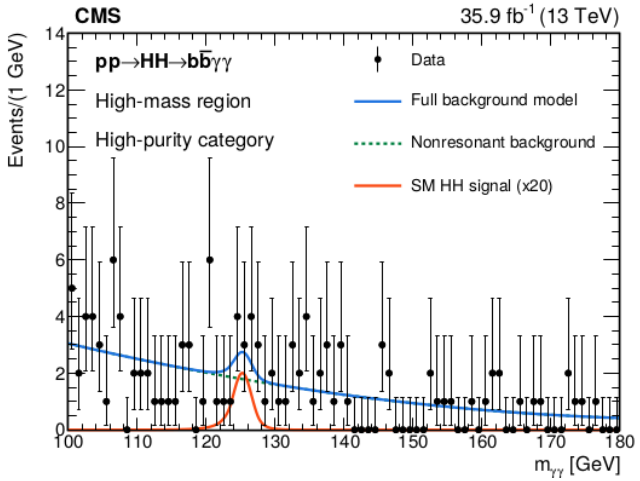
3) Results

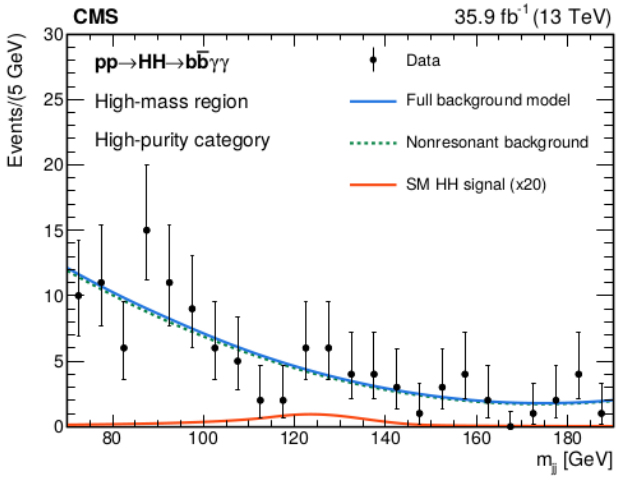
3.1) Analysis method

- 2D analysis. It was verified that within stat uncertainties signal and background shapes are uncorrelated.
- 2D improves compared to 1D by ~ 10%.

