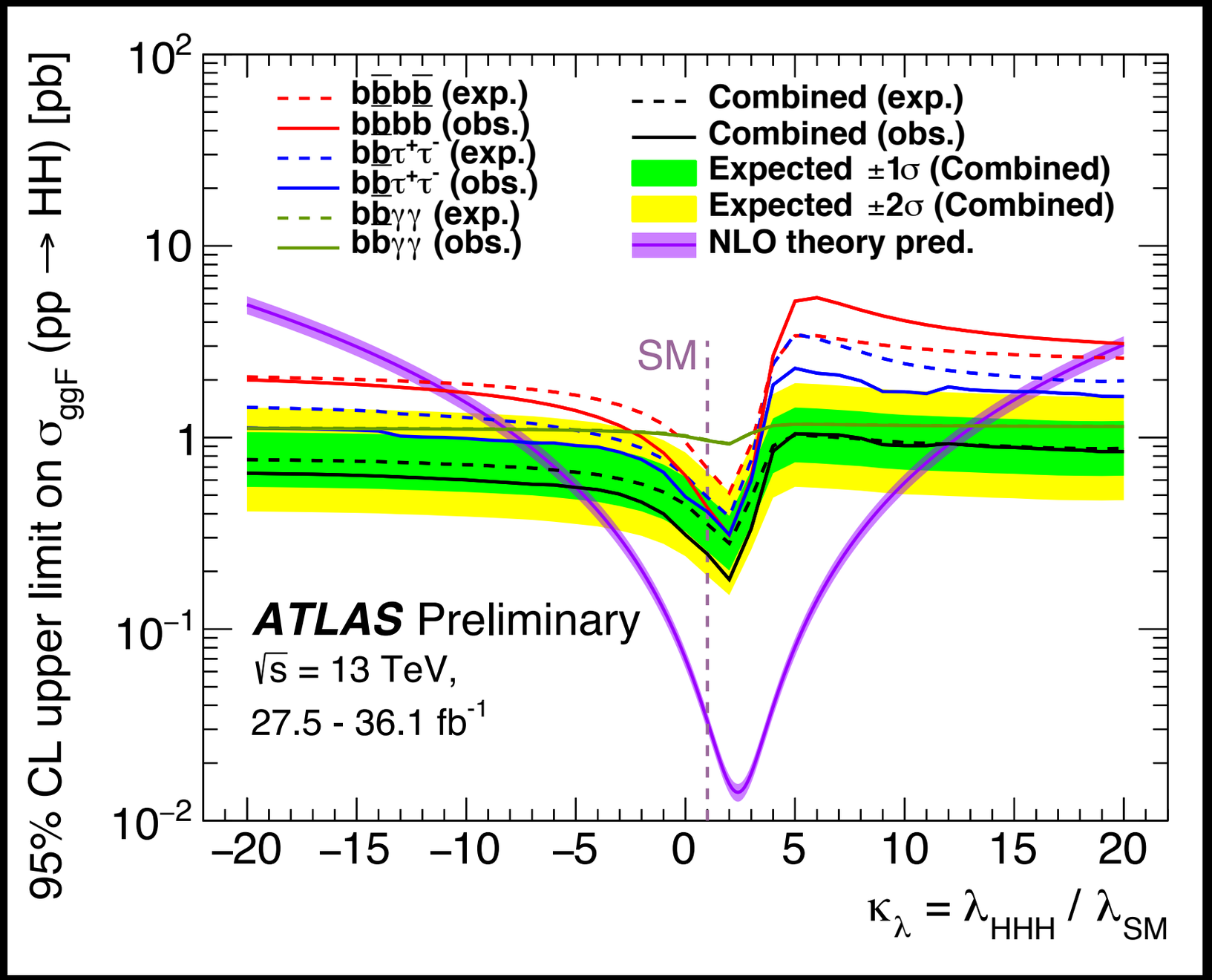
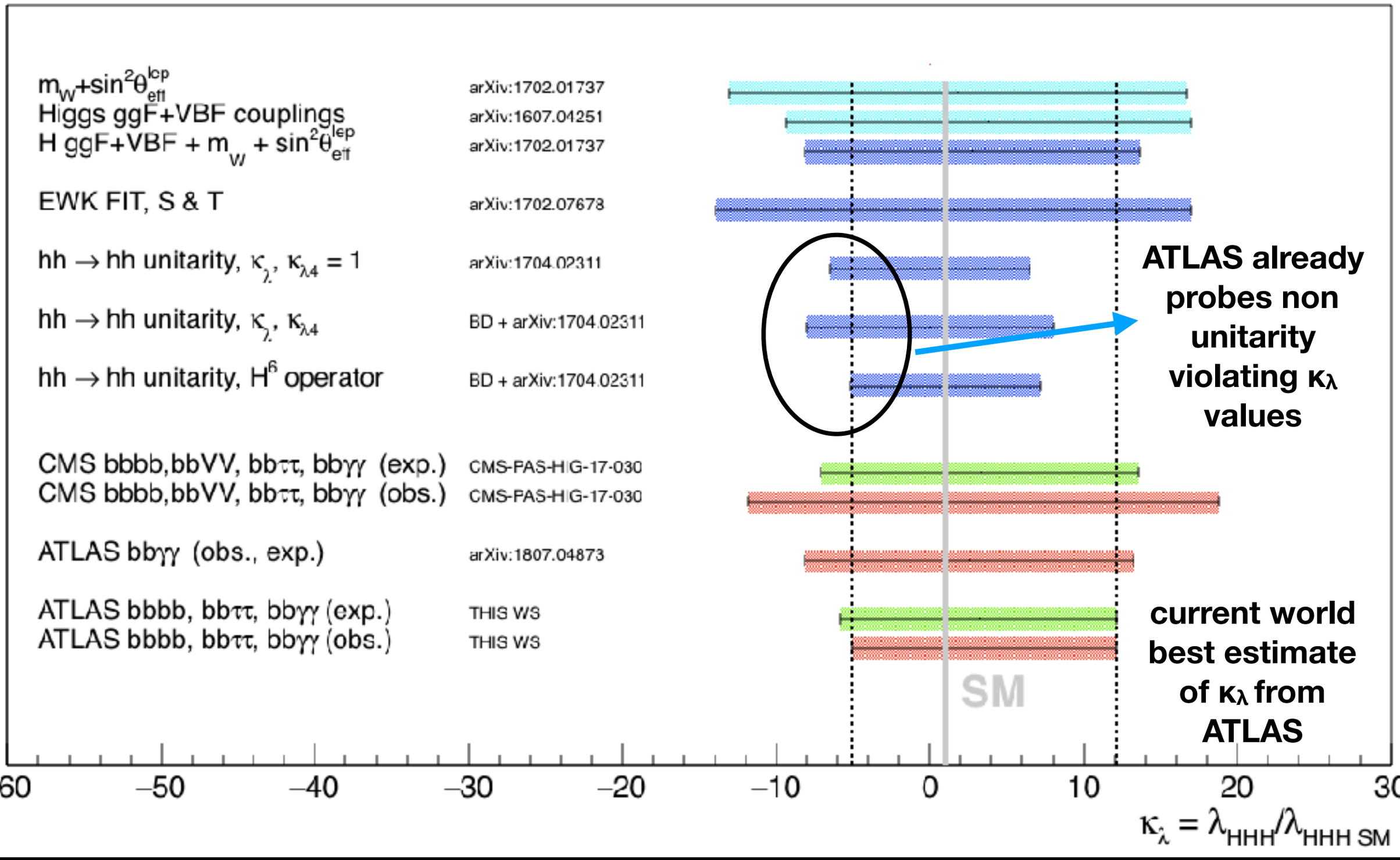
Discussion summary and followups

