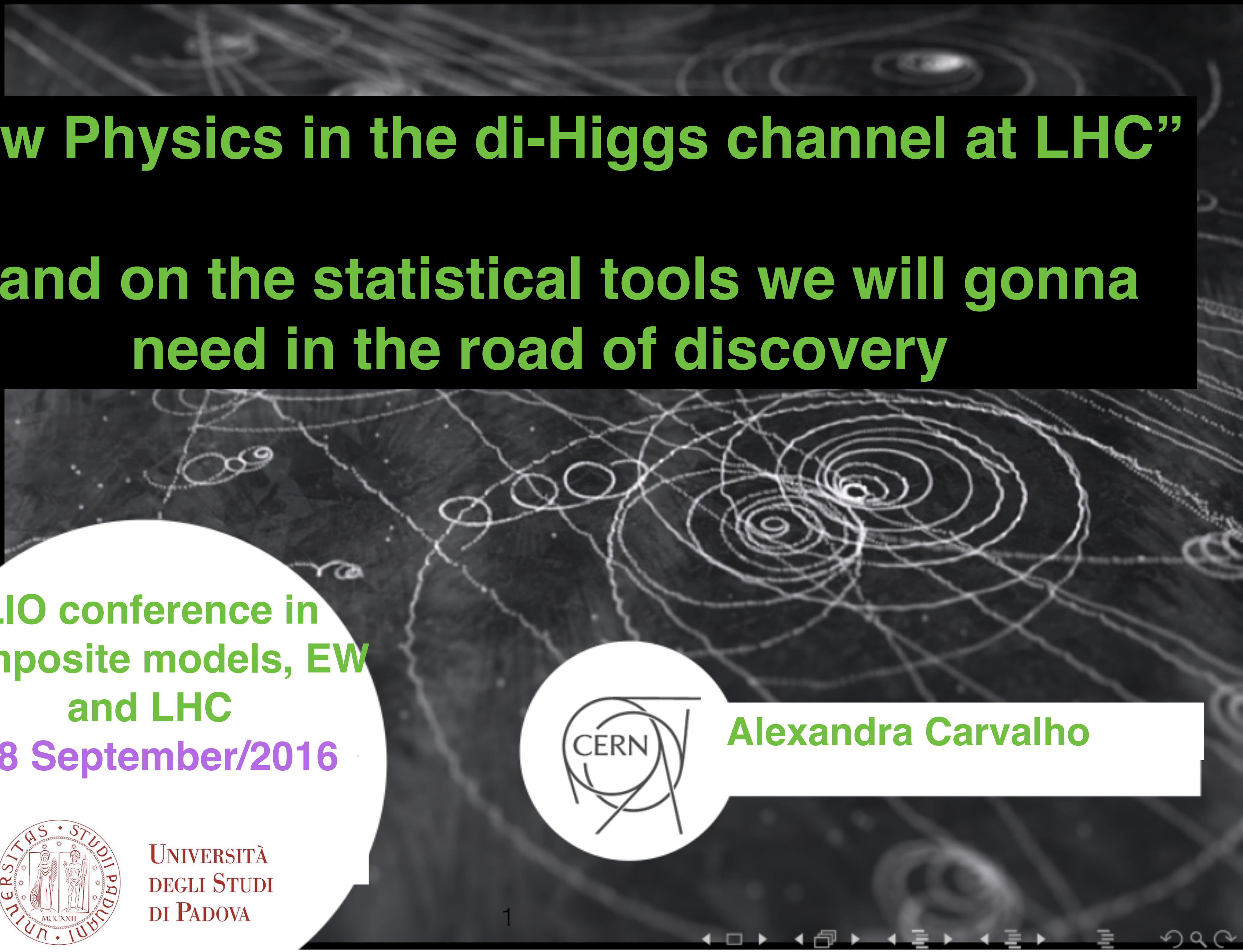


“New Physics in the di-Higgs channel at LHC”

…and on the statistical tools we will gonna need in the road of discovery

LIO conference in composite models, EW and LHC

5-8 September/2016

UNIVERSITÀ DEGLI STUDI DI PADOVA

CERN

Alexandra Carvalho

Outline

- Non resonant diHiggs production: Theory context
- Experimental context
- The clustering technique
- Towards general EFT interpretation
- Going even further: (no!) recast of explicitly models

Theory context:

The role of the Higgs boson in the Standard Model => ElectroWeak Symmetry Breaking

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad \mathcal{L}_{\rm scalar} = \partial_\mu \Phi^\dagger \, \partial^\mu \Phi - V(\Phi^\dagger \Phi) \ ,$$

$$\boxed{V(\Phi^\dagger\Phi) = \mu^2 \, \Phi^\dagger\Phi + \lambda \, (\Phi^\dagger\Phi)^2}$$

given that we had discovered that the Higgs boson mass is ~125 GeV,

if \lambda = 0.012 <u>and the SM particles and forces are all that exists</u>:

Voila! We 'predict' the masses of the foton, W and Z bosons. Keeping an elegant and renormalizable formulation for their interactions…

… fixing some more parameters we 'predict' the masses also of all the known fermions.

Theory context:

The role of the Higgs boson in the Standard Model => ElectroWeak Symmetry Breaking

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad \mathcal{L}_{\rm scalar} = \partial_\mu \Phi^\dagger \, \partial^\mu \Phi - V(\Phi^\dagger \Phi) \,,$$

$$\boxed{V(\Phi^\dagger\Phi) = \mu^2 \, \Phi^\dagger\Phi + \lambda \, (\Phi^\dagger\Phi)^2}$$

given that we had discovered that the Higgs boson mass is ~125 GeV,

if \lambda = 0.012 <u>and the SM particles and forces are all that exists</u>:

Voila! We 'predict' the masses of the foton, W and Z bosons. Keeping an elegant and renormalizable formulation for their interactions…

… fixing some more parameters we 'predict' the masses also of all the known fermions.

Only in fairy tales things are simple…

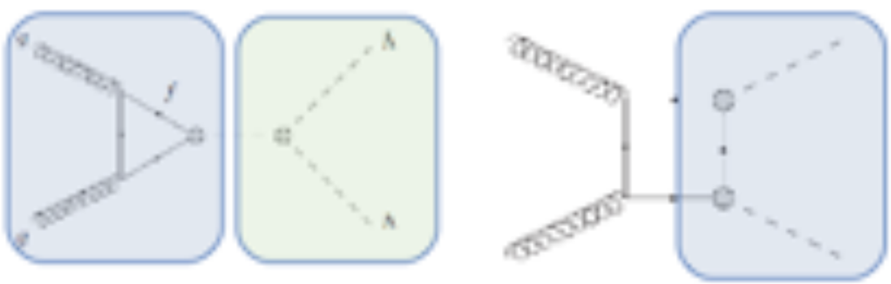
1) Why is \lambda = 0.012 ? Let measure it! (boh.. at some point we could arrive there)

2) **This talk:** What if if the EWSB is not as we expect? if there are other particles? other forces?
   2.1) What/how to expect to see it in LHC in this case?

Experimental challenge (in a fairy tale world):

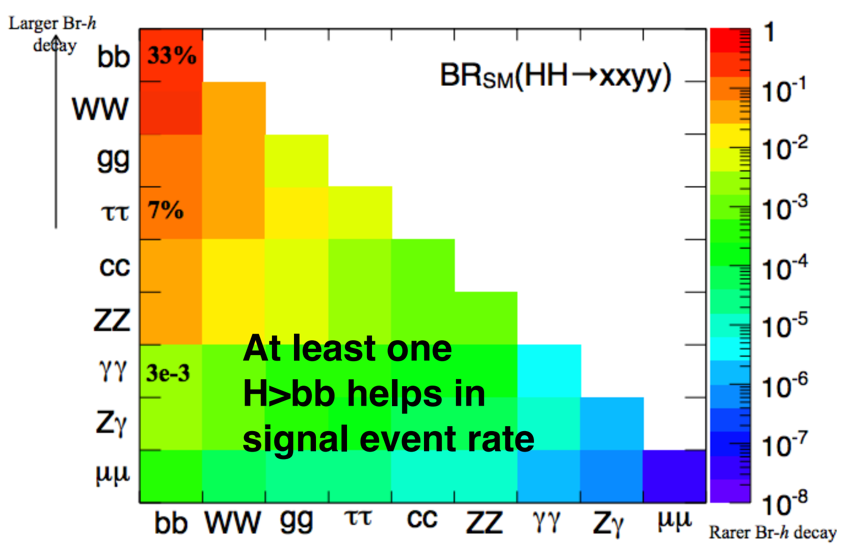
The first channel that allows one to probe the Higgs boson trilinear coupling is HH production

The most promising production mode at LHC is throughout gluon fusion,
this is a loop induced process at leading order (the loops contain mainly top quarks)



if \lambda = 0.012 and the SM particles and forces are all that exists:
the XS of this process is ~ 37 fb @LHC13 (pretty low)

The experimental search channels are classified buy the HH decay



The BKGs for the channels with larger BR tend to be huge and not easily separable (QCD).