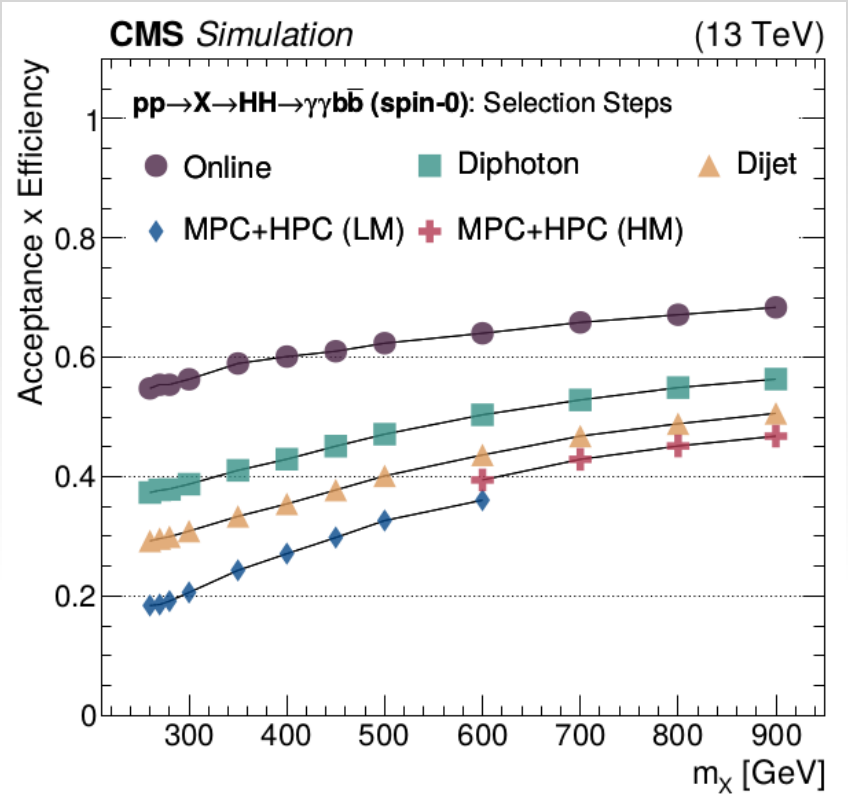


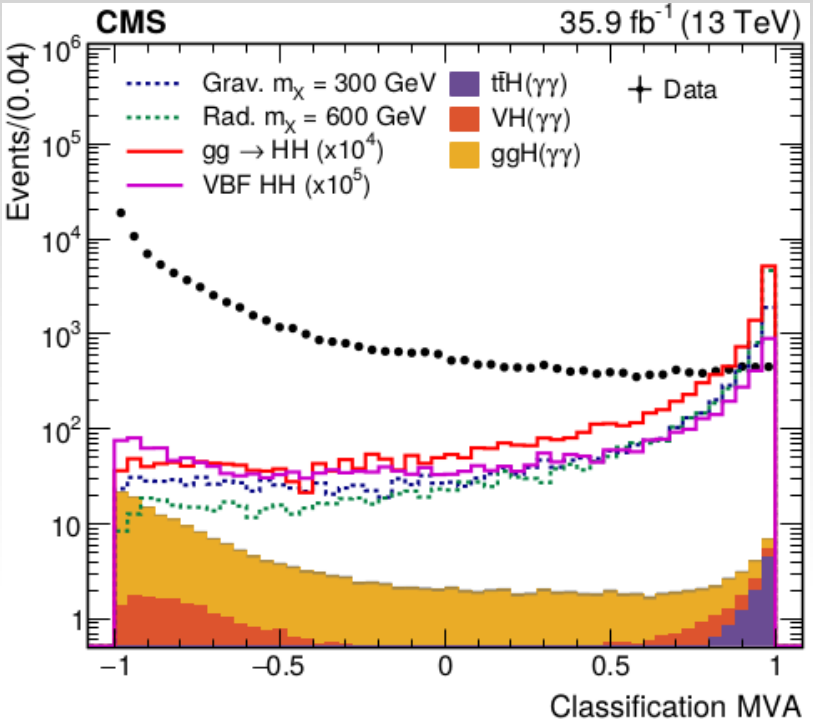






3.2) Efficiencies





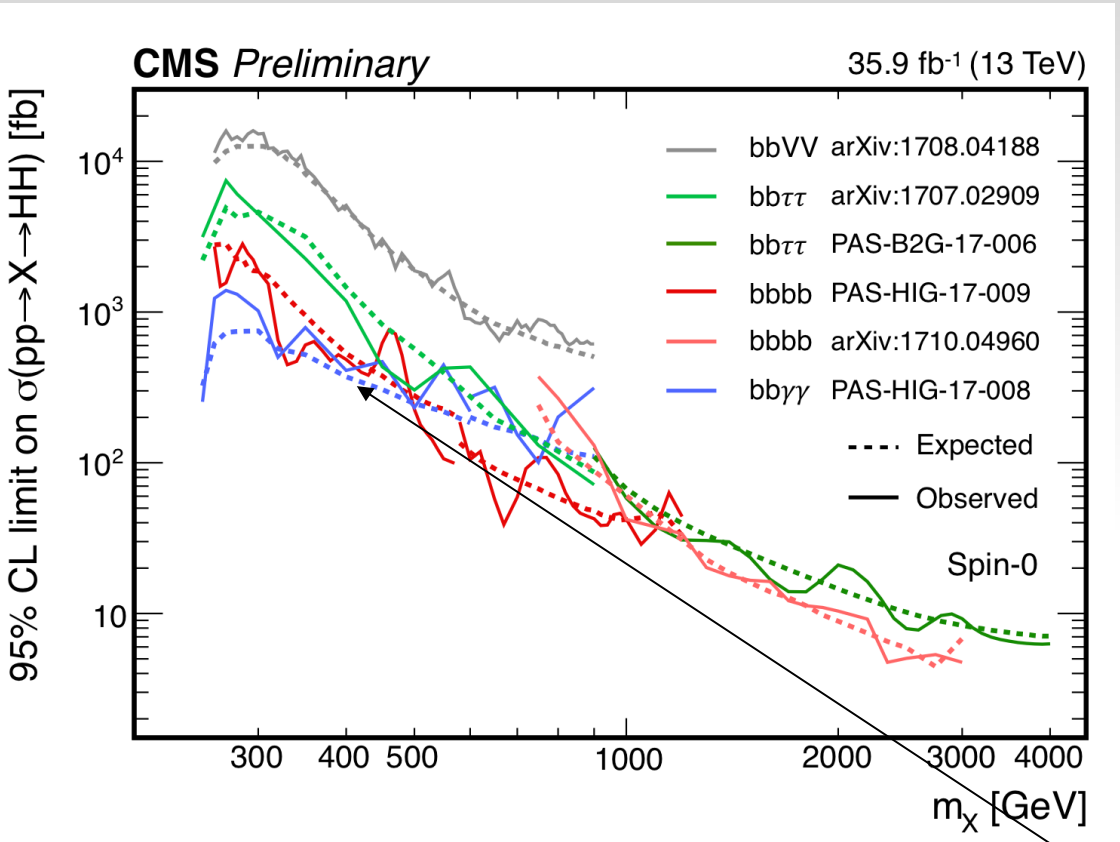
The efficiency times acceptance is 30% (13%) for the SM-like ggHH (VBF HH) signal hypothesis, with 25% (10%) in the high-mass region and 5% (3%) in the low-mass one. The difference between the two production mechanisms mainly comes from the fact that the MVA was trained assuming ggHH signal.

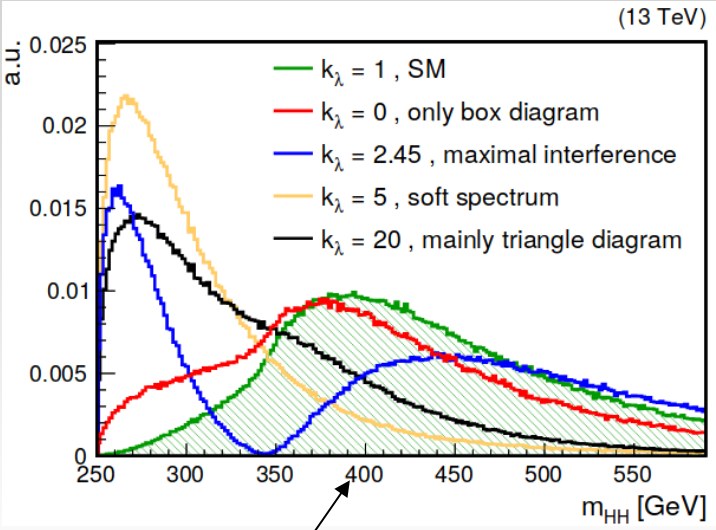
3.3) Systematics

Table 4: Summary of systematic uncertainties.

Sources of systematic uncertainties	Type	Value (%)
Integrated luminosity	Normalization	2.5
Photon related uncertainties		
Diphoton selection (with trigger uncertainties and PES)	Normalization	2.0
Photon identification	Normalization	1.0
PES ($\frac{\Delta m_{\gamma\gamma}}{m_{\gamma\gamma}}$)	Shape	0.5
PER ($\frac{\Delta \sigma_{\gamma\gamma}}{\sigma_{\gamma\gamma}}$)	Shape	5.0
Jet related uncertainties		
Dijet selection (JES+JER)	Normalization	0.5
JES ($\frac{\Delta m_{jj}}{m_{jj}}$)	Shape	1.0
JER ($\frac{\Delta \sigma_{jj}}{\sigma_{jj}}$)	Shape	5.0
Resonant analysis specific uncertainties		
Mass window selection (JES+JER)	Normalization	3.0
Classification MVA – b tagging (HPC)	Normalization	11–19
Classification MVA – b tagging (MPC)	Normalization	3–9
Nonresonant analysis specific uncertainties		
$\widetilde{M}_X$ Classification	Normalization	0.5
Classification MVA – b tagging (HPC)	Normalization	10–18
Classification MVA – b tagging (MPC)	Normalization	3–9
Theoretical uncertainties in the SM single-Higgs boson production		
QCD missing orders (ggH, VBF H , VH, $t\bar{t}H$)	Normalization	0.4–5.8
PDF and α_S uncertainties (ggH, VBF H , VH, $t\bar{t}H$)	Normalization	1.6–3.6
Theoretical uncertainty $b\bar{b}H$	Normalization	20
Theoretical uncertainties in the SM HH boson production		
QCD missing orders	Normalization	4.3–6
PDF and α_S uncertainties	Normalization	3.1
m_t effects	Normalization	5