k_λ status after this workshop

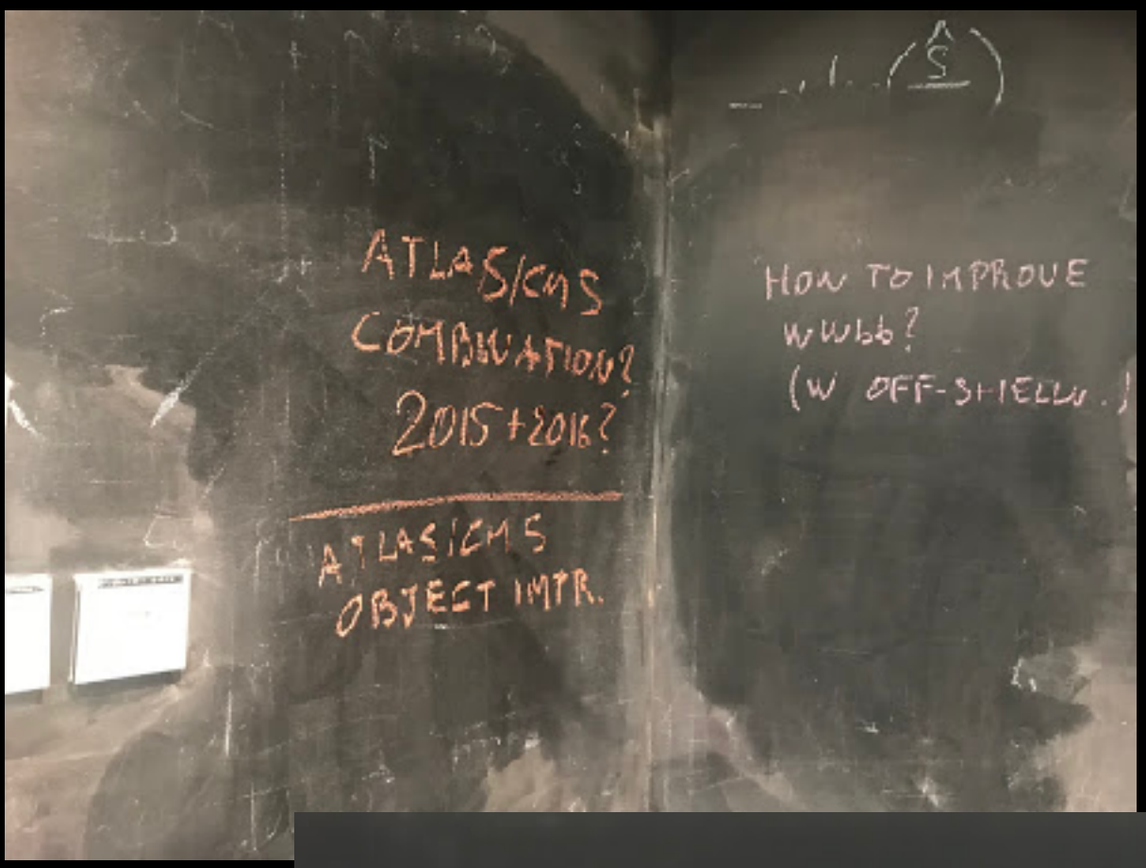


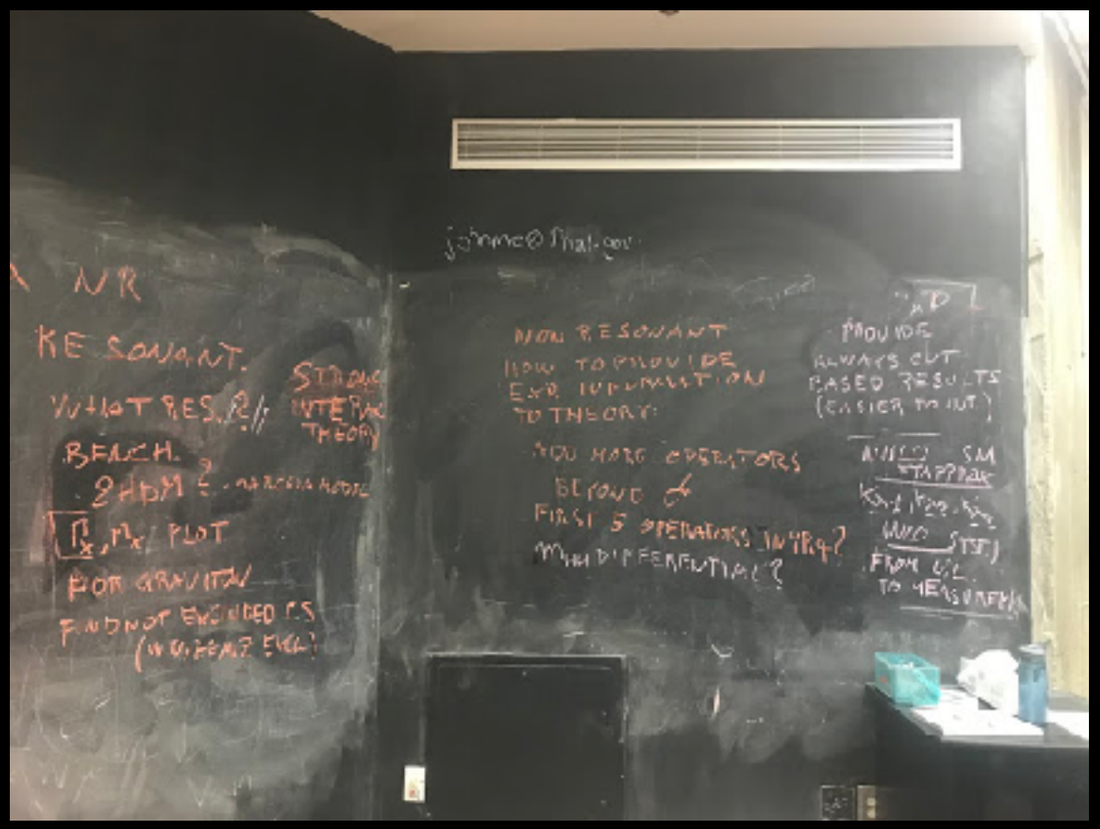
$-5 < \kappa_\lambda < 12.1$ **@95% CL (missing theory systematics)**

