Challenges in Higgs Combination

Andrea Carlo Marini, Roberto Di Nardo

HIGGS2022

Introduction

Combinations of Higgs boson experimental measurements are extremely powerful to shade the light on the Higgs sector

- Allow for a coherent interpretation of the overall Higgs sector
- Often measurements in multiple channels
  bring complementary information
- The best sensitivity achieved by exploiting the full power of our recorded data

Combinations are also a challenging effort:

- Complex input analyses imply a (even) more complex combined model
- When starting to think about cross-experiment combination

  → Important to set a common ground that can ease the combination process

  - both on physics and technical

Outline

- Historical perspective
  - Single experiment combinations
  - First Higgs LHC combinations
- Combination Tools
- Recent examples of single experiment combination
  - Higgs boson mass
  - Differential cross section measurements
  - STXS
- Interpretations: k-framework, EFT
- Towards combined ATLAS-CMS Higgs measurements (+interpretations) and global combinations

- Disclaimer: some of the considerations/points raised reflects our personal view and open for discussion

The journey to Higgs comb.

- Long journey started way long the Higgs boson discovery

- First step (back to ~2008) was to set a common ground in terms of statistics tools to being able to interpret the first data
  - RooStats interfaced to RooFit
  - Common discussion between ATLAS & CMS Stat. Committee

- Not enough to have a Tool → need to set up also a consistent procedure to being able to combine measurements
  - Particular important for cross experiment combinations

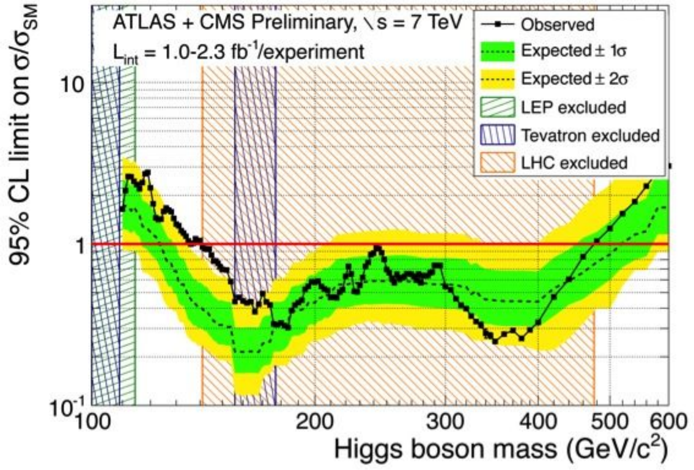
Not only a matter of tools…

- In Dec 2010 formed the LHC-HCG
  - Charged to combine CMS & ATLAS
  - Be ready to combine across experiment the first ~1/fb of LHC data
  - established on common methods to report exclusion limits, quantify excesses
  - Agreed on common syst. uncertainties together with their modelling and correlations
- The LHC H(XS) WG (born in Jan 2010) played a crucial role in defining samples, uncertainties, general procedural recommendations, and interaction with the theory community

Procedure for the LHC Higgs boson search combination in Summer 2011

The ATLAS Collaboration
The CMS Collaboration
The LHC Higgs Combination Group

CMS-NOTE-2011-005
ATL-PHYS-PUB-2011-11



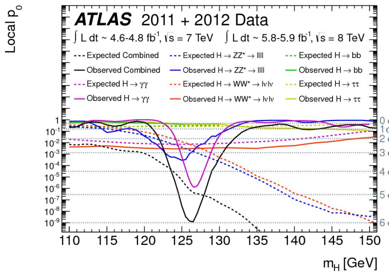
Historical perspective

Higgs Combination (single experiment)

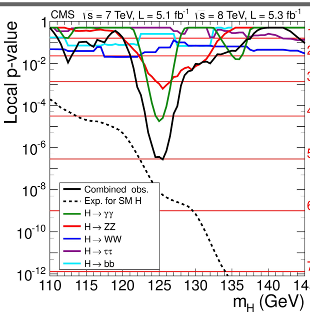
- In Run I started to be evident that many channels Share a similar excess.
- Need to understand the overall compatibility with the SM Higgs boson
- Define overall standards and procedure (Run I)
- Fitting framework

Run 2:

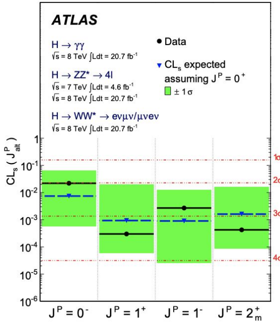
- Increased analyses' complexity: STXS in many final states
- Larger and larger sensitivity to differential Observables
- Deeper interpretations and connection to BSM Effects and operators
- Differential measurements combinations
- HH combinations



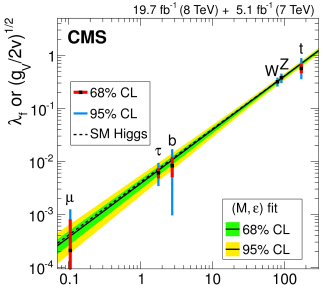
PLB 716 (2012) 1-29



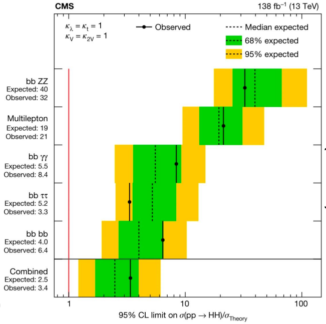
JHEP 06(2013)081



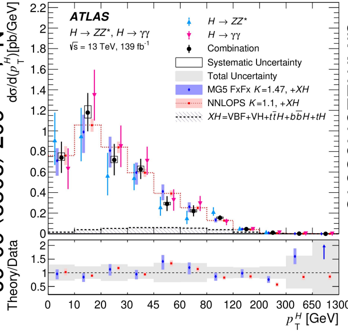
PLB 726 (2013) 120



EPJC 75(2015)212



Nature 607 (2022) 60-68



arXiv:2207.08615

Historical perspective – LHC

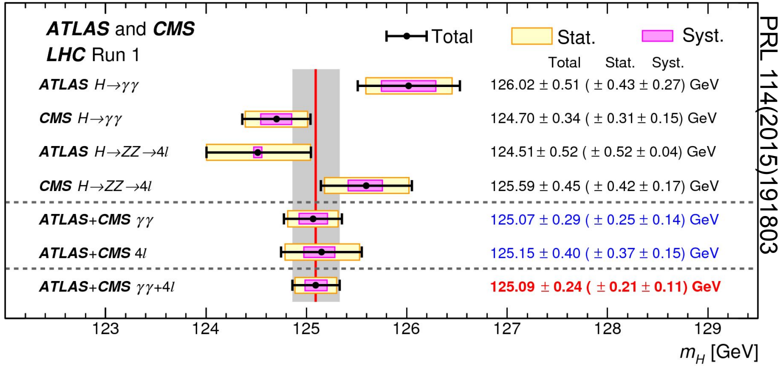
At the end of Run I, the experiments envision the effort of joint results:

- Technical description of the likelihood function
- Discussion and agreement on methodology,
- meaning of systematics,
- parameters of interest

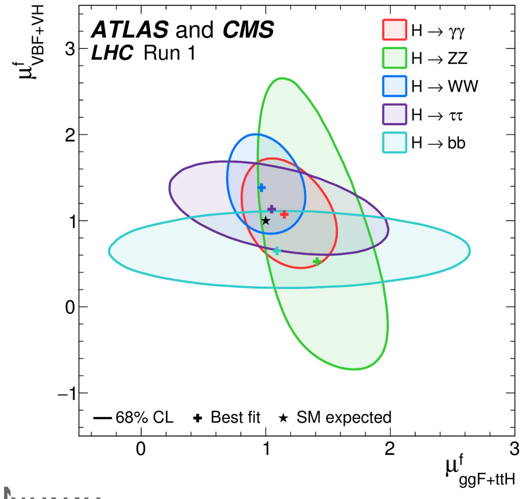
Two major results:

- Mass combination - m_H=125.09±0.24 GeV
- Couplings Measurements
  - ~4200 nuisance parameters, ~600 signal/control regions

Set common ground for Run II measurements



PRL 114(2015)191803



JHEP 08(2016)045

Higgs Combination Tools

CMS from the beginning developed a coherent effort on the tool.