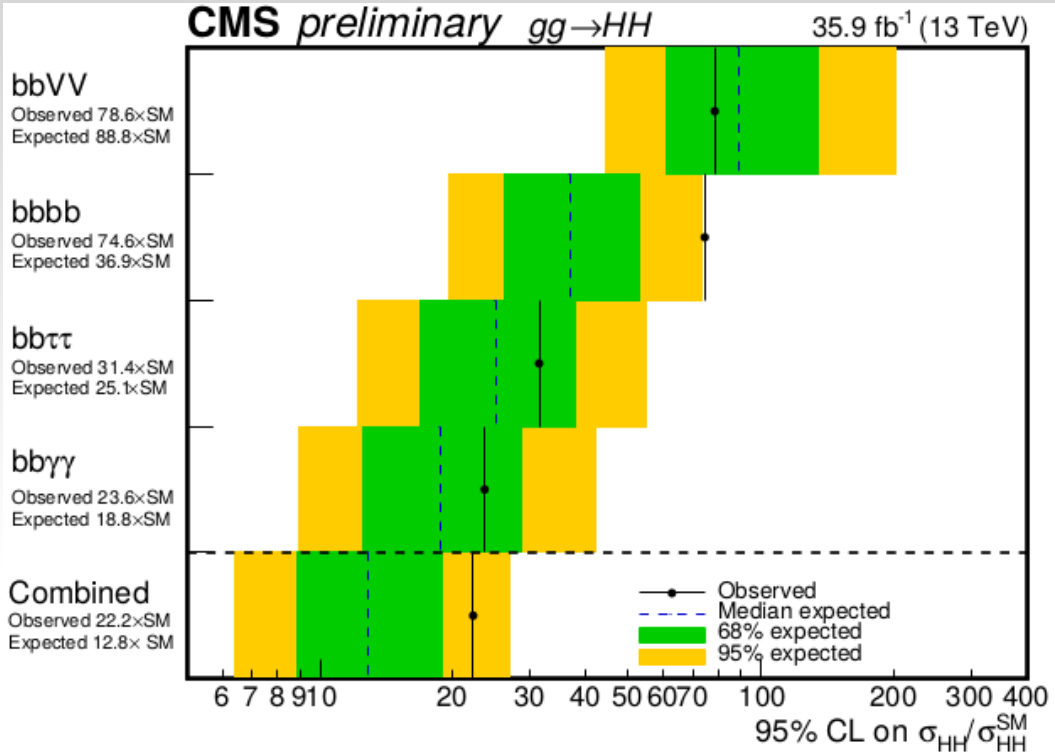
3.4) SM-like “the magic spot”





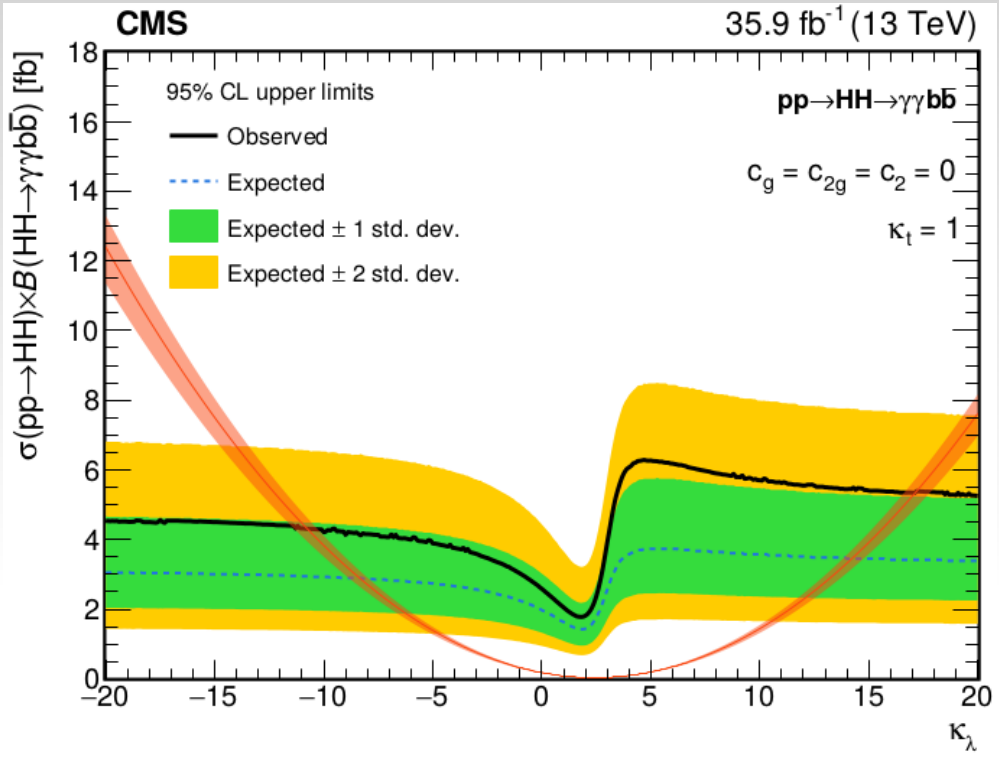
- The peak in SM HH spectrum is at 400 GeV.
- This is the place where are crossing the sensitivities of 2 or best 3 channels.
- The measurement of SM HH→ 2b2g at CMS is in fact a large team effort.