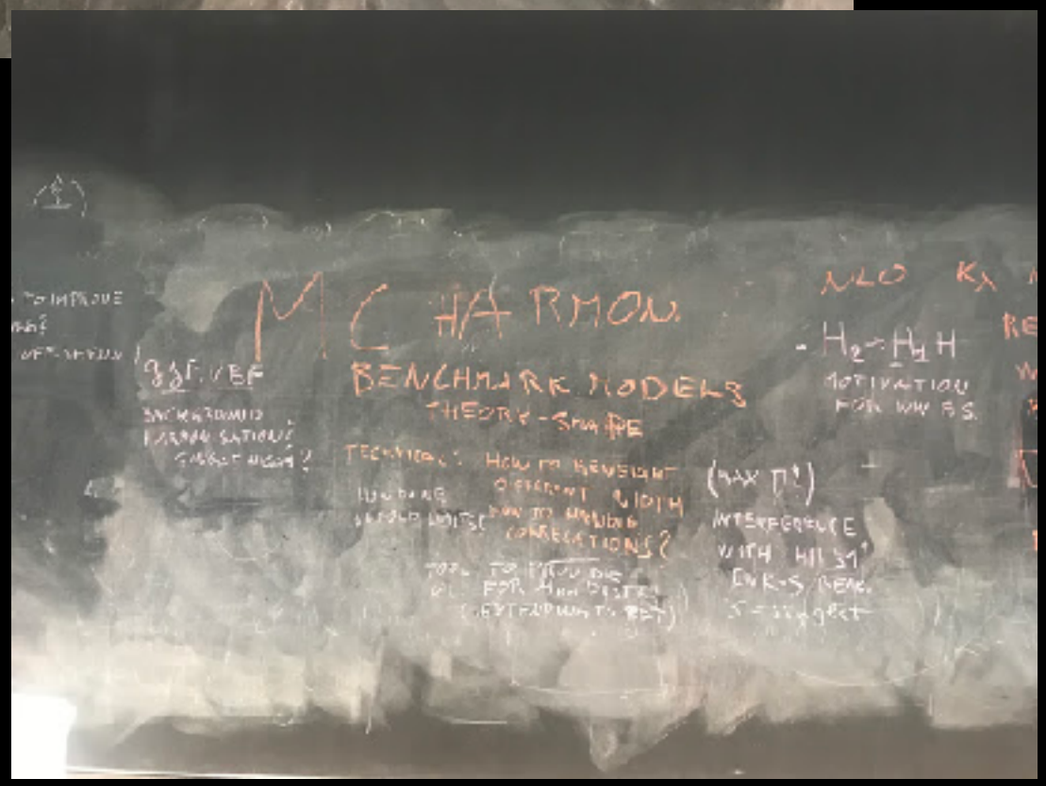
k_λ global status after this workshop



Discussion session







Starting from Tuesday blackboard many nice ideas that needed to be developed

High partecipation at the discussions (even on Friday night after the wine and cheese)

Discussion points grouped by the conveners into discussion sessions

1. ATLAS+CMS combination (Luca C., David W., Javier M., B. Di Micco) - Wednesday Room Ramsey

- based on 2015+2016 in preparation of the Run-2 legacy
- MC settings (NLO vs LO)
- Single H+HH combination
- Total cross section vs k_lambda + uncertainties

2) How to make results public (J. Allison, Max S., K. Leney) - Friday 1W (17:30 - 18:30)

- Tools to provide unfolded UL
  - How to handle bbb correlations?
- Resonant: Gamma vs Mx plot?
  - How to reweight different widths?
- cut based results
- differential results in m_HH (truth vs. reco)
- Special care for BDT-based results

3) EFT (S. Di Vita, M. Gouzevitch, J. Robinson) [17:30 - 18:30] 11th floor ROC

- Which framework? More operators beyond O6
- How to make EFT useful for model testing?
- Which inputs from H and HH?
- Which topology? ggF/VBF single H background?
- Usage of shape benchmarks

4) BSM (M. Carena, K. Tschann-Grimm, Ian Lewis, Lian-Tao Wang, X. Carvalho) Wednesday 11th floor ROC - Room One West

- Benchmark models : which one ?
  - Resonant: Is graviton still a good benchmark?
  - Interplay with VV
- Motivations for H1->H2 h
- Interference with SM HH (EWK-S, 2HDM) benchmark

5) ATLAS/CMS objects/analysis strategies (M.Kagan, F. Micheli, C.Vernieri) [Thursday 1W (5:00-5:45)]

- Trigger strategies
- B-tagging and b-jets (regression)
- MET

6) How to improve WWbb (W off-shell) [S. Shrestha, N. De Filippis - Thursday 11 ROC]

Discussion 1

ATLAS/CMS combination:

Combining 2015+2016 is not needed: each experiment will have 4 times more statistics at the end of Run2, the priority is to have full data results now.

The timeline for CMS could be relatively short, but CMS want (is obliged to) work on analysis optimisation
ATLAS thinks that the priority is to have full reoptimised result: realistic timeline is in 2 years from now, unless there is some pressing from PC.

Given this time line, it is not useful to start a concrete combination effort now: analyses could change, systematics too, stat framework is already quite similar. A realistic time is 2020, after all individual channels legacy papers are out.