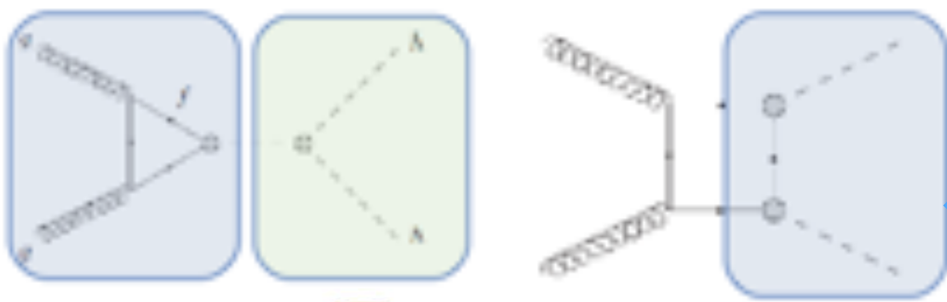
limit/SM	CMS	ATLAS
4b	342	29
ggbb	91	115
WW(ll)bb	410	-
tatabb	200	-
ggWW	-	700

See more details in this presentation

Theory challenge:

The first channel that allows one to probe the Higgs boson trilinear coupling is HH production

The most promising production mode at LHC is throughout gluon fusion,
this is a loop induced process at leading order (the loops contain mainly top quarks)



The **htt** coupling is also a SM parameter, that it better constrained by single H measurements, but in any case hides a bit our prize as price of signal kinematics

The presence of new particles and/or new forces complicates the picture

For all that we don't know (New Physics) we can take the **Effective Field Theory** approach as a good start

BSM diHiggs production: the humble approach

We parametrize our NP ignorance by modifications in the Higgs couplings (in terms of the physical Higgs boson). Truncated to dimension 6 operators

$$\mathcal{L} \supset \boxed{\frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{2} m_h^2 h^2 - \kappa_\lambda \lambda_{SM} v h^3}$$

$$\boxed{-\frac{m_t}{v}\left(v + \kappa_t h + \frac{c_2}{v} h h\right)\left(\bar{t}_L t_R + h.c.\right)}$$

$$\boxed{+\frac{1}{4}\frac{\alpha_s}{3\pi v}\left(c_g h - \frac{c_{2g}}{2v} h h\right) G^{\mu\nu} G_{\mu\nu}}$$

pure Higgs

Yukawa type

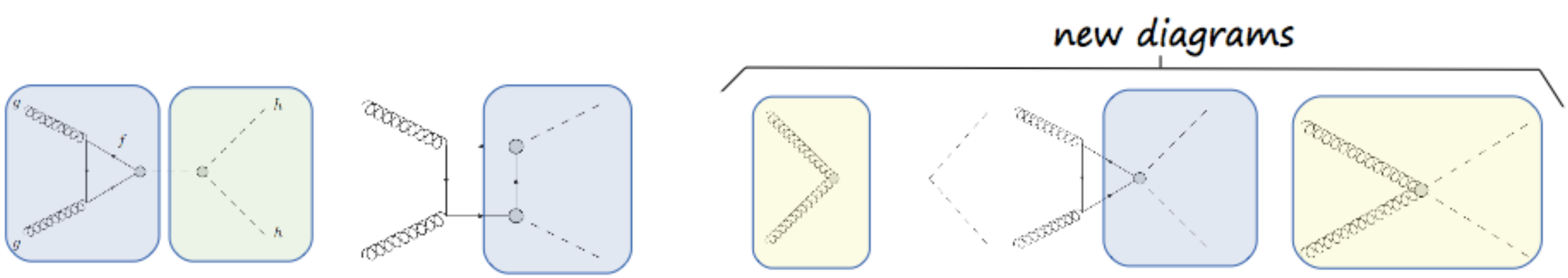
Gluon contact interactions

5 parameters:

$\boxed{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}}$

Defined as ratios with respect to the SM quantities

We use the most general parametrization (non linear lagrangian)
=> it includes the case the Higgs boson is not completely part of a doublet*



*The linear case is recovered if cg = - c2g

The different diagrams and their interferences produces a non-linear pattern of interference

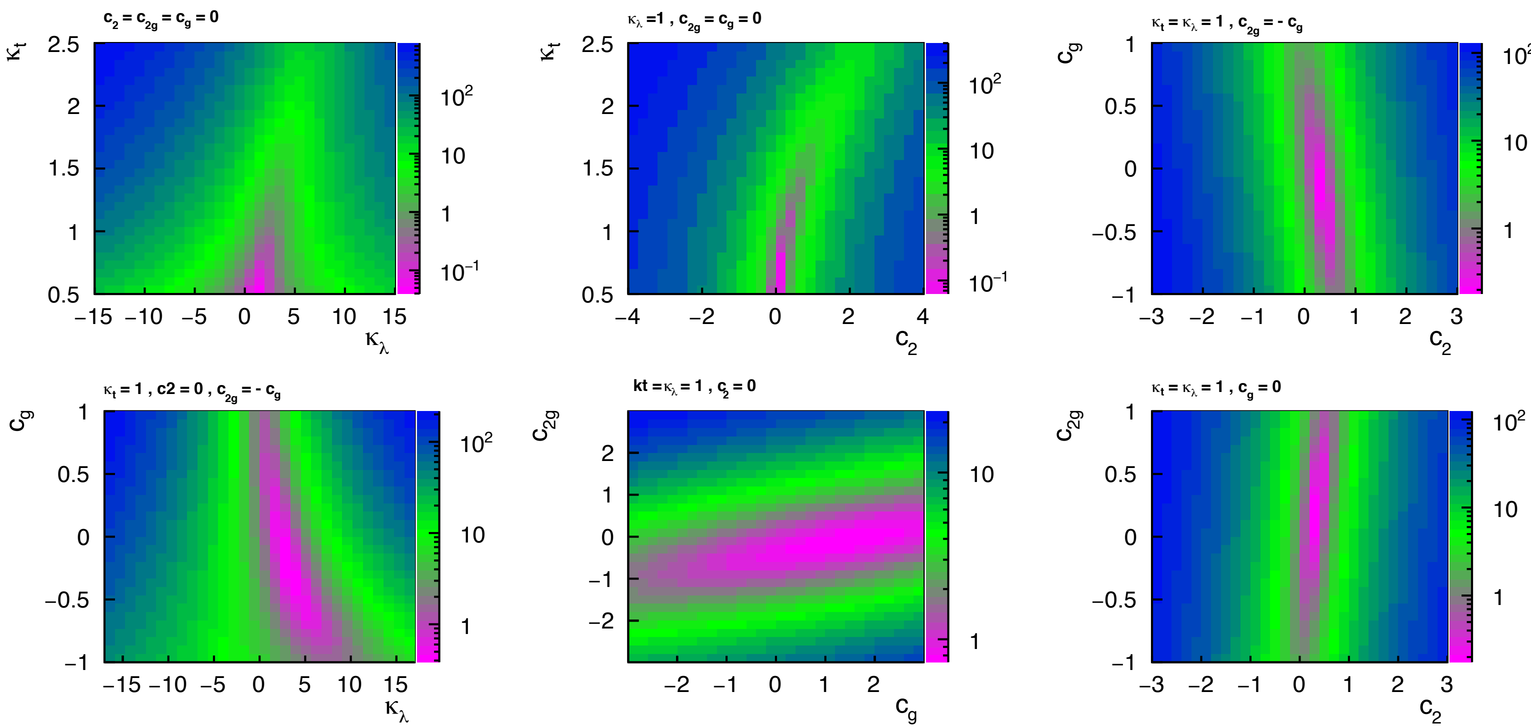
Cross sections for anomalous Higgs couplings

We have a theory framework, how much is the cross section rate in BSM models?

The process is too complicated to be calculated explicitly, we had enrolled an analytical cross section parametrization, based in Monte Carlo

LHCHXSWG-INT-2016-001

We variate the EFT parameters inside the range of parameters variation allowed by single H production:



XS/XS$_{SM}$

The XS modifications can arrive to 100 enhancement in the HH rate with respect to the SM!

From the behavior of the cross sections with the parameters scan we understand there are strong correlations among the different parameters.

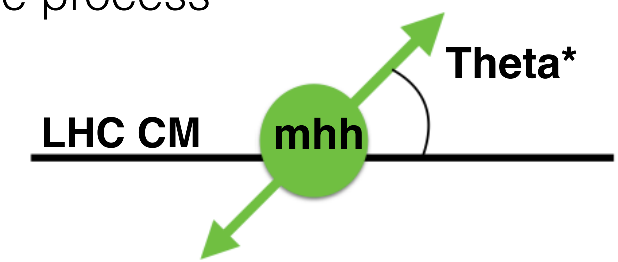
M. Dall'Osso, T. Dorigo, C. Gottardo, A. C., M. Tosi, F. Goertz

Cross sections are the tip of the iceberg as theory input differential information are crucial! (SM and BSM)

Modifications of the higgs couplings wrt its SM values can dramatically change signal topology
=> Interference patterns from theory parameters
=> Modifications from trigger acceptances to definition of the best analysis methods

While theory does not catch up, we want to rely in the simplicity of the process

At LO we have a simple 2 -> 2 process, where only two kinematic variables define the BSM piece of the process



longitudinal boost from proton PDF, it does not carry NP information

A purely statistical method to define benchmarks for non-resonant HH production at LHC, mapping the BSM H couplings to the signal kinematical properties in different points of parameter space.