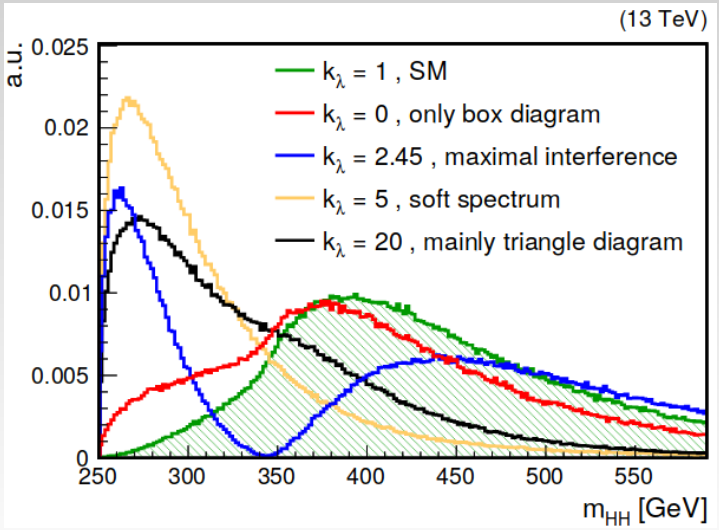
3.5) SM-like results



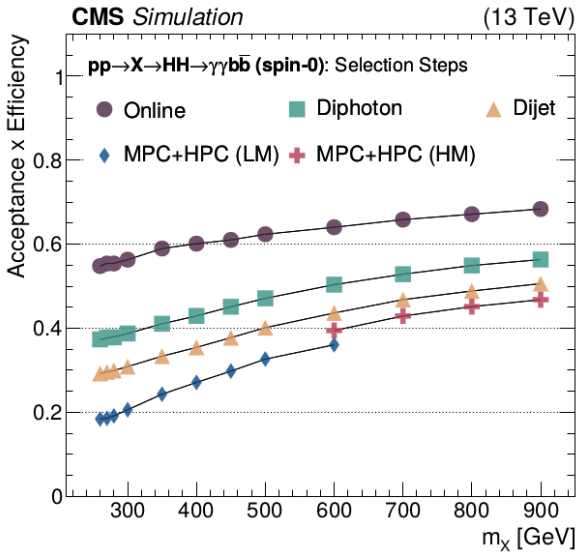
- The γγbb channel is the most sensitive for SM production.
- We checked that simple add-on of VBF sample (4% of the total cross section) improves the SM-like sensitivity by 1.3%. Need more work to take maximal advantage of VBF.

3.6) κ_λ scan

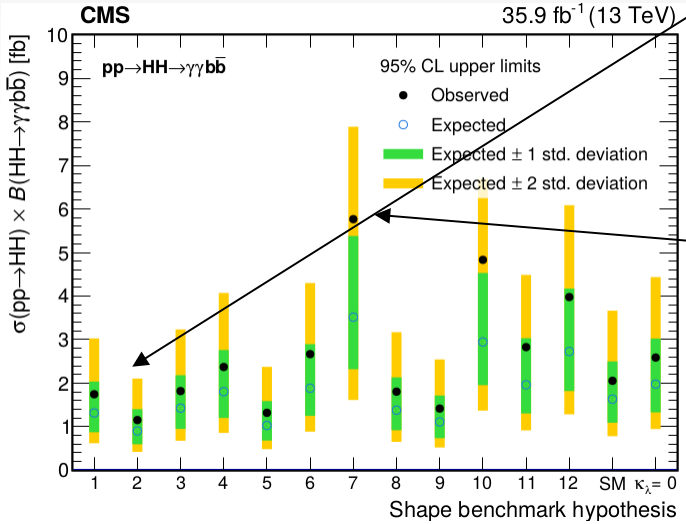


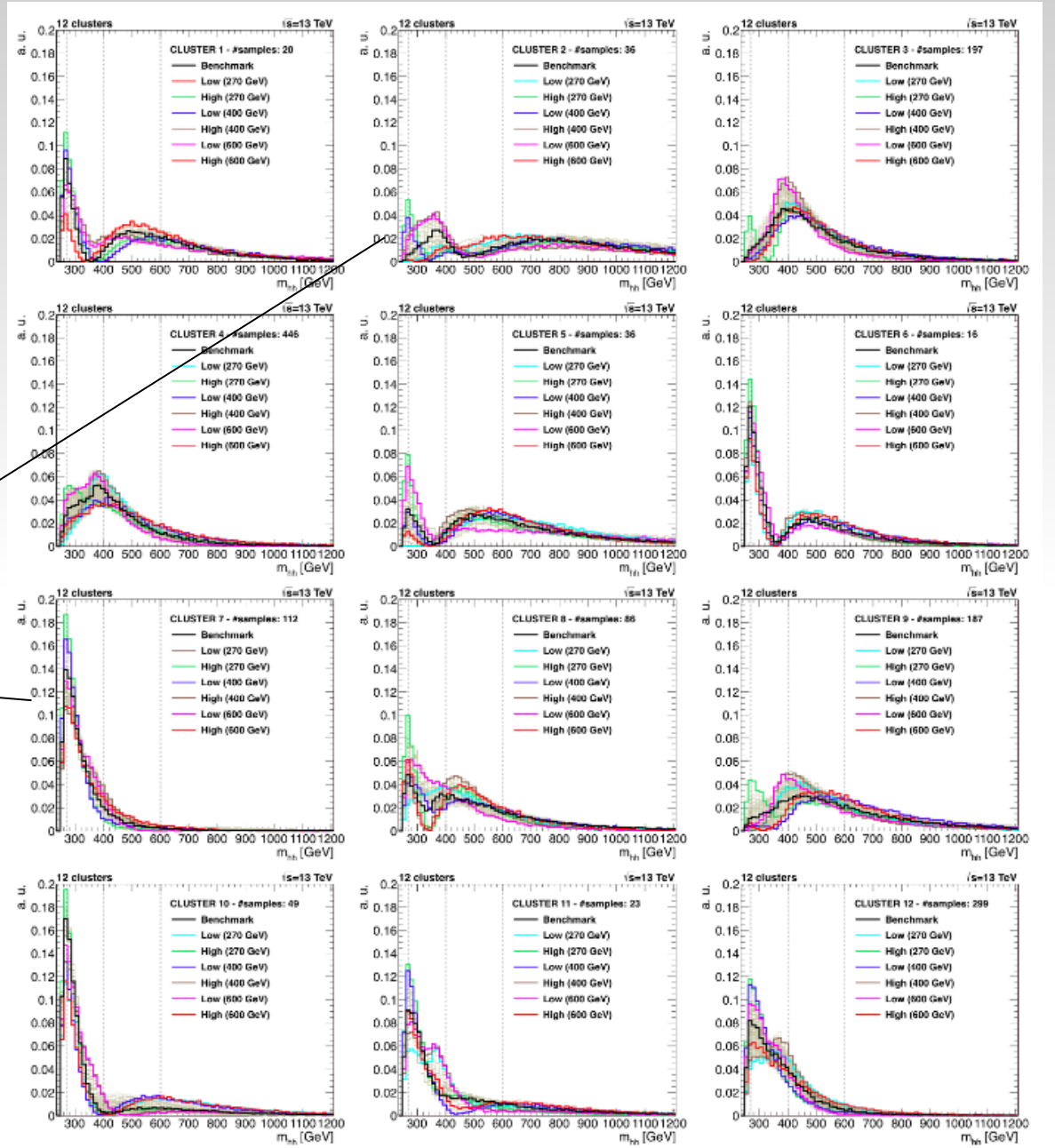


- The sensitivity is the best close to the maximal interference.
- Purely signal efficiency effect (see M_X efficiency figure).