ACTION ITEMS FOR COMBINATION
Talk to PC, ask to build up an ATLAS/CMS hh combination group with the mandate
- **define the software for the combination**
- **define MC for signal, systematics, and correlation**

Start to think to interpretations and benchmark models to insert in the paper
Address a possible H+HH combination

2) MC generation

- CMS using LO+Pythia8, it is known that emission from Pythia8 is hard, even harder than NLO, plan is to move to exact NLO finite mt correction
- ATLAS is using MG5_aMC@NLO + Herwig++, with FTapprox

Concerning k_Lambda variation, ATLAS re-weigths NLO FTApprox using LO weights, CMS uses plain LO

Theory: in order to have NLO with exact finite mt NLO as a function of k_Lambda, grids need to be provided by the theorists, this could become available by the legacy papers timeline

ACTION ITEMS FOR MC

ATLAS and CMS continue developing NLO with exact mt, using both Pythia8 and Herwig7 PS and looking at Powheg and aMC@NLO, decision to be made when validation distribution will be available

- set from now NLO with finite mt correction as baseline for combination
- concerning k_Lambda variations, ideally we would use NNLO+NNL, in case varied samples will not be available FTApprox could be used to reweigth the SM sample to different k_Lambdas.

3) H + HH combination

- Using only the HH channel it is not possible to decouple k_t from k_λ, needs to make an arbitrary choice of k_t for HH only results (it can be set at 1 given the present results on Higgs coupling)

- another option is to make a combined fit H+HH, but in HL-LHC projections k_λ precision is barely improved by the addition of the single H channels, nevertheless it allows to constraint the other Higgs couplings removing degeneracy on k_t

- one issue is the cross section variation of ggF/VBF Higgs production as a function of k_λ, it can be parameterised but a full NLO EWK analysis needs to be performed in order to get the k_Lambda dependence

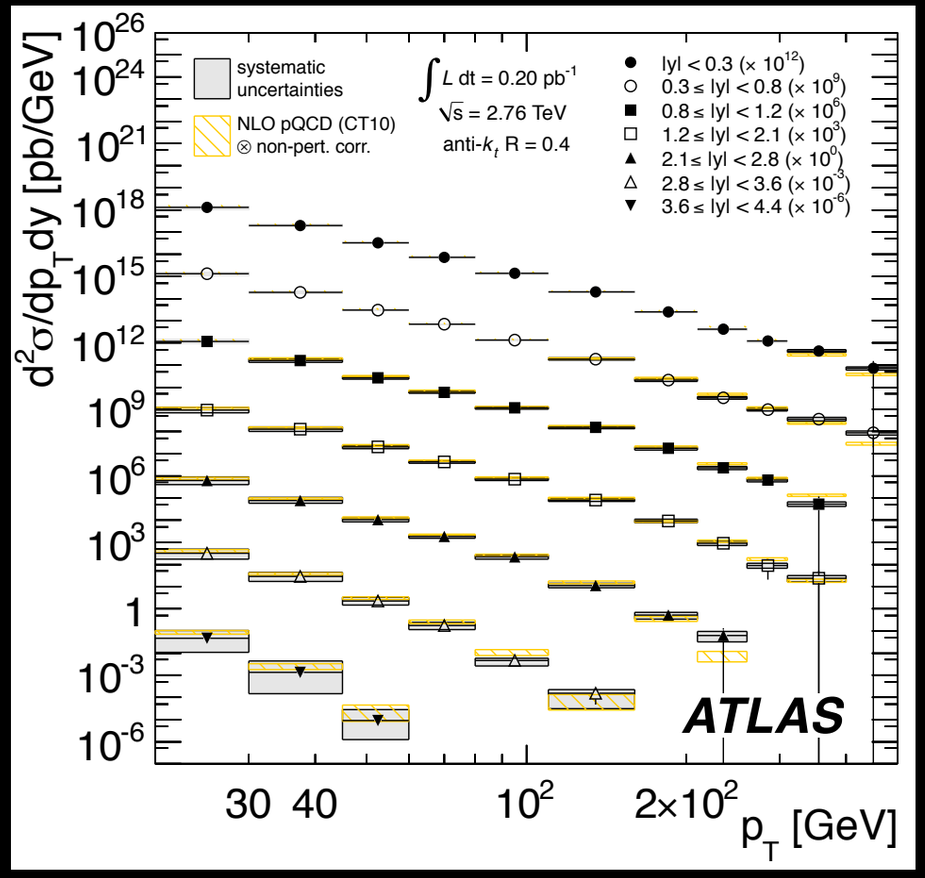
- additional issue for VH and ttH channels is the kinematic variation of the final state as a function of k_λ (NLO EWK MC would be needed to properly take into account acceptance effects), finally, single Higgs background in hh (particularly relevant for γγbb) should be properly rescaled as a function of the k modifiers, and proper MC generators should be setup to allow it

4) Total cross section vs k_λ + uncertainties
The cross-section σ(pp →HH) (κ_L) is available at LO (finite m_t) and at NNLO+NNLL ($m_t \to \infty$)

5) NLO as a function of k_λ values with exact finite mt correction will be provided in the next year

2) How to make results public (J. Allison, Max Swiatlowski, K. Leney)

Theoreticians ask to have the differential m_{hh} distribution
They would like to have something like this



The equivalent would be $\frac{d\sigma}{dm_{hh}}$

- **but we don't observe any hh pair, therefore , we cannot build that distribution**
- **we can give m_{hh}^{reco} of background events, but cannot unfold (there isn't any hh at truth level)**

Solution 1

We can give them the full likelihood $\mathcal{L}(x_i, \theta_j, \mu)$

parameter of interest

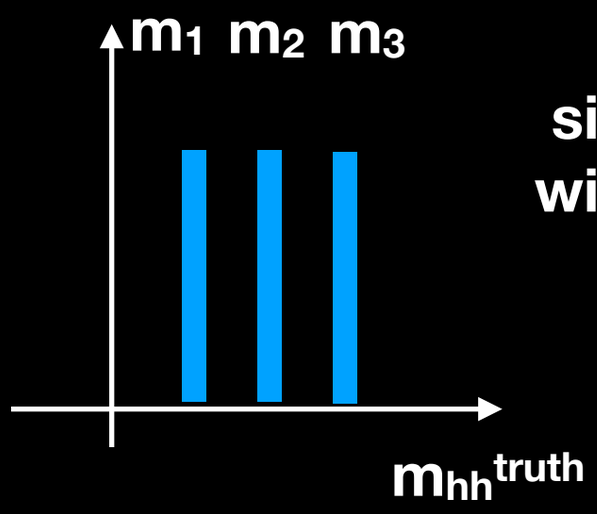
nuisance parameters

data yield in a given bin

Typically:

$$\mathcal{L}(x_i, \theta_j, \mu) \sim \prod_i \frac{(B_i + \mu S_i)^{x^i}}{x^i!} e^{-B_i - \mu S_i}$$