The samples similarity is tested with a likelihood ratio based on Poisson counts

M. Dall'Osso, T. Dorigo, C. Gottardo, A. C., M. Tosi, F. Goertz (2015)

The Likelihood ratio Test Statistic

We test samples similarity with a likelihood ratio based on Poisson counts

If two samples under test share the same parent distribution the probability to observe n1,i and n2,i in the i-th bin is given by:

$$Pois(n_{i,1}) \times Pois(n_{i,2}) = \boxed{Pois(n_{i,1} + n_{i,2})} \times \boxed{Binomial(n_{i,1}/(n_{i,1} + n_{i,2}))}$$

All information about samples similarity is contained here

Ancillary information not relevant to sample comparison

It is possible to define the log-likelihood ratio (Test Statistic):

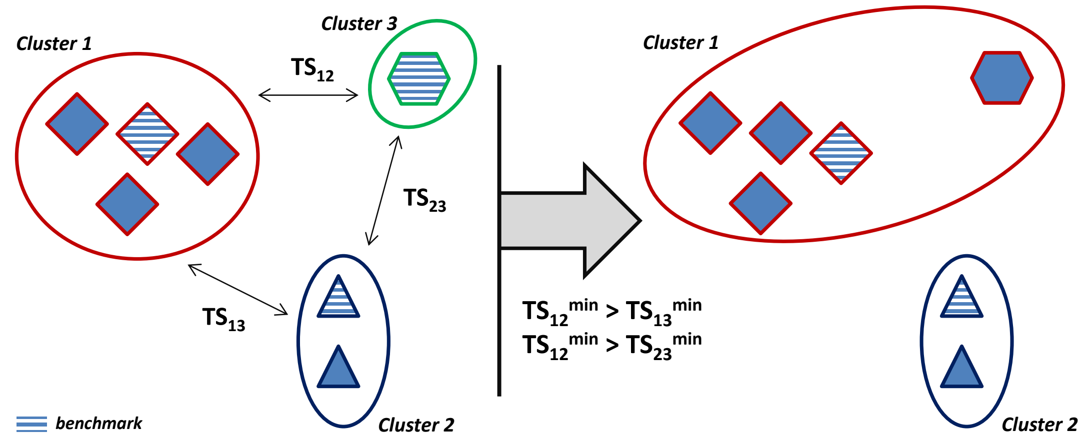
TS(1,2) = $= 2\ log\left(\frac{L}{L_S}\right) = 2\sum_{i=1}^{N_{bins}} log(n_{i,1}!) + log(n_{i,2}!) - 2log\left(\frac{n_{i,1}+n_{i,2}}{2}!\right)$

the denominator in the TS is the "saturated model" ($n_{i,1} = n_{i,2}$)

The TS is chi2 distributed, and can be used as an ordering parameter when comparing two samples (Wilks theorem !)

The cluster analysis:

Given 3 samples (i,j,k) if
TS(i,k) > TS (j,k) then the samples **(i,k) are more similar between then than (j,k)**



Steps for clustering samples:

1 - each sample is identified as one-element cluster

2 - the cluster to cluster similarity is defined as TSmin = min(TS(i,j))
where i runs on the first cluster elements and j in the second one

3 - The pair of clusters with the highest TSmin is clustered together

input parameter

==> repeat steps 2 and 3 up to a fixed number of clusters (**Nclus**) is reached

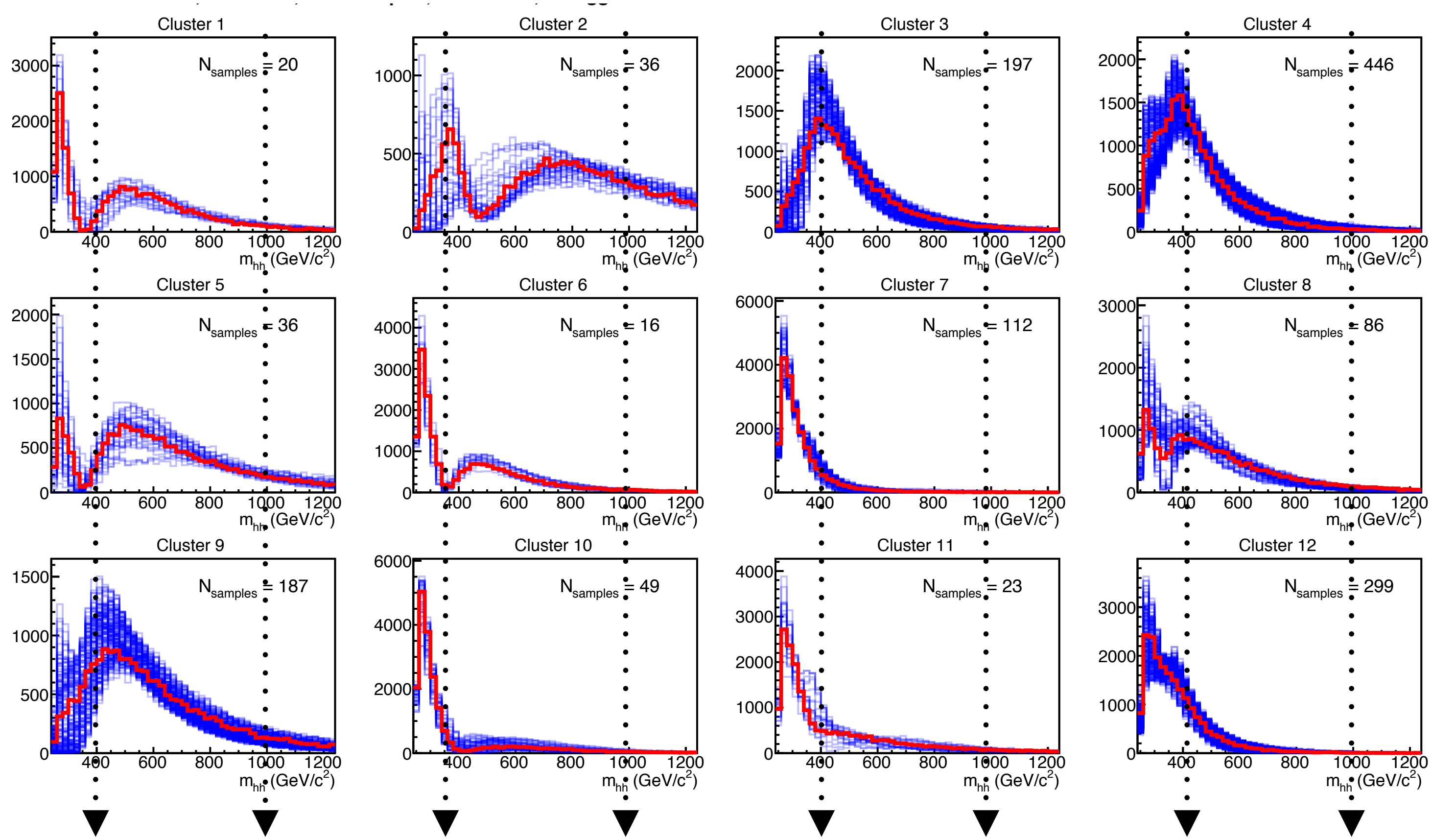
The **benchmark** is defined as the element with the highest TSmin(k) = min (TS(k,i))
where i runs over the elements of the cluster k
= the element most similar to all the other samples of the cluster

M. Dall'Osso, T. Dorigo, C. Gottardo, A. C., M. Tosi, F. Goertz (2015)

The cluster analysis solution

From a scan of 1507 parameter space points smartly chosen to spam the 5D parameter space, and using LO MC simulations we arrive in automated way to a kinematic classification to define benchmark points.