3.7) Benchmark limits





Conclusions

- This channel is the most sensitive at low mX at CMS.
- It is also the most sensitive for SM-non resonant production even if ττbb is quite close.
- It is the most sensitive for k_λ scan.

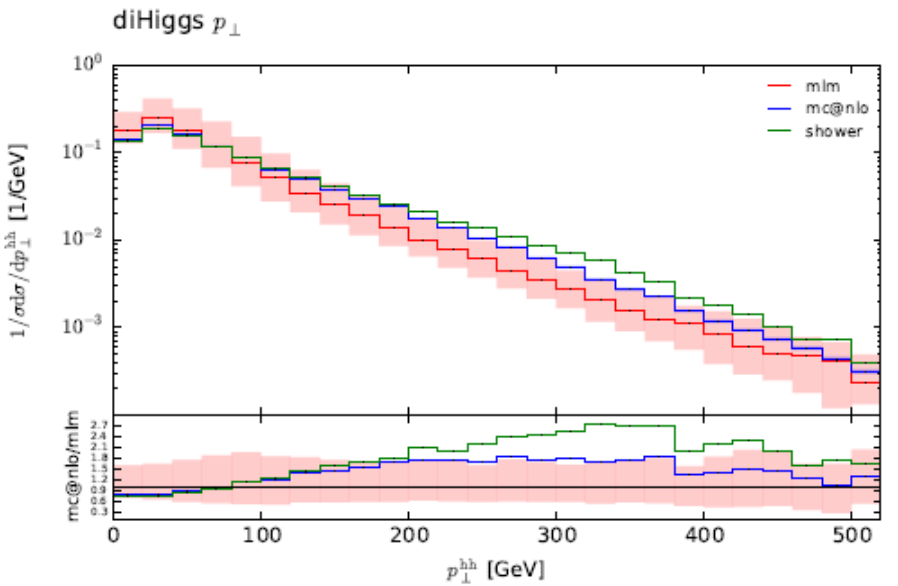
- There are interesting ideas how to improve it using more subtle BDT for classification.

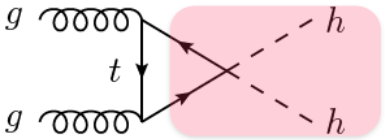
- Extremely important for this analysis to keep the $M_{\gamma\gamma}$ resolution as good as possible.

BACKUP

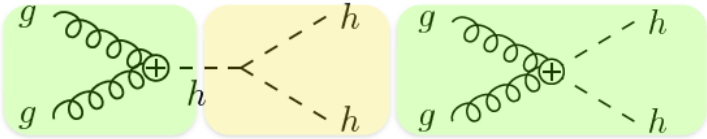
2.3) MVA: Additionnal order effects

- N(...)LO QCD effects:
  - Presence of additional IRS jets:
    - Described by extra variables such as
      - H_T-$H_{T,\,xxyy}$, p_T, η of extra jets.
      - Reduced FSR activity (gap) if x or y have no color charge.
  - Modification of HH kinematics:
    - $p_{T,HH} > 0$, $\Delta\varphi_{HH}$!= π.
    - Characteristic of color charge of the production mechanism.
- Be careful:
  - Variables description depends on the accuracy of the MC prediction.
  - May be not so different between signal and background since the Initial State color structure may be rather similar.



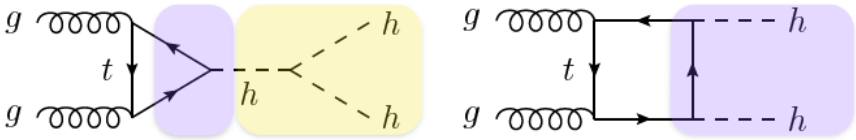


ttHH non-linear interaction



Higgs-gluon contact interactions

4) Exploration of HH phase-space in EFT



SM diagrams

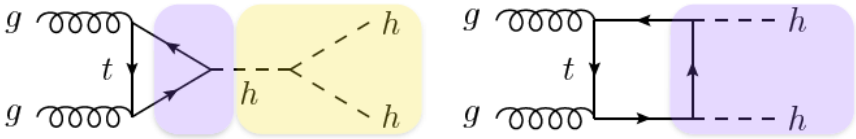
4.1) Non-resonant HH production: EFT and BSM

The relevant lagrangian terms of gg→HH production in D=6 EFT

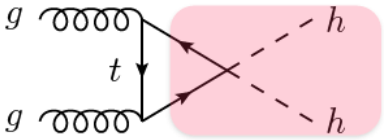
$$\mathcal{L}_{hh} = -\frac{m_h^2}{2v}\left(1 - \frac{3}{2}c_H + c_6\right)h^3 + \frac{\alpha_s c_g}{4\pi}\left(\frac{h}{v} + \frac{h^2}{2v^2}\right)G^a_{\mu\nu}G_a^{\mu\nu}$$
$$- \left[\frac{m_t}{v}\left(1 - \frac{c_H}{2} + c_t\right)\bar{t}_L t_R h + \text{h.c.}\right] - \left[\frac{m_t}{v^2}\left(\frac{3c_t}{2} - \frac{c_H}{2}\right)\bar{t}_L t_R h^2 + \text{h.c.}\right]$$

arXiv:1410.3471

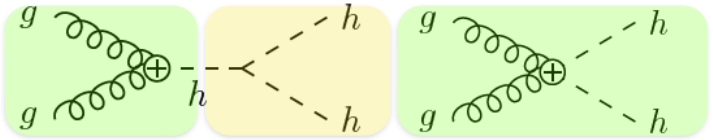
Non SM Yukawa coupling is not considered