B_i background S_i signal



signal generated with flat m_{hh}

build one likelihood for each $m_{1,2,3...}$

$\mathcal{L}(x_i, \theta_j, \mu, m_1)$ $\mathcal{L}(x_i, \theta_j, \mu, m_2)$

$\mathcal{L}(x_i, \theta_j, \mu, m_3)$

Use RooMorph to interpolate the signal template

$$\mathcal{L}(x_i, \theta_j, \mu, m)$$

$\mathcal{L}(x_i, \theta_j, \mu, m)$ likelihood for a NWA signal of mass m

The theoretician provides $\frac{dP}{dm}(m)$

We compute: $\int \mathcal{L}(x_i, \theta_j, \mu, m) \frac{dP}{dm} dm = \mathcal{L}_{int}(x_i, \theta_j, \mu)$

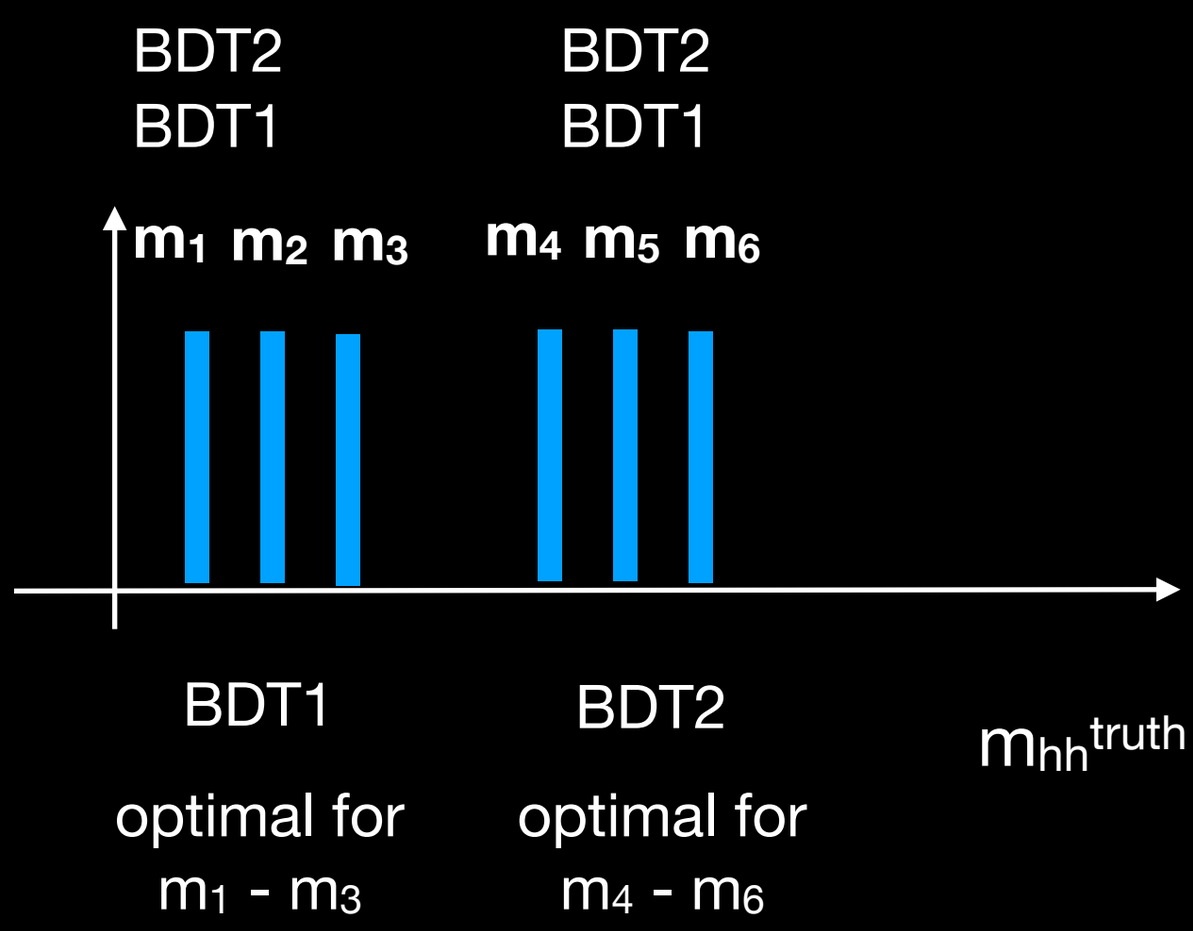
compute CLs upper limit cross section

Caveats:

- need to rename NPs to not give too many details on systematics
- theory systematics could be mantained
- it needs a granular mass grid of MC samples
- need to check behavior for low MC stat. (probably limit degradation)
- limited a single-state search
- problems with mass optimized BDT

How to handle mass optimized BDTs

$\mathcal{L}^{\mathrm{BDT1}}(x_i, \theta_j, \mu, m)$ $\mathcal{L}^{\mathrm{BDT2}}(x_i, \theta_j, \mu, m)$

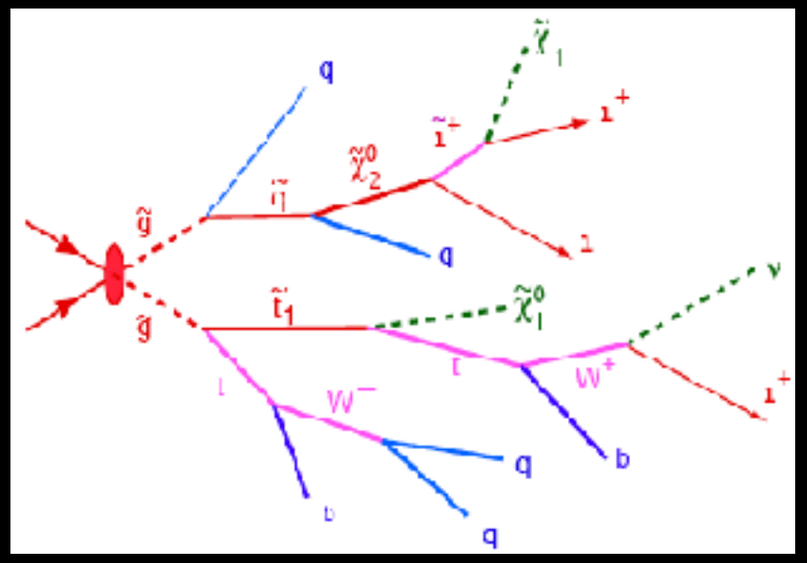


BDT1 and BDT2 are computed for all mass points (not only for their optimal points)

the signal can be probed with both BDT discriminants: BDT1, BDT2 and the most performant for the desired signal can be taken