The Mhh distributions in the clusters:



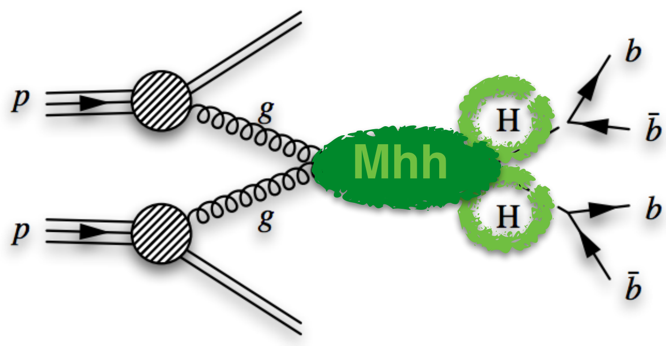
Red: the benchmark
Blue: the other samples in the cluster

The benchmarks are part of the LHCXSWG YR4

Mhh = 400, 1000, 400, 1000, 400, 1000, 400, 1000, GeV

M. Dall'Osso, T. Dorigo, C. Gottardo, A. C., M. Tosi, F. Goertz (2015)

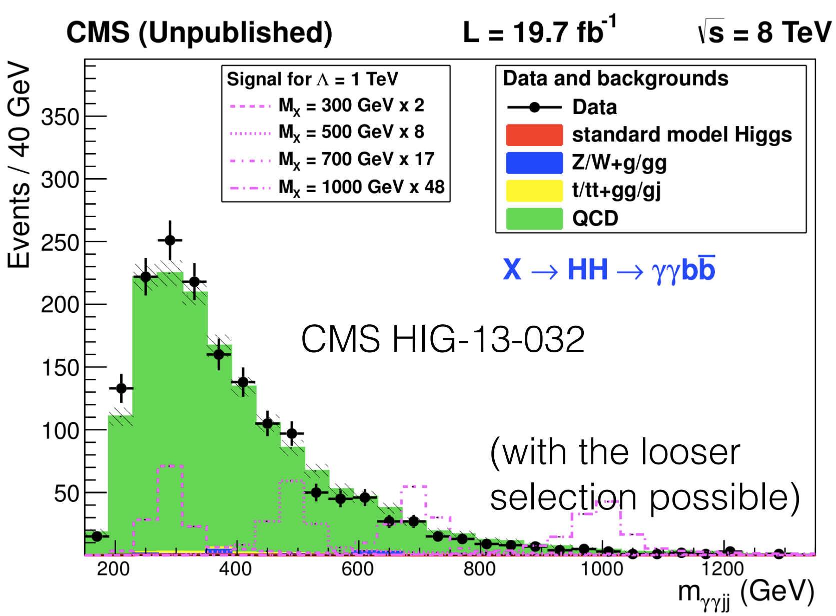
Experimental challenge (II):



250 GeV ← Mhh → few TeV

Depending on the mass of the X the best signal extraction method may change

Let's start with the hh > gamma gamma bb example



The BKG have kinematical peak on MX ~ 400 GeV.

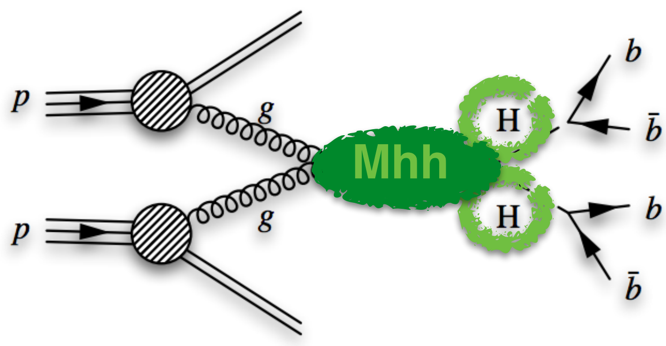
=> Near this region it is more powerful to extract signal exploring the line shape of the H bosons

To MX ~ 400 GeV. the BKG is a falling function

The picture for BKQ is ~ the same to all the HH full reconstructible channels

Experimental challenge (III):

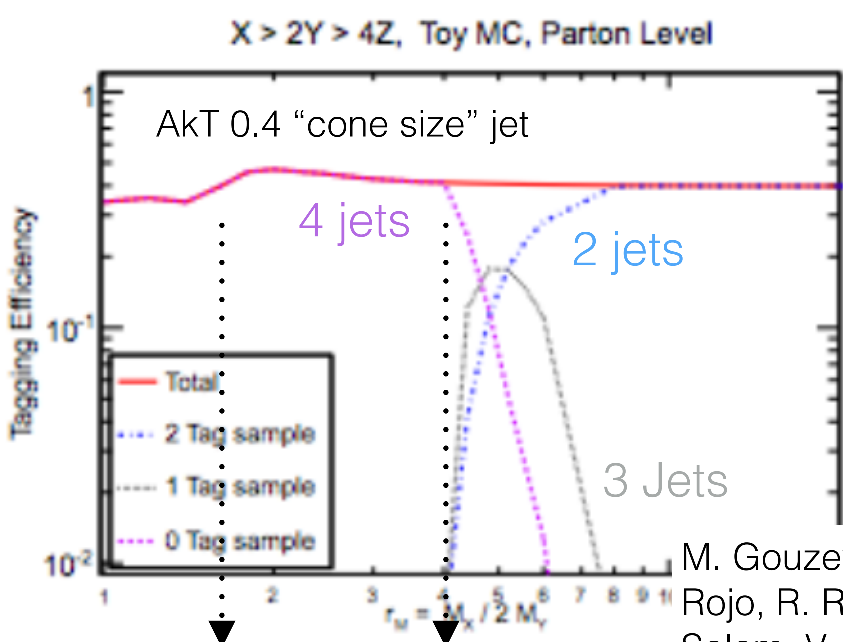
250 GeV ⟵ Mhh ⟶ few TeV



Depending on the Mhh the final objects reconstruction may change,
(in extreme cases as well the best trigger)

If the H is boosted its decays can merge in one unique object in the detector

Let's go to the 4b as a raw example:



Mhh = 400, 1000 GeV

M. Gouzevitch, A. C., J. Rojo, R. Rosenfeld, G. Salam, V. Sanz (2013)

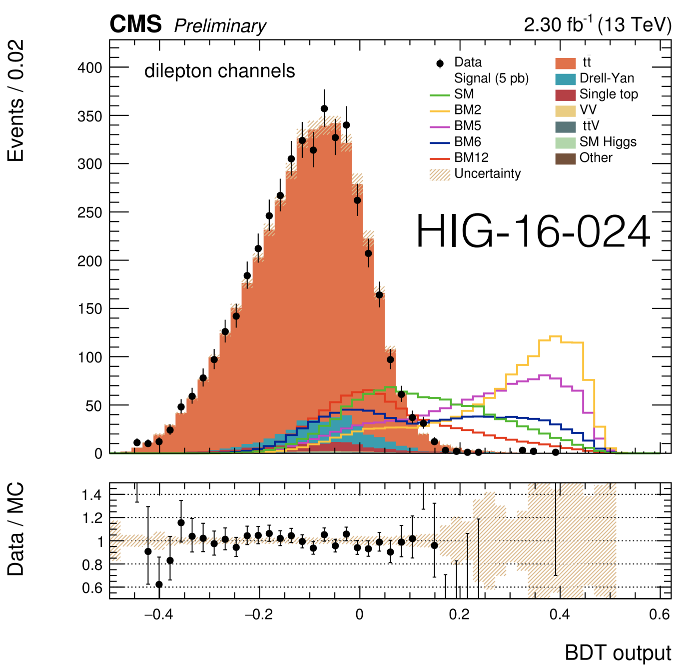
Similar picture happens in tau tau bb, WW bb and ZZ bb final states

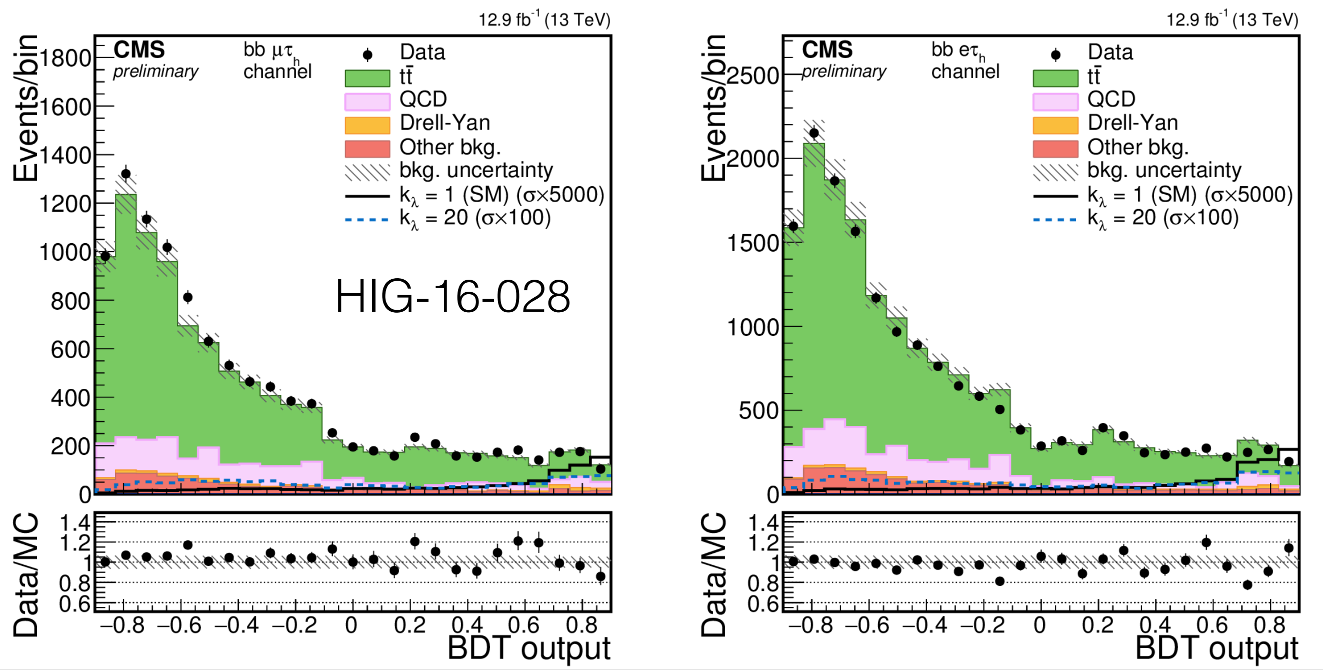
Experimental results:

We already have two results designed using the benchmarks

- WW (ll) bb
- tau tau bb (lsemi-leptonic + hadronic).

not fully reconstructable final states





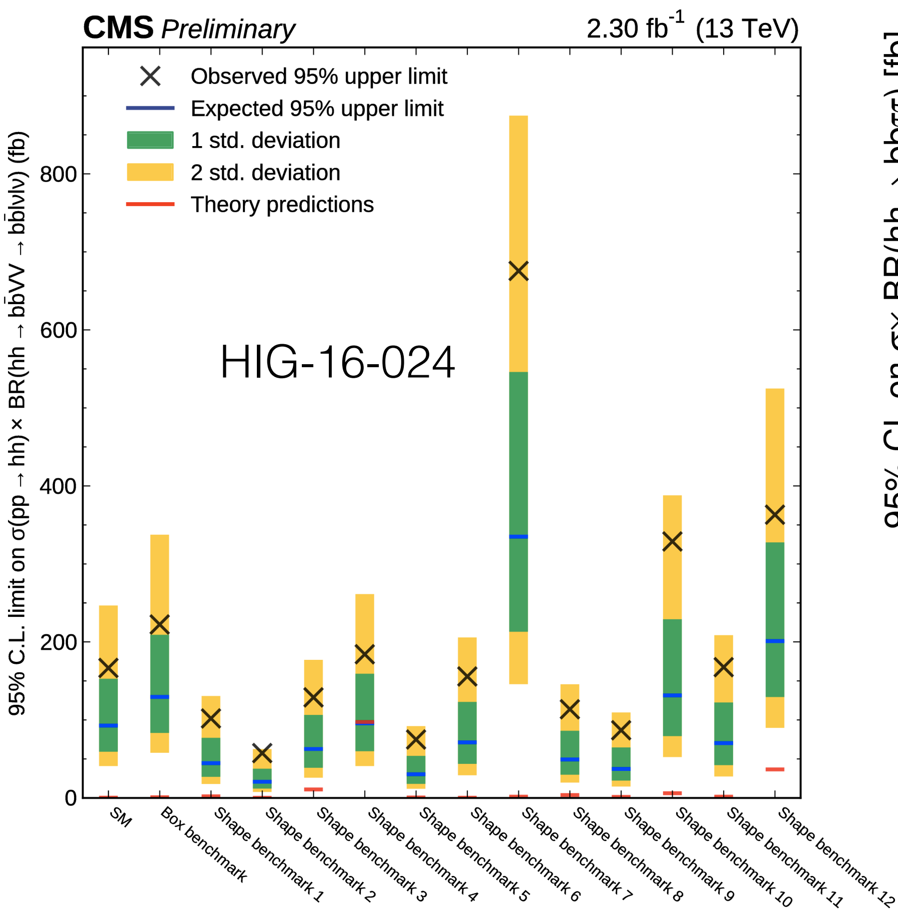
(a) bb$\mu\tau_h$ (b) bb$e\tau_h$

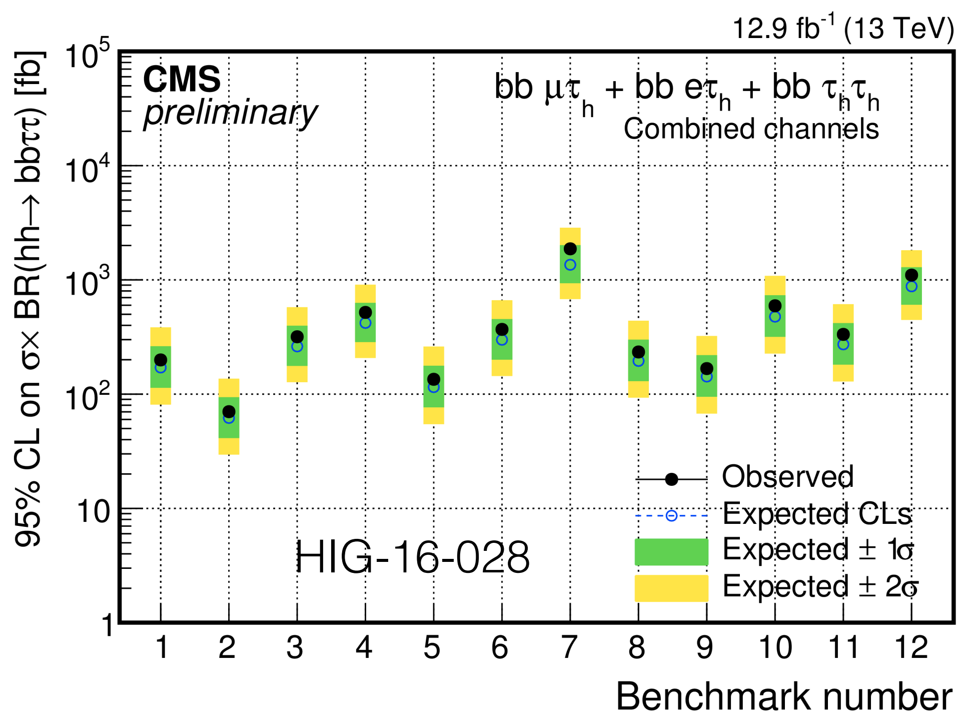
Preliminary results are made based in categorizations and cut-and-count.
Both searches need to rely in a BDT to separate signal, those were made based in a single benchmark and applied to .
In future this BDT can be done based in benchmarks, using more kinematical variables

Experimental results in benchmarks:

All the shape hypotheses are compatible with data.

The variation in final limits are striking. As expected, threshold-like samples are the worst constrained by those searches (not good triggers at low pt's)



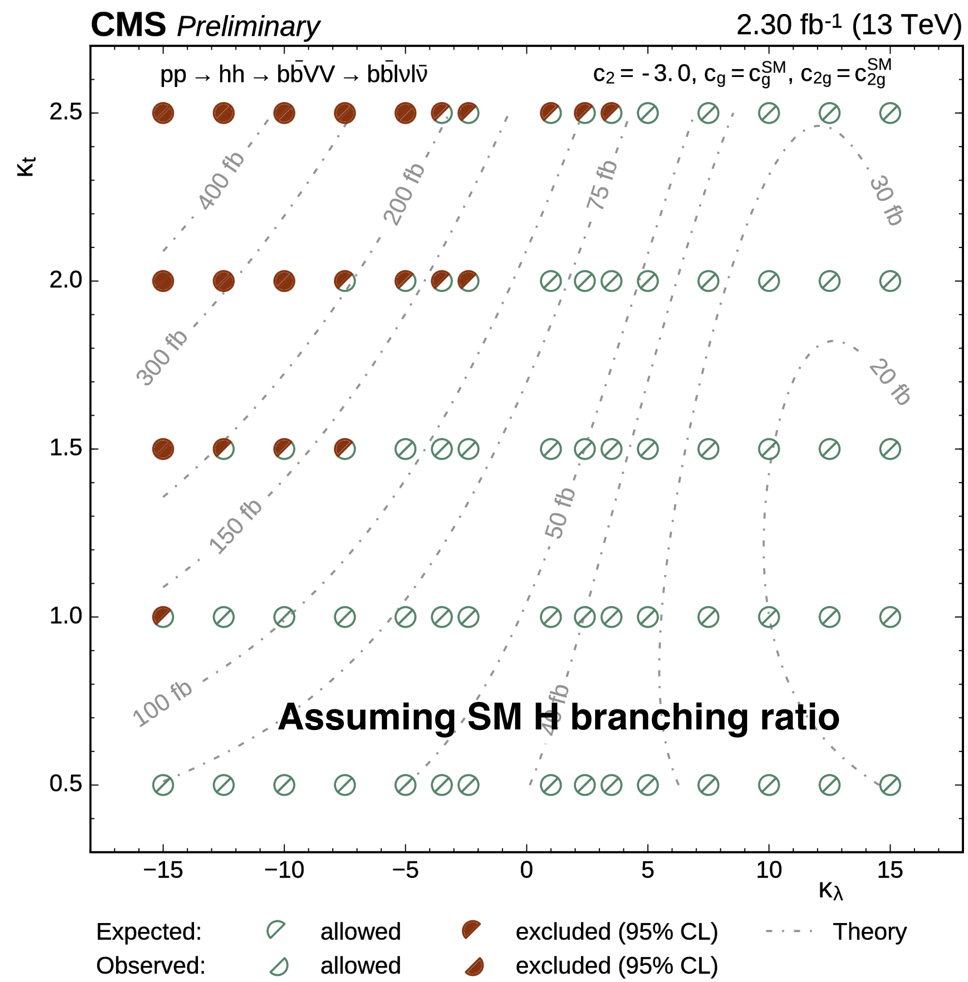


Preliminary EFT interpretation

The parton-level simulations made to define the benchmarks can be mapped back to the parameter space of H anomalous couplings.

If there are enough full-simulated events in the sensible we can always recover one parton level simulation with a simple reweighing.