SM diagrams



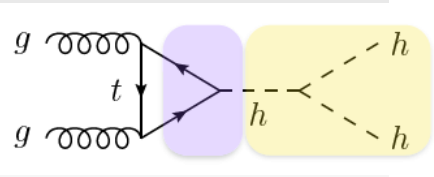
ttHH non-linear interaction

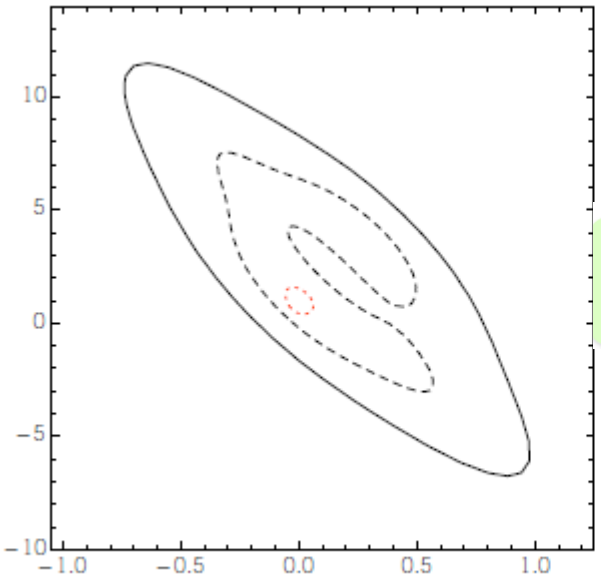


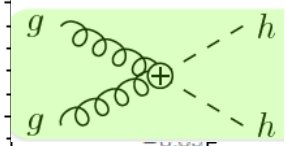
Higgs-gluon contact interactions

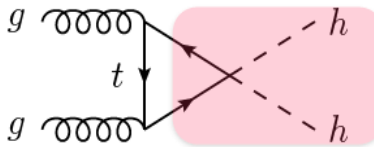
- Five D6 operators for HH sector.

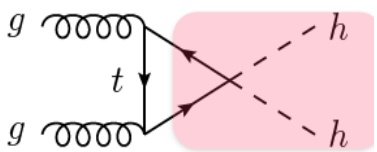
4.1) Non-resonant HH production: EFT and BSM











	LHC_{14}	HL-LHC	FCC_{100}
$\sqrt{s}$	14 TeV	14 TeV	100 TeV
Luminosity	$L = 300\,\mathrm{fb}^{-1}$	$L = 3\,\mathrm{ab}^{-1}$	$L = 3\,\mathrm{ab}^{-1}$

arxiv:1502.00539

- Idea: set limits in 5-D parameters space and constraint progressively the phase-space with ellipses converging (if no new physics observed) to SM point.

4.2) We need: predicted cross section

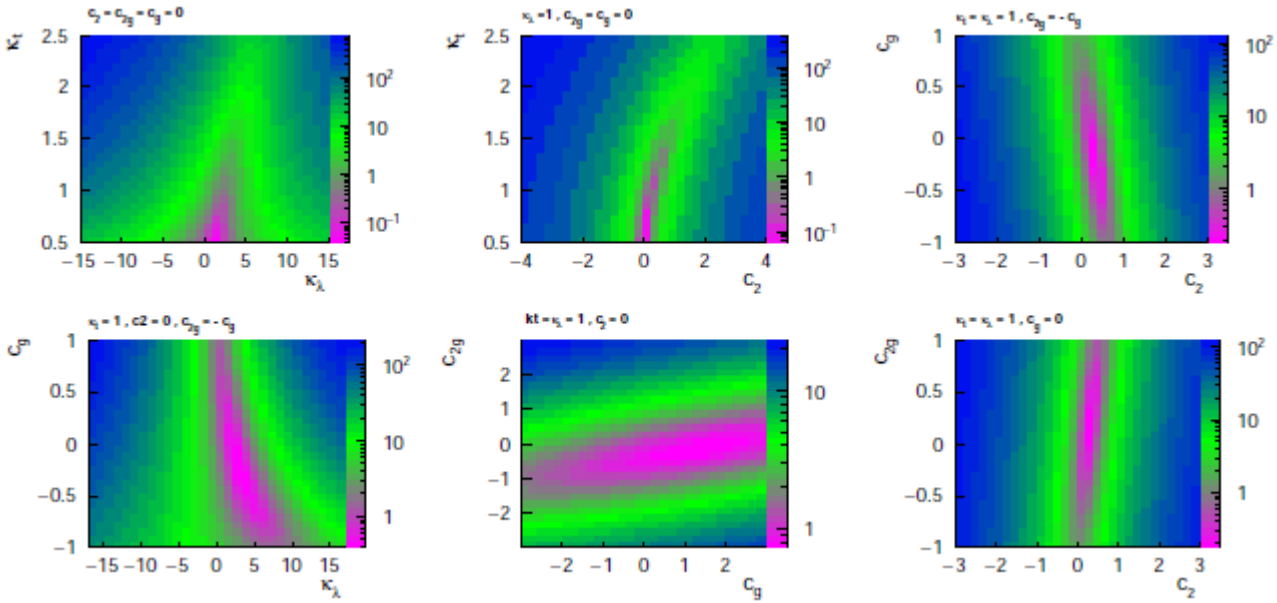
- Within some approximation (top loop predominant contribution)

k = NNLO+NNLL/LO is expected to be similar within 5% to the one of SM.

$$\sigma_{HH} = \sigma^{SM}_{HH,NNLO+NNLL} \cdot R_{HH}$$

YR4
arXiv:1608.06578

$$R_{HH} \equiv \frac{\sigma_{HH}}{\sigma^{SM}_{HH}} \overset{LO}{=} A_1\,\kappa_t^4 + A_2\,c_2^2 + (A_3\,\kappa_t^2 + A_4\,c_g^2)\,\kappa_\lambda^2 + A_5\,c_{2g}^2 + (A_6\,c_2 + A_7\,\kappa_t\kappa_\lambda)\kappa_t^2 + (A_8\,\kappa_t\kappa_\lambda + A_9\,c_g\kappa_\lambda)c_2 + A_{10}\,c_2c_{2g} + (A_{11}\,c_g\kappa_\lambda + A_{12}\,c_{2g})\,\kappa_t^2 + (A_{13}\,\kappa_\lambda c_g + A_{14}\,c_{2g})\,\kappa_t\kappa_\lambda + A_{15}\,c_gc_{2g}\kappa_\lambda\,.$$