- we should anyway publish some model interpretation
- M_X - Γ_X scan, benchmarks and so on (results could be produced with the framework itself)

The signal topology could be, in general, more complex



- some discussion was on going that one should do less optimized 1 bin analysis to make interpretations easy (John wanted a beer, discussion ended)

full simulation

Theoreticians provides LHE files from MC generation.
The full simulation chain for signal is applied.
This signal is used to compute S_i , build the likelihood and extract the upper limit

smearing

Theoreticians provides LHE files from MC generation.
The objects are smeared according their resolution.
This signal is used to compute S_i , build the likelihood and extract the upper limit

truth level

Theoreticians provides LHE files from MC generation.
The truth objects are used instead of the reco ones.

It has been shown that in some cases it works
(Good Luck!!)

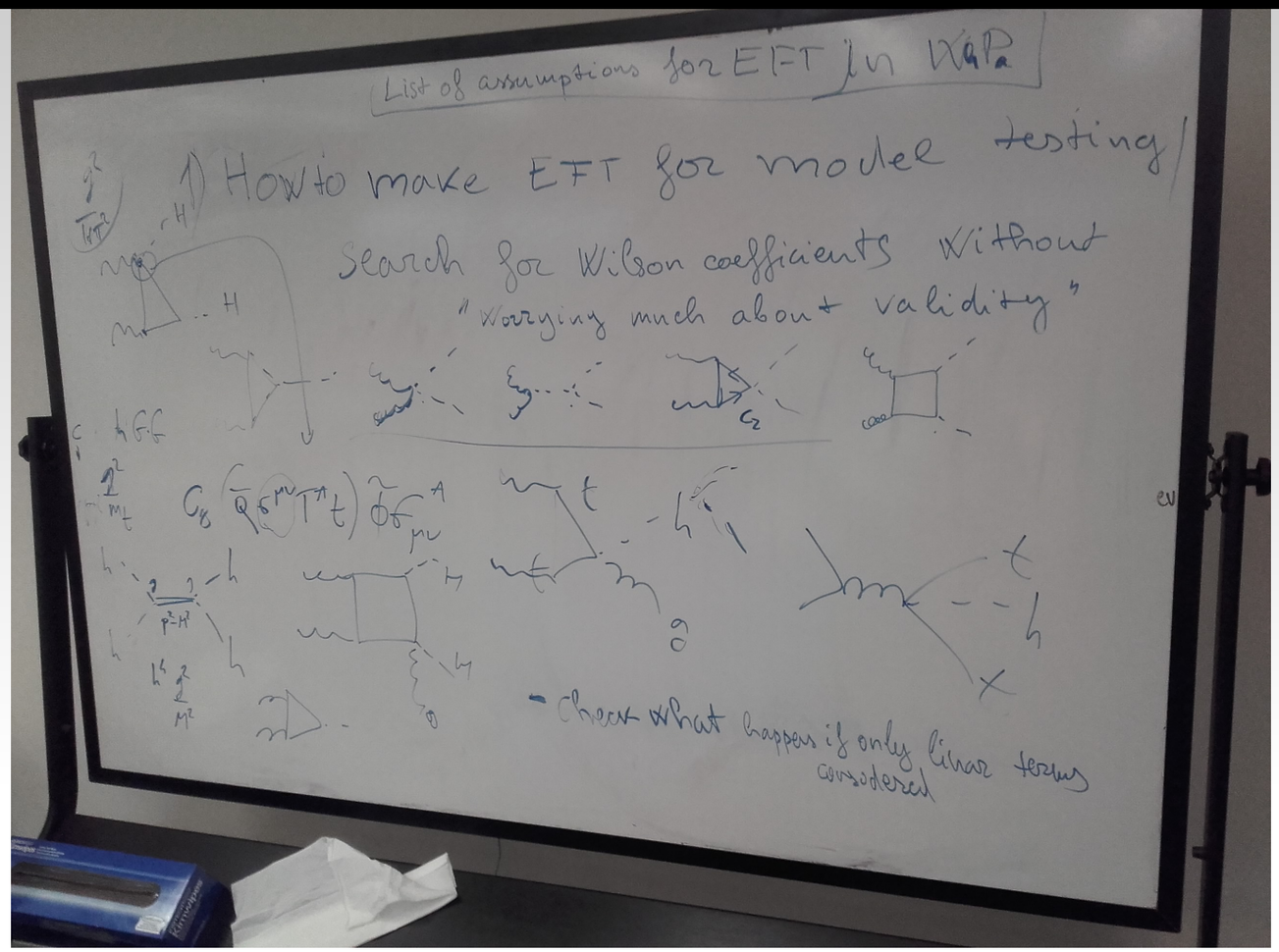
EFT

S. De Vita, M. Gouzevitch, J. Robinson

1) Is EFT useful

YES!!

We believe it is the right way to address anomalous couplings.



- EFT is “model independent” but requires assumptions that we shall clearly recall in the WhP.
- Define classes of models → and operators to include
  - Operators that affect Higgs physics « only » to be included at least.
  - Chromomagnetic or 4 fermions (with tops) operators not considered till now (expected to be small in some classes of models) but can also be considered in future analyses.
    - Linear, not linear; Weakly/Strongly coupled models; Dim6^2 vs Dim8*SM.
- EFT validity: energy << cut off; maximal size of coefficients for perturbativity. Headache for theorists. Experiments shall provide maximal scale used for constraints.