The 2D histograms necessary to make samples reweighing are public, this made possible to the CMS experiment to provide asymptotic limits in a large quantity of hypotheses as a preliminary result in BSM HH.



The strategy for the final result for interpretation in anomalous couplings should look more continuous. This is work in progress.

If you also happen to have a striking nice how to fold limits by points there are ~1500 limits to play with in a preliminary parameter scan in the HIG-16-024 twiki

EFT interpretations to the glory

We had made the choice of benchmarks to optimize experimental coverage of New Physics effects (targeting discovery).

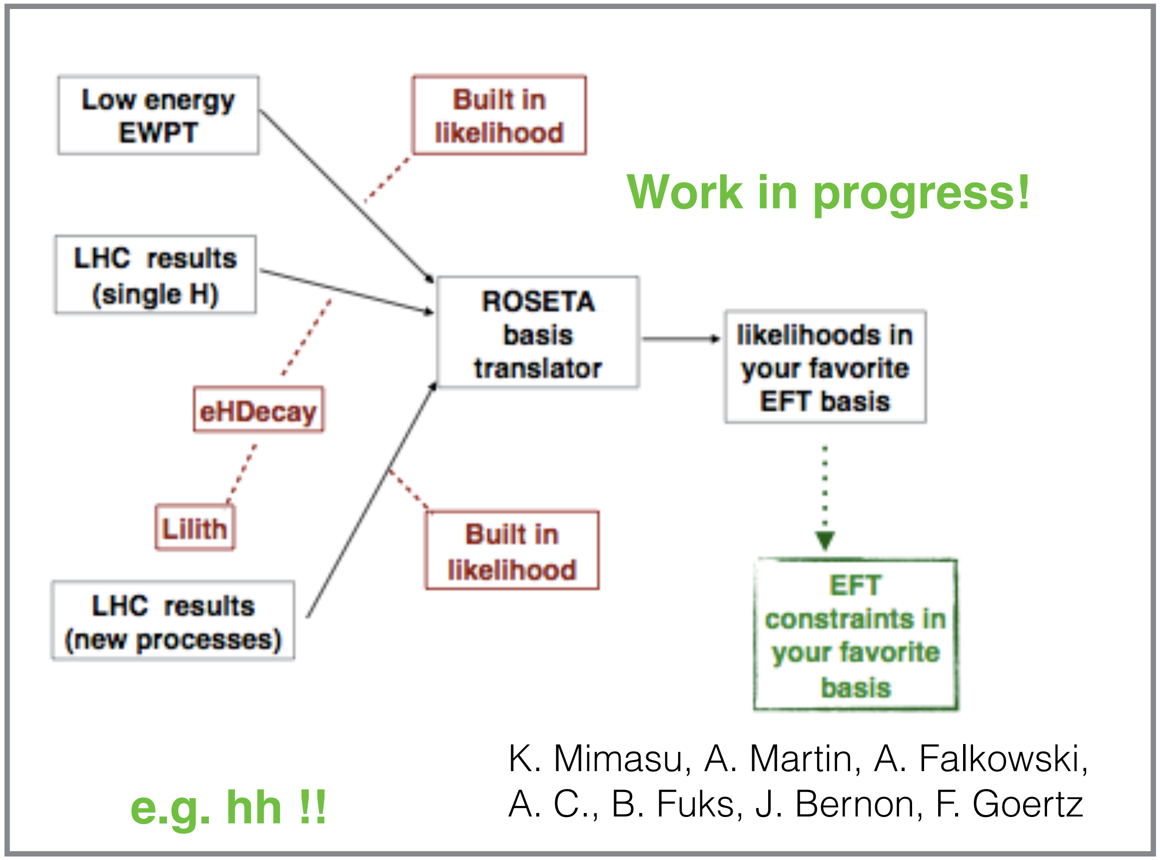
While we find nothing, will we be able to interpret the results in a fully flagged theory framework? YES!

A complete dim6 EFT treatment of the SM contain at least 69 operators.

ROSETTA is an automated routine that translate between the different EFT languages.

A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva, V. Sanz (2015)

The EFT bases translation is the first step towards to unify the different SM precision tests and LHC measurements and searches in robust theory constraints.

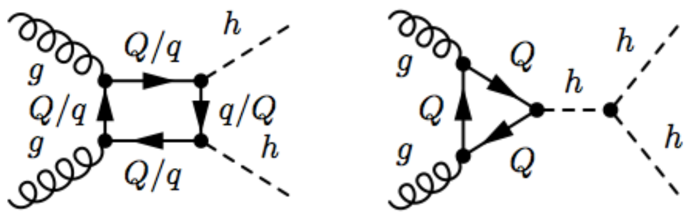


Preliminary automats in ROSETA: pick a point in the parameter space and find the cluster it belongs,

https://github.com/kenmimasu/Rosetta/tree/master/dihiggs

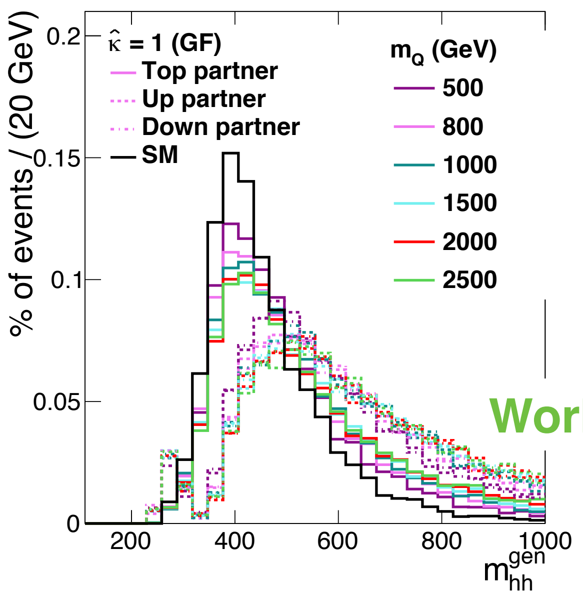
And if NP is not decoupled? Proof of principle gg > HH (loop induced with VLQs)

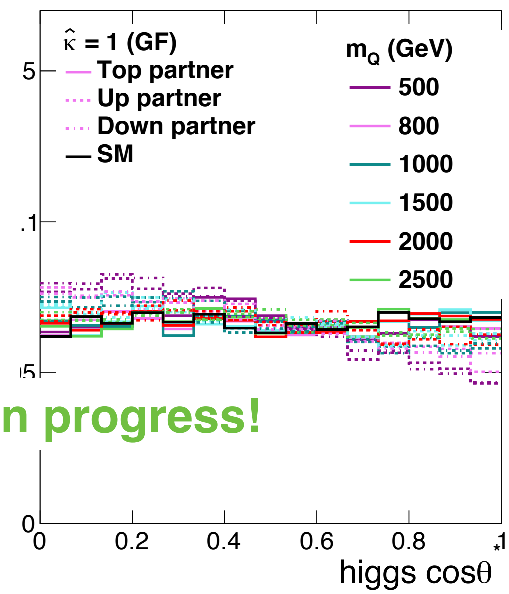
In the Mhh < 1 TeV range the process is dominated by the loop induced processes



$$\hat{\kappa}_L^{T/B} = \kappa \, \frac{M_Q}{v_{\rm SM}} \, ,$$

While k is the coupling between Qqh in the unitary gauge





Work in progress!

Light quarks partners in the loop produce interesting features:

- The interference with the SM is clear form the Mhh distribution

- More angular asymmetry (to be studied analytically as well)

Direct correspondence with the previously defined benchmarks
=> In principle no recast is necessary to use the experimental result!

G.Cacciapaglia, H.Cai, A.C, A.Deandrea, T.Flacke, B.Fuks, D.Majumder, H.-S. Shao

There are many others explicit model examples

Heavy particles can add production processes - loop and cascades
(colored scalars as well, for example SUSY like Stops)

Extended Higgs sectors modify the tree level relations among the couplings in the Higgs sector (again the SUSY-like example)

Extended symmetry or extra dimensions induce high order operators in the Lagrangian, other coupling structures

Non imaginable surprises?!