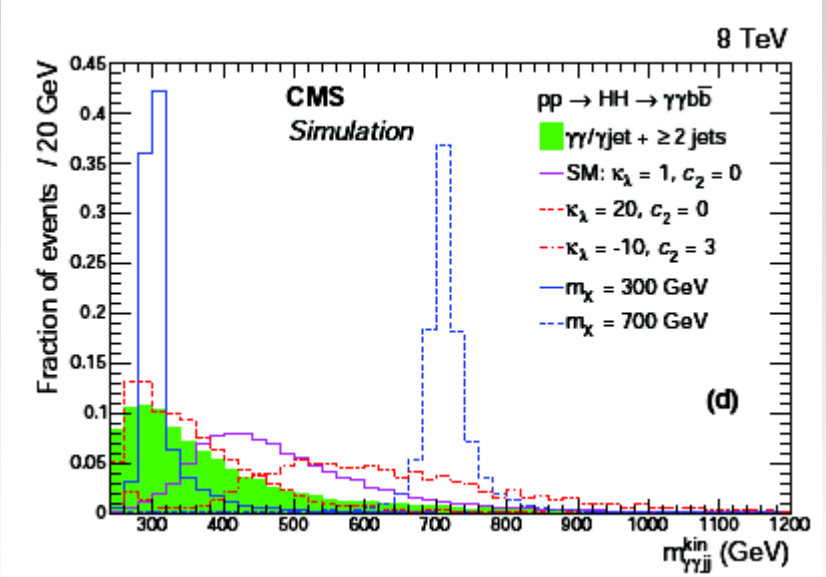


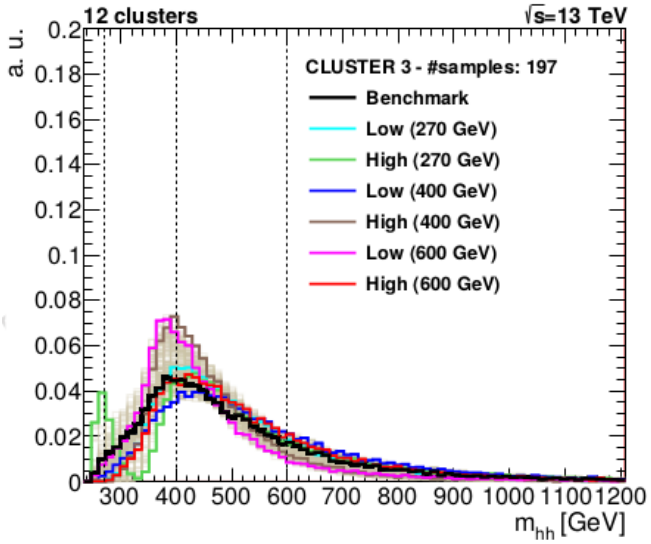
4.3) We need: kinematics.

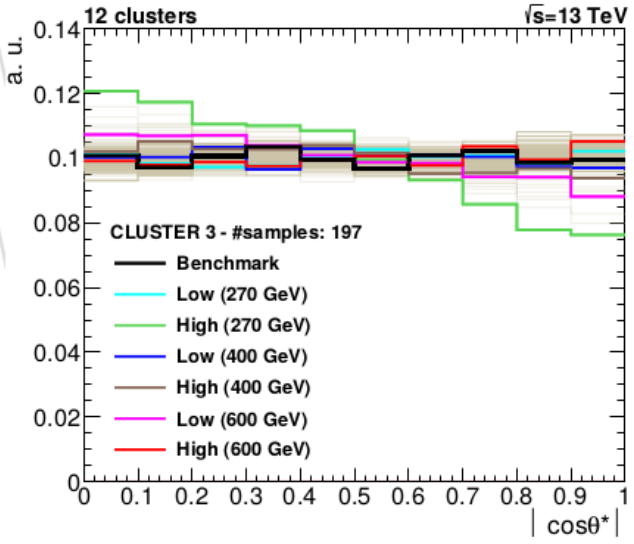
- When you change parameters you change interference pattern. So you change M_{HH} shape.

  And analyses are very sensitives to M_{HH} shape.

  Remember λ scans.
- Also cos θ^* is also interference dependent but mainly dominated by s-wave.







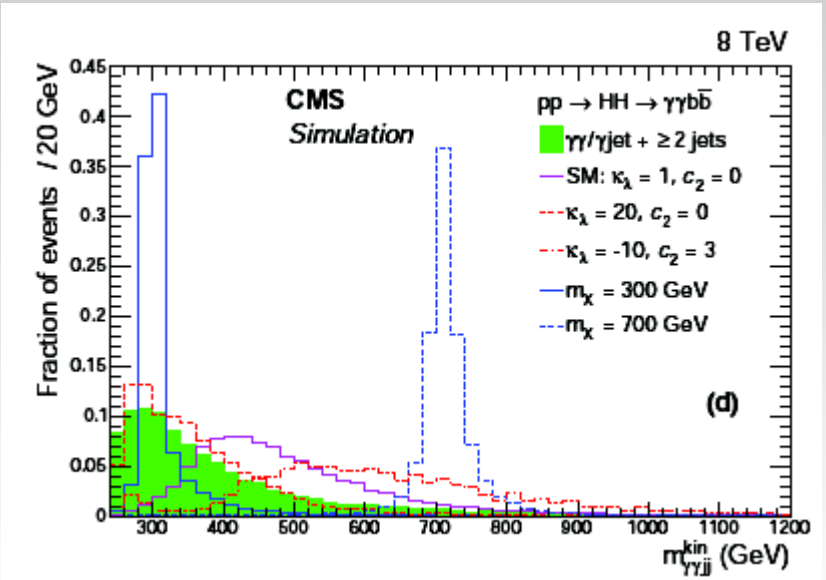
4.3) We need: kinematics.

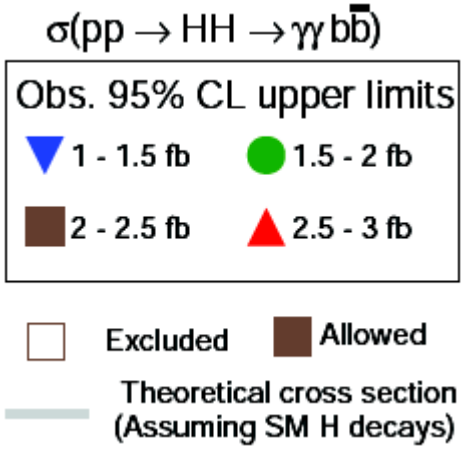
- When you change parameters you change interference pattern. So you change M_{HH} shape.