1) Example of EFT assumptions (other can be made)

My working assumptions

- Linearly realized EW symmetry (h belongs to Higgs doublet) ⇒ SMEFT
- Keep operators O_i up to dimension-6
- Operators tested in processes w/o Higgs assumed to be constrained
- Work in the Higgs basis ⇒ trilinear interaction $\lambda_3=\kappa_\lambda\lambda_{SM}=(1+\delta\kappa_\lambda)\lambda_{SM}$
- Further simplifying assumptions (just to limit # of O_i)
  - no CP,L,B-L, violating O_i
  - no dipole O_i
  - flavor universality
  - no ψ^4 (t^4,ttqq,q^4)

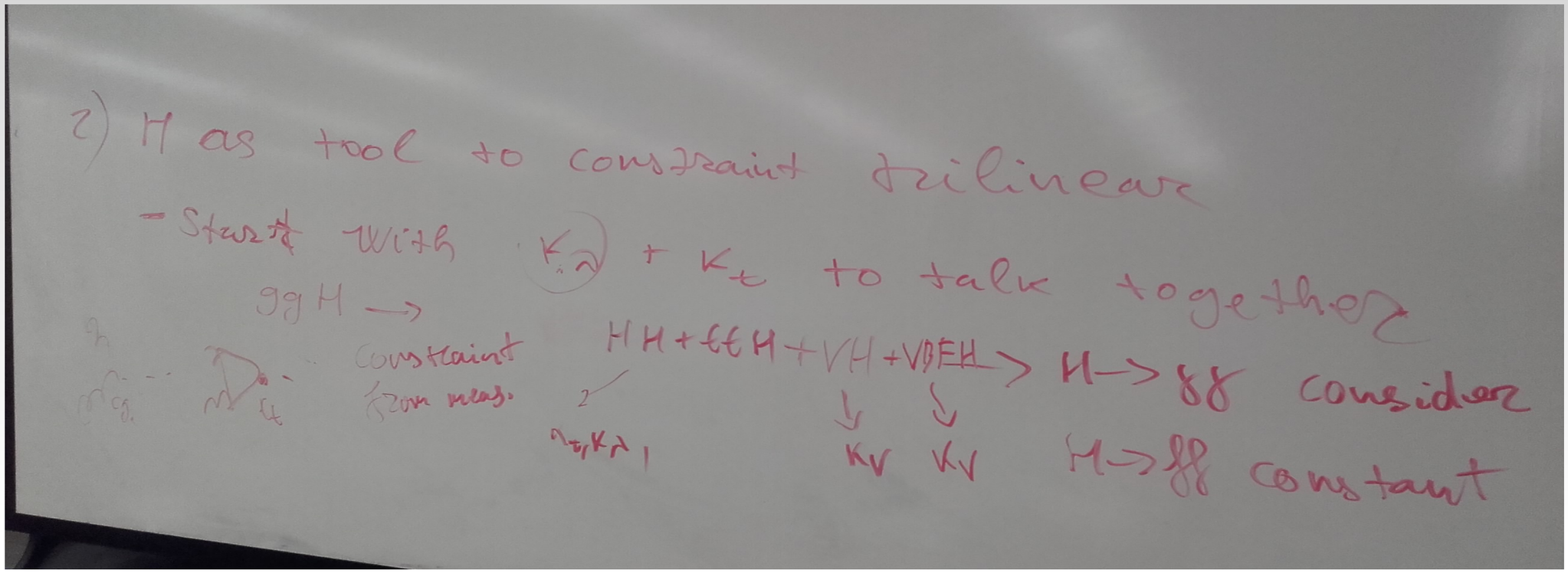
Pomarol '14; +Gupta,Riva '14; Falkowski '15; HXSWG YR4

$$\mathcal{L} \supset \boxed{\mathcal{L}_{\rm SM}} + \cancel{\mathcal{L}_{d=5}} + \boxed{\mathcal{L}_{d=6}} + \cancel{\mathcal{L}_{d=7}} + \cancel{\mathcal{L}_{d=8}} + \dots$$

L violating — B-L violating — subleading wrt d=6

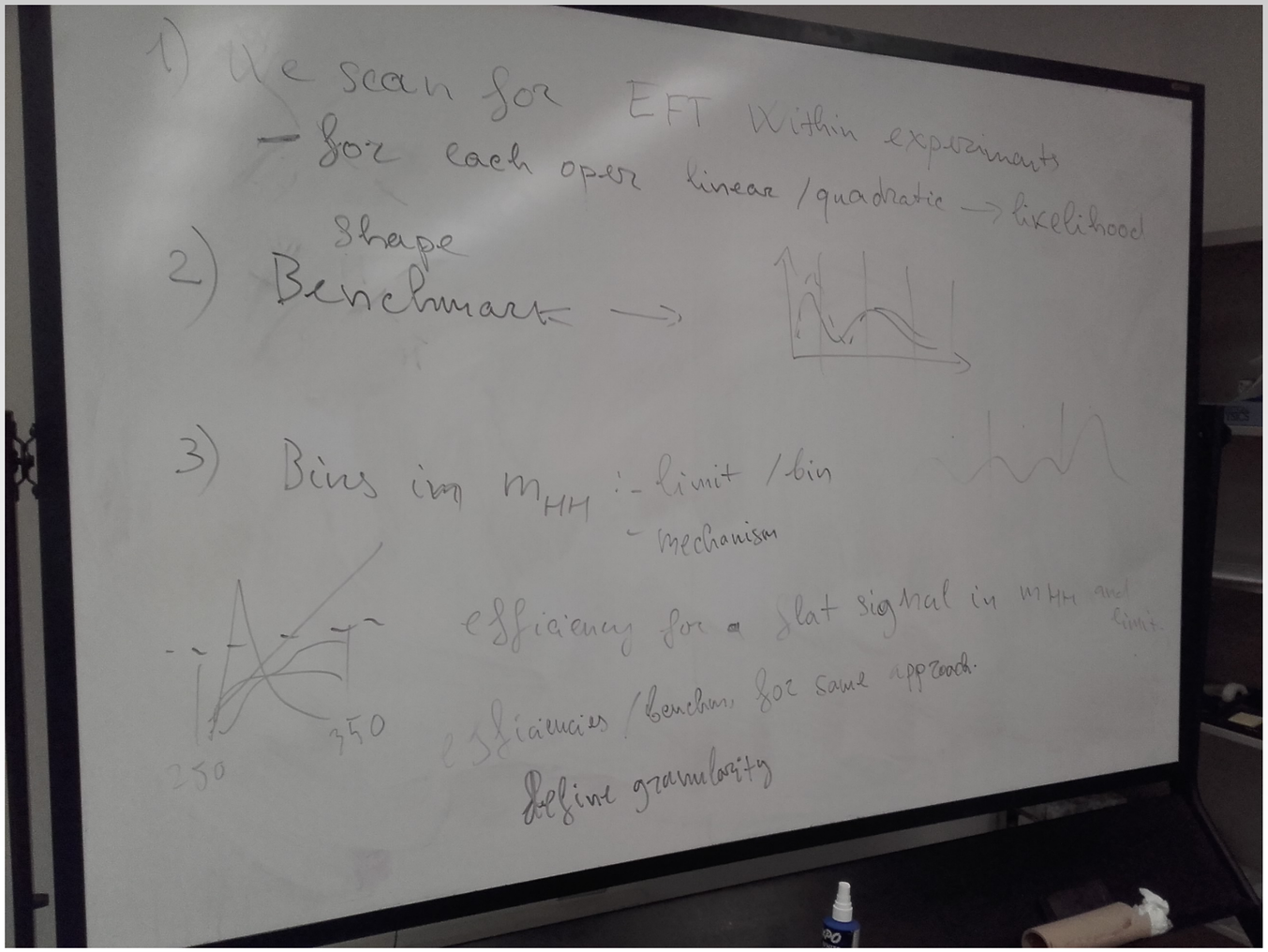
Focus on 10 O_i relevant at the LHC (not just SM tensor structures! EFT ≠ k-framework)
⇒ 10 independent deformations of hGG, hψψ, hWW, hZZ, hγγ, hZγ, hhGG, hhψψ, **hhh**