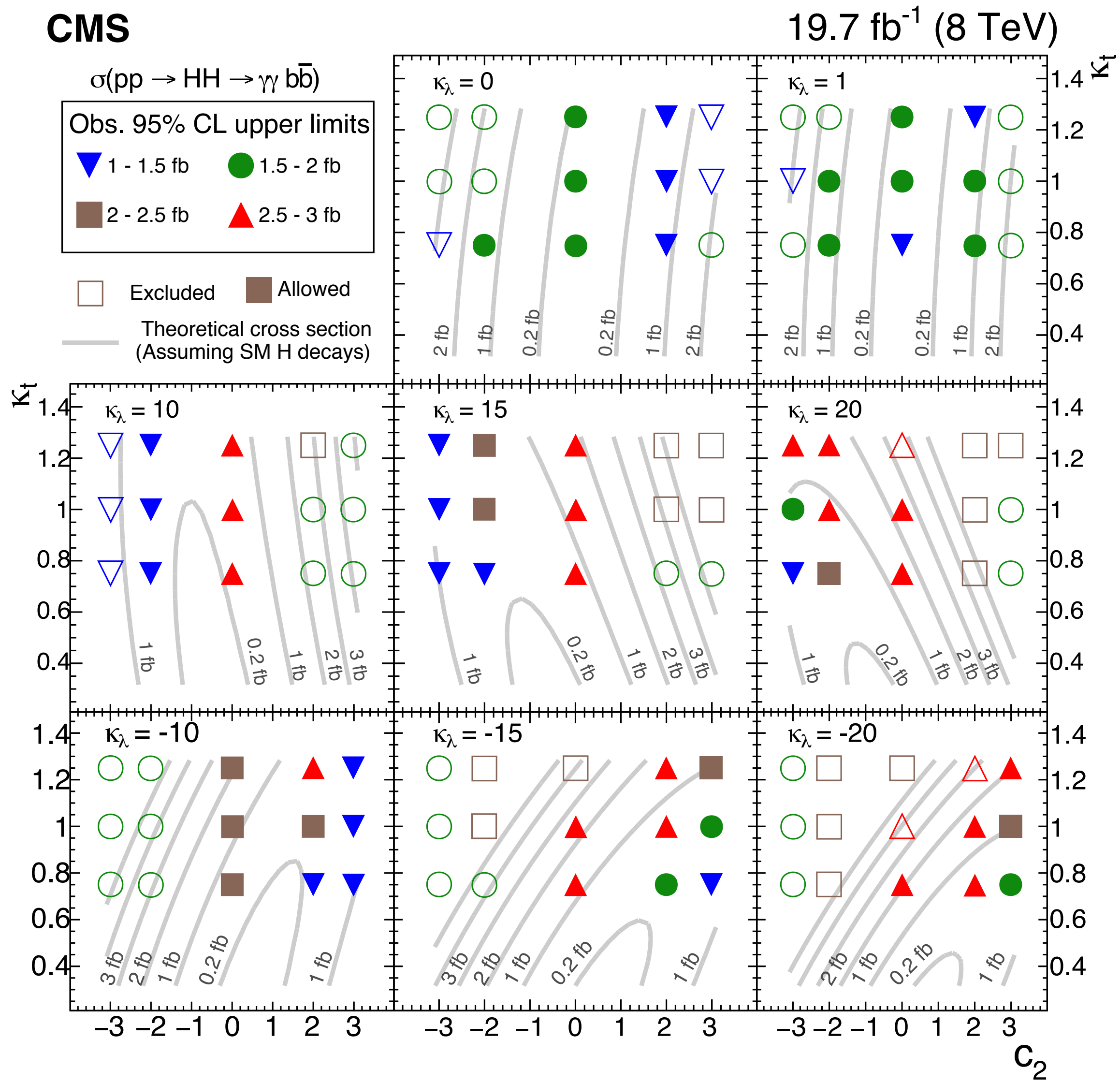
Conclusion

We are preparing the nest to find and interpret the NP signals that will (or not) come!

On experimental side too! See next two talks

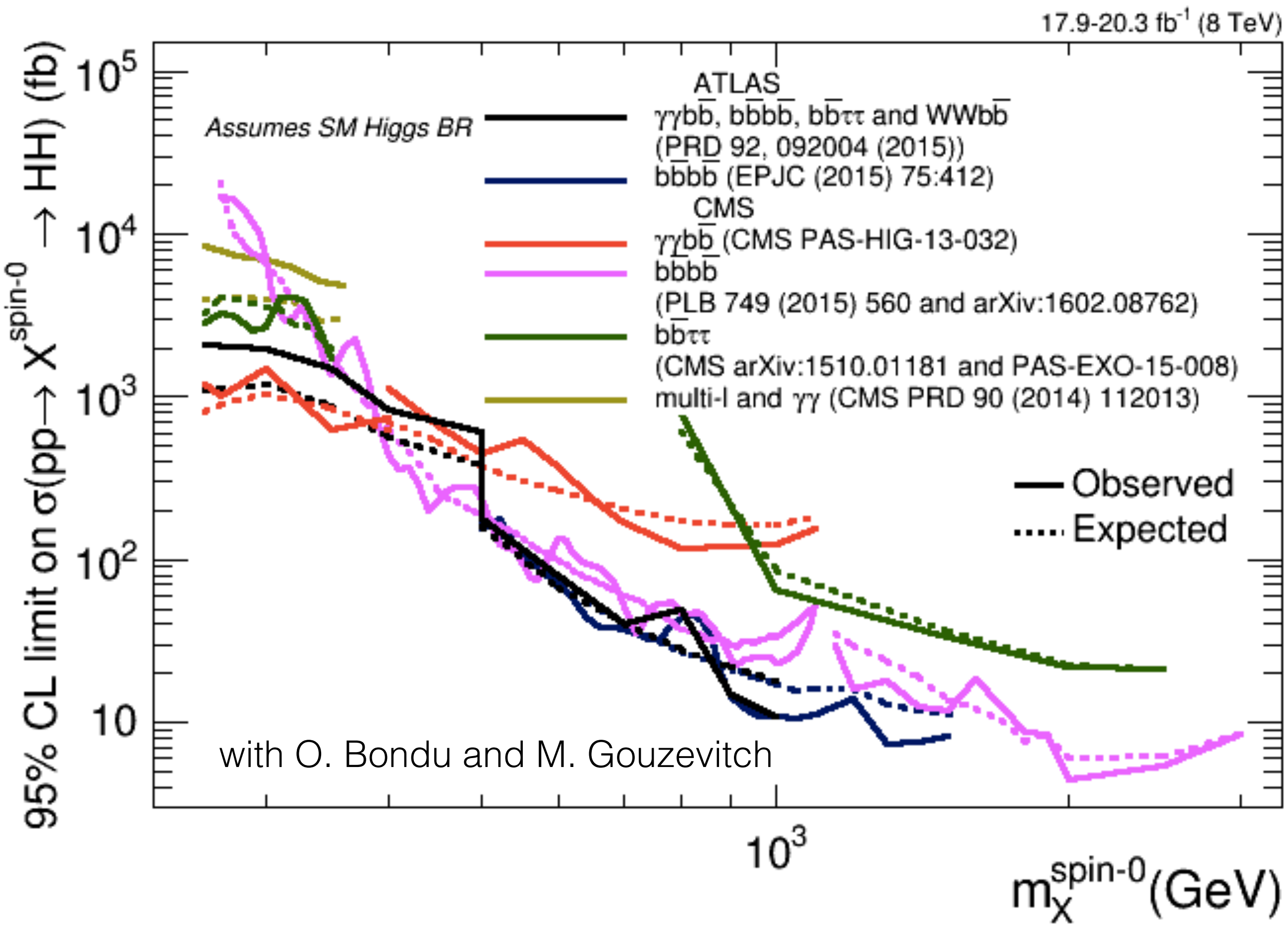
A ggbb preliminary result in BSM from CMS

Categories in Mhh



Run 1 legacy:

The HH enterprise started from experimental searches for resonances in the hh final state:



=> In different energy ranges
different final states dominate

**The H branching ratios can be modified in specific models, this is out of the scope of this talk

Cross sections for anomalous Higgs couplings

We would compare the results with the total cross section, rather than the enhancement

$$\overline{\sigma_{hh}^{SM}} \times R_{hh}$$

The state of the art calculations (NNLO) + uncertainties.	Result of our fit (LO) + uncertainties

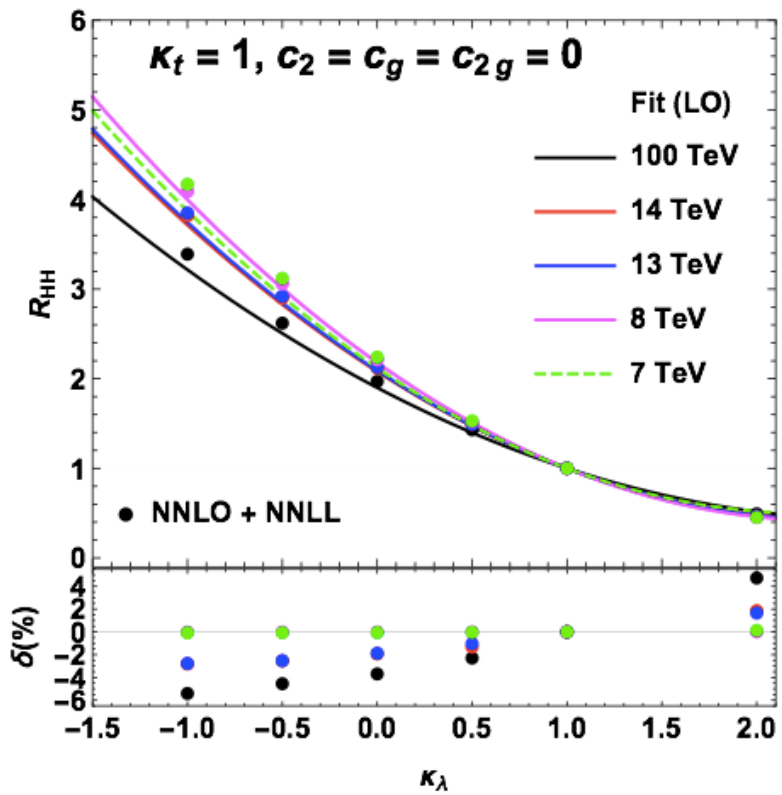
Both parts have theory errors from the same source, with a degree of correlation (PDF and scale)

We would like to estimate the part of the theory errors that act ONLY on R_{hh}, that we will call **BSM-like systematics**

factorizing the bulk of the theory uncertainty to be taken into account by the normalization $\overline{\sigma_{hh}^{SM}}$

**The fit errors in the coefficients of R_{hh} are estimated by the residuals wrt the central value, and are found to be negligible (~ 0.1% / point).