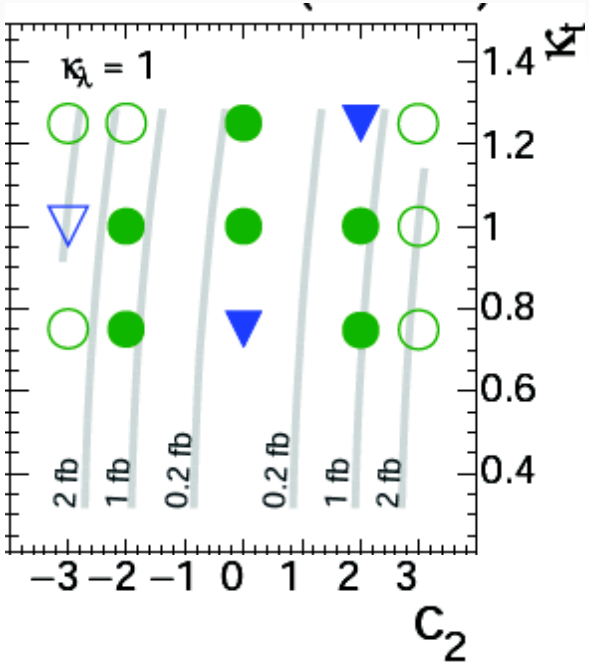
And analyses are very sensitives to M_{HH} shape.

Remember λ scans...

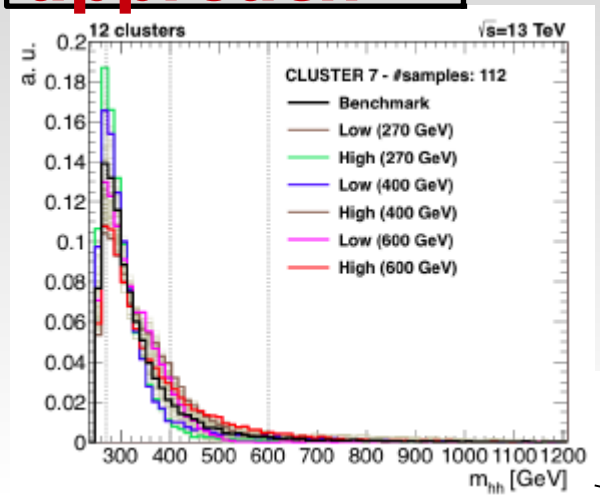
- Pedestrian method:
  - Generate a grid of BSM point. Fully Simulate them and put limits.
  - Was tried with 3 parameters in γγbb at 8 TeV (128 samples).
  - Very computing and time intensive. Impossible to refine the grid a posteriori. Impossible in 5D.

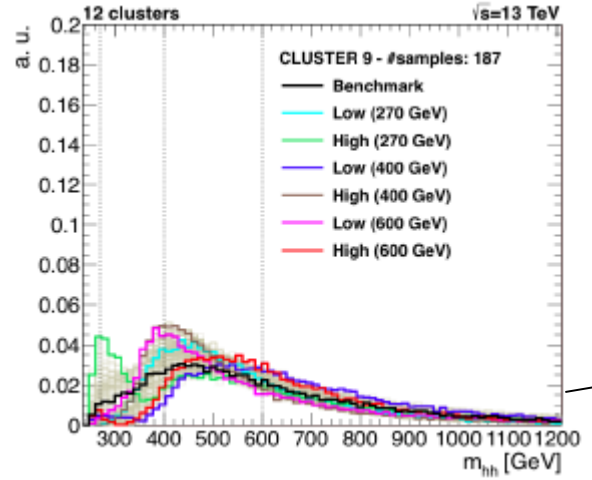




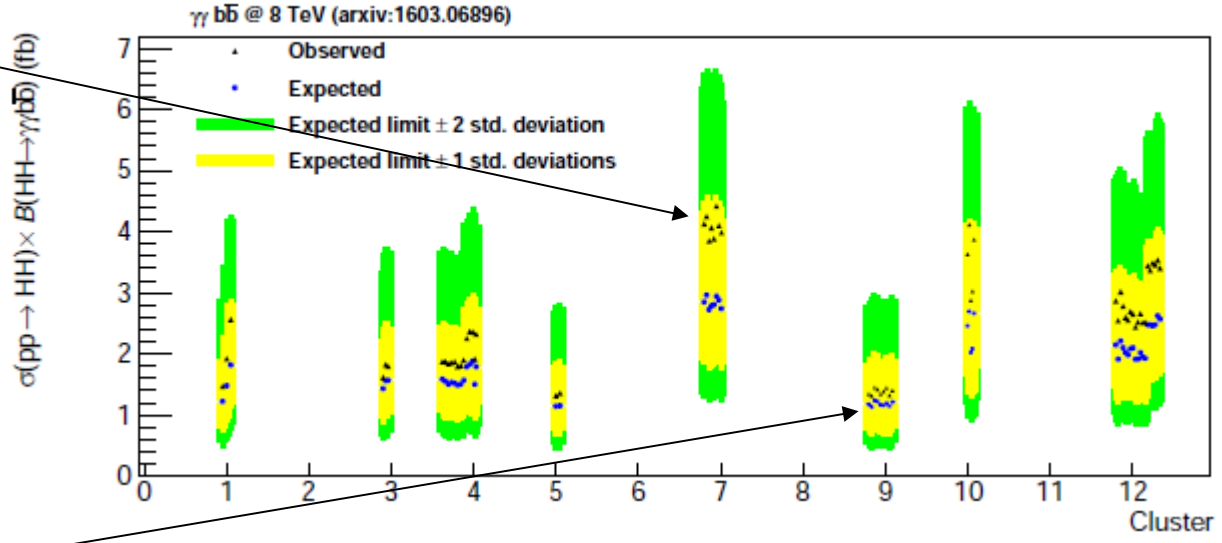


4.4) Benchmark approach





- Idea: use clustering mathematical method to classify the EFT points per shape:
  - Only 12 representatives shapes needed.
  - Optimize the analysis for each of those shapes.
  - Any EFT kinematics can be obtained by Generator reweighing of those shapes.
  - Being smart we can obtain 1500 limits with only 12 generated samples!!!



arXiv:1507.02245
arXiv:1608.06578

4.4) Benchmark approach

B.	1	2	3	4	5	6	7	8	9	10	11	12	SM	$\kappa_\lambda = 0$
κ_λ	7.5	1.0	1.0	-3.5	1.0	2.4	5.0	15.0	1.0	10.0	2.4	15.0	1.0	0.0
κ_t	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0
c_2	-1.0	0.5	-1.5	-3.0	0.0	0.0	0.0	0.0	1.0	-1.0	0.0	1.0	0.0	0.0
c_g	0.0	-0.8	0.0	0.0	0.8	0.2	0.2	-1.0	-0.6	0.0	1.0	0.0	0.0	0.0
c_{gg}	0.0	0.6	-0.8	0.0	-1.0	-0.2	-0.2	1.0	0.6	0.0	-1.0	0.0	0.0	0.0