2) HH + H



- For studies with large kL: crucial to study together H and HH. KL would affect bounds on single Higgs couplings.
  - Ultimately we need a global fit: H + HH + aTGC.
  - As a first (training step): HH + ttH (+ VH + VBF H)
    - Let kL, kT (kV) float for a combined fit including 1 loop EWK corrections.

3) Presentation of the results



- Direct scans of Wilson coefficients by experiments: always welcome.
- Shape benchmarks:
  - Check if Chromomagnetic shape is covered by one of the benchmarks.
  - Keep providing benchmarks and maintain a tool to convert easily limit on benchmarks to limit for a particular combination of Wilson coefficients.
- Provide information on MHH:
  - Limits or profile likelihood per bin.
  - Info on efficiency for different shape benchmarks or even « flat sample ».
  - Guarantee a sufficient granularity to have small efficiency fluctuation inside each bin.

BSM 1 M. Carena, K. Tschann-Grimm, Ian Lewis, Lian-Tao Wang, X. Carvalho

Simple models that could be included in searches:

S-channel resonances: 1) spin-0 2) spin-2

higgsino→ hh+MET or higgsino→ hh+jets

X→hh where one h is 125 GeV and one is NOT 125 GeV, also X→ Vh where h is not 125 GeV

We should move away from RS models, which were firstly introduced to have sizable cross sections to be probed,

- on the other side, models where hh is a leading channel are difficult to find, unless tuned like the EWK singlet model, but needs tuning to survive; one could build up simplified models. Simplified models, like in DM searches, that gradually develops could be a way

- other possibilities follow a similar path as for the program for Higgs exotic decays. Eg for the resonant case: list the several free parameters (in all generality 3 masses, 3 widths, 1 spin structure, and 1 cross section) and then see the several signature topologies not yet look for. Start with those similar to signatures already looked for.

ACTION ITEM: GO FOR SIMPLIFIED MODELS

BSM

Other di-Higgs searches: H1+H2; Resonance->Sh, Resonance->Zh;

m_S = 60 - 250 GeV

- interference between resonant and non resonant could become important in configuration with large yukawa couplings, being able to produce a signal at ~60xSM
    - not already excluded by Higgs couplings??

objects/analysis strategies (M. Kagan, F. Micheli, K. Leney, C. Vernieri)

How can analyses do classifications based on b-tagging scores? How do you derive/apply scale factors?

— Derive scale factors in different bins (pseudo-continuous)
— CMS has fine binning and interpolates scale factors between bins to consider in a continuous

Can we estimate an uncertainty on the BDT output score directly?

- done by CMS using continuous calibration for distributions

ATLAS b-tagging seems to outperform CMS', possible reasons:

- different algorithms (one extra algorithm: JetFitter), extra detector IBL
  - results are shown with different p_T threshold, it is probably needed a coordination to ensure that plots are comparable between experiments;
- unify b-tagging plot style ? rejection vs. mistag, vs. tau background?
- unify ATLAS/CMS definition of H-tagging, b-jet for labelling **(MC truth)**

BDT versus cut based

- lively discussion on BDT versus cut based
- BDT always suspicious because look magic, need to understand what event feature the BDT is catching
- in ATLAS analysis, big improvement seems to come from the fit shape (careful shape systematics), if this is the case it is in principle possible to design a cut based analysis performing as the BDS (how much complex it would be, do you remember run-1 H->bb??)

6) How to improve WWbb (W off-shell) [S. Shrestha, N. De Filippis]

Question: Since the sensitivity (upper limit on mu) is of around 100 is it worth to continue to work on bbWW and bbZZ?

- high WWbb motivates a new pheno studies, trying to improve background rejections, we should expect new results on pheno studies for bbWW by the end of this year, 4b seems more difficult to study,

From phenomenological point of view, there is a growing motivation to study channels with higher BR given the need for higher statistics to reach SM sensitivity. Because of the phenomenological challenges to study 4b final state, bbWW is being explored with new cuts+MVA techniques, and the results are expected by the end of this year, same technique was used in the bbgamgam channel (arxiv: 1704.07395)

From experimental point of view, it was generally agreed that bbWW continues to be of interest because the two experiments show very different sensitivity, and there is much that remains to be improved.

Both ATLAS and CMS are trying various methods to reconstruct neutrino longitudinal momentum in case of single and two lepton channels.

likelihood-based neutrino pz reconstruction has been studied in CMS and seems to be promising. Build likelihood once offshell and again with onshell leptonic W decay, and let the likelihood decide which one is onshell/offshell. CMS has been requested to do truth-matching studies and get the fraction of times the nu pz is correctly reconstructed.
ATLAS going to test the same method.

Timeline: new results from CMS on bbZZ and bbWW with 2 leptons, different analysis due to different kinematics and background contributions

Good discussion time and everyone learnt from each other

How to ramp up

ideally: try to write a white paper, with the idea expressed and follow up on their development

it should be a sort of text book, summarizing all the on-going effort in harmonization and combination and setting it on stones

it needs coordination with HXSWG (think at it as a HXSWG note? better as proceedings with follow up discussed inside the Higgs cross section wg?)

what to put inside the white paper? (talk proceedings? are they needed? development of discussed topics and finalization ?)

what timeline?

where to publish (arXIv, Fermilab Note, CERN Note) ?

Why to do an LHC related workshop at FERMILAB

you can have real fun!!

Thanks to all for coming!!