Cross sections for anomalous Higgs couplings



D. de Florian and J. Mazzitelli

LHCHXSWG-INT-2015-003

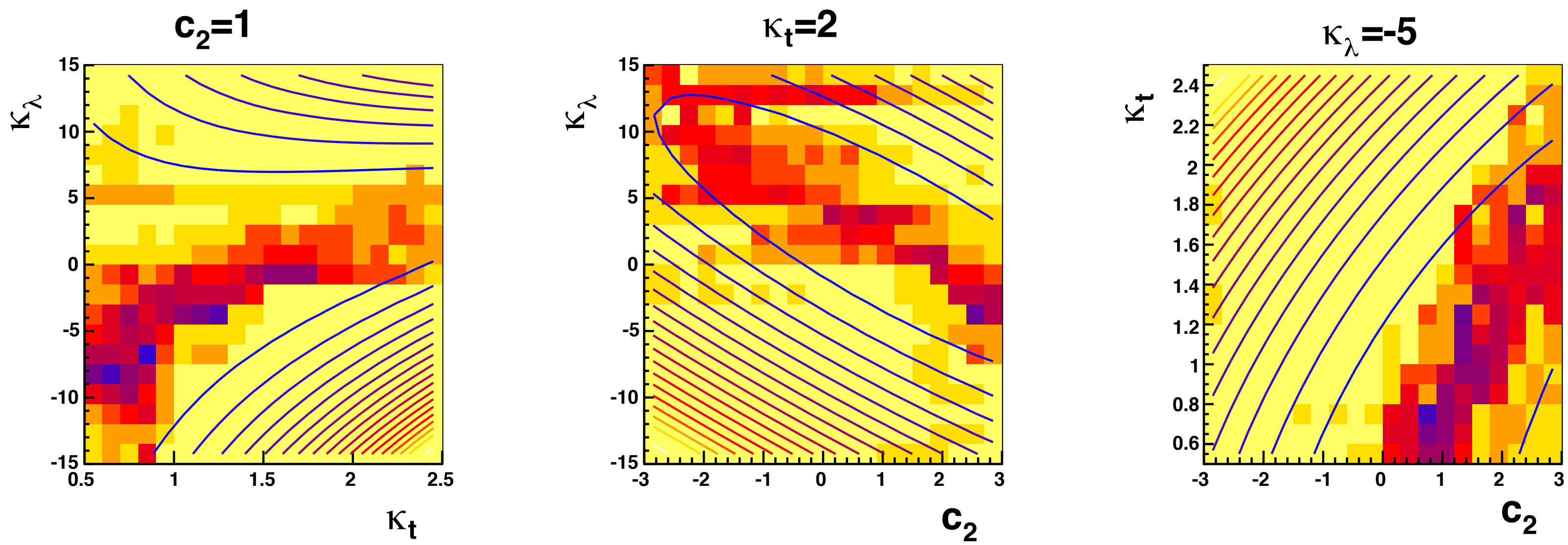
Benchmark	1	2	3	4	5	6	7	8	9	10	11	12
8 TeV δ_p	0.4	2.4	0.1	0.2	0.2	1.2	0.4	1.1	0.4	0.9	1.0	0.2
13 TeV δ_p	0.5	2.2	0.1	0.3	0.2	1.3	0.2	0.9	0.5	0.9	0.6	0.2
14 TeV δ_p	0.5	2.2	0.1	0.3	0.1	1.2	0.2	0.8	0.5	0.8	0.6	0.1
100 TeV δ_p	0.6	2.2	0.1	0.4	0.1	1.7	0.4	0.6	1.0	1.0	0.2	0.2

Table 3: Theory uncertainty (%) estimated due PDF and α_S variations to each one of the benchmark points (j) of table 2 to the four center of mass energies we consider.

Sampling in the anomalous couplings parameters space

A larger variability in kinematic topologies is correlated with local minima of cross sections (where apparent cancelations among different processes holds)

When we overlap the values of the TS between two nearest neighbors samples with the isolines of cross section we directly see the correlation

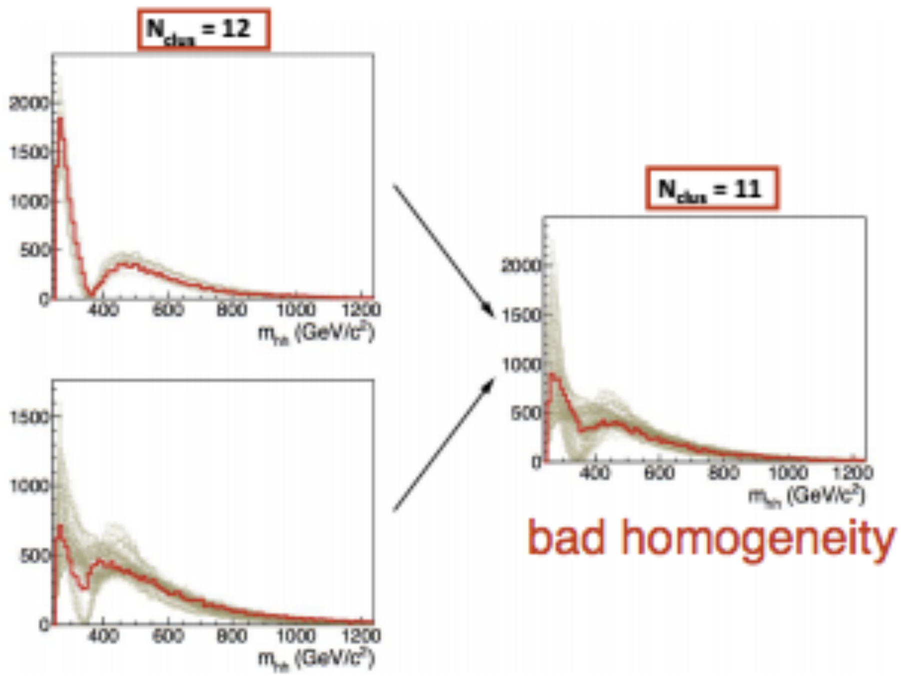


Blue/Red stand for higher values of the TS between nearest neighbors

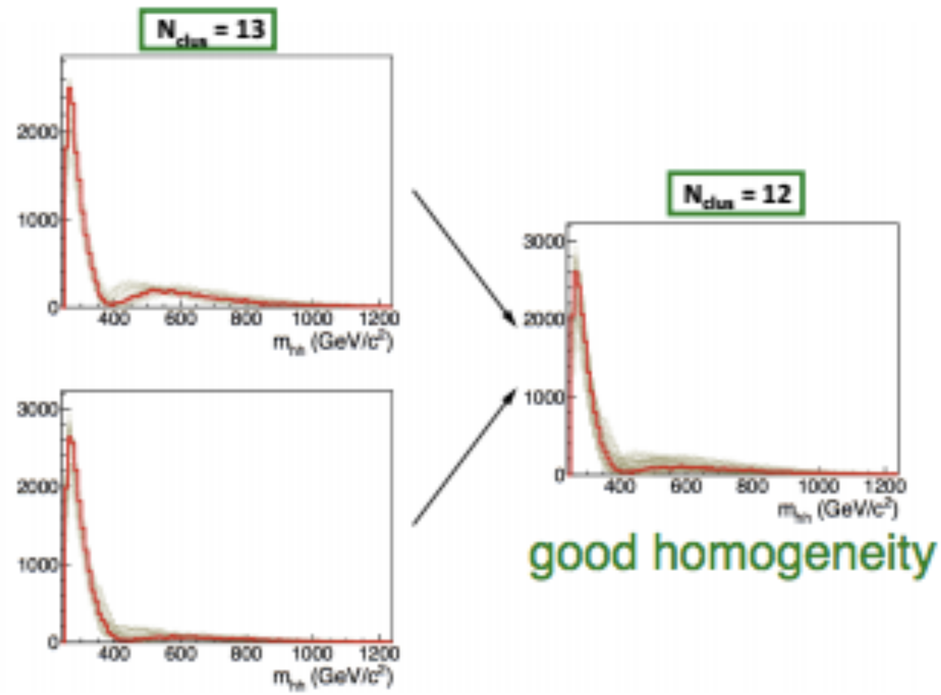
Prioritizing variety of kinematic configurations to our dataset we sample a larger density of points close to the cross section local minima

Why to stop with 12 clusters?

11 clusters are visually too few



13 clusters are visually too much



Trade off between number of benchmarks and intra-cluster homogeneity