- Inputs (datacards) used to construct the likelihood functions are saved and parsed
- Code has been optimized to boost performances

ATLAS uses a tool fully based on Roofit

- Powerful tool allowing for combination and manipulation of likelihood models (workspaces)

Run 2 take home:

- Big step in Run 2 with the implementation of analytical minimization of the Barlow-Beeston light uncertainties: typically a factor x1000 faster in mid-sized analyses
- modify datacards content on the fly
- Interaction with RooFit community is beneficial

The Higgs boson mass

First LHC Higgs combination with full run1 LHC data

- Unknown parameter in the SM, input for couplings and differential measurements

Currently available ZZ and γγ ATLAS and CMS results with partial or full Run2 data

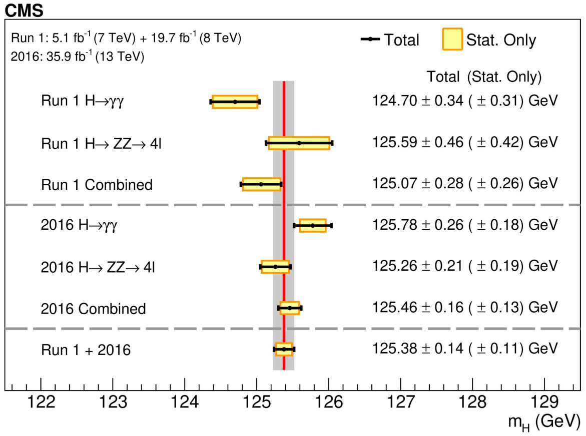
- Dedicated studies: assessment of correlation schemes for exp. systematics
- Stat. uncertainty dominant

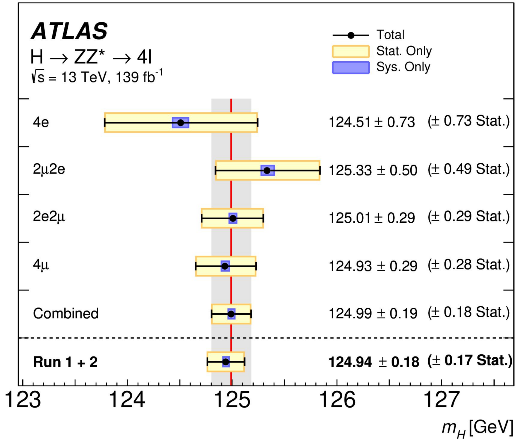
→ An ATLAS-CMS mass combination with full Run2 data could set a new m_H reference

- Combination with Run1 profiting from correlation schemes studied in single experiment

Per-event models used → additional challenges for the combination

PLB 805(2020)135425





arXiv:2207.00320

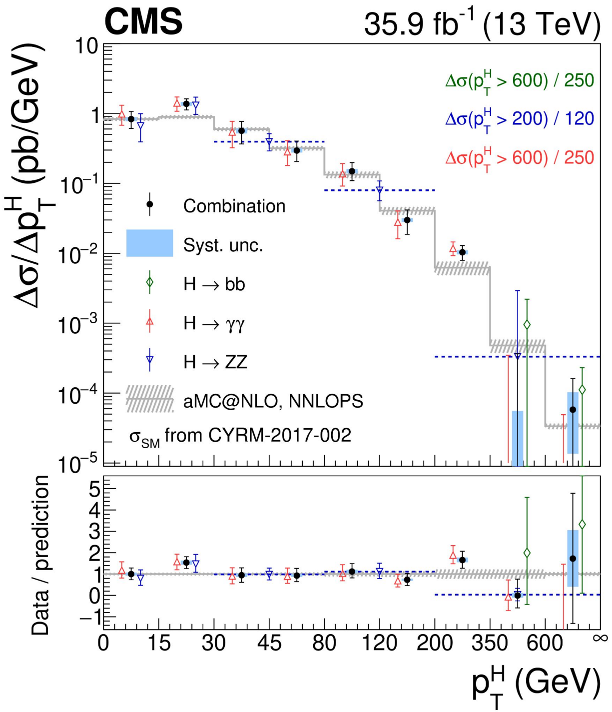
Differential measurements

Differential cross-section measurements (usually) performed in the fiducial phase space.

- bb, γγ, ττ, ZZ, WW final states available
- Impact of model dependence minimal
- Performed using likelihood unfolding (+ eventual regularization when needed)

Combinations performed for few specific observables (e.g. pTH, Njet) within single experiment

- Extrapolation to the full phase space needed
- SM assumption on the BR



PLB 792(2019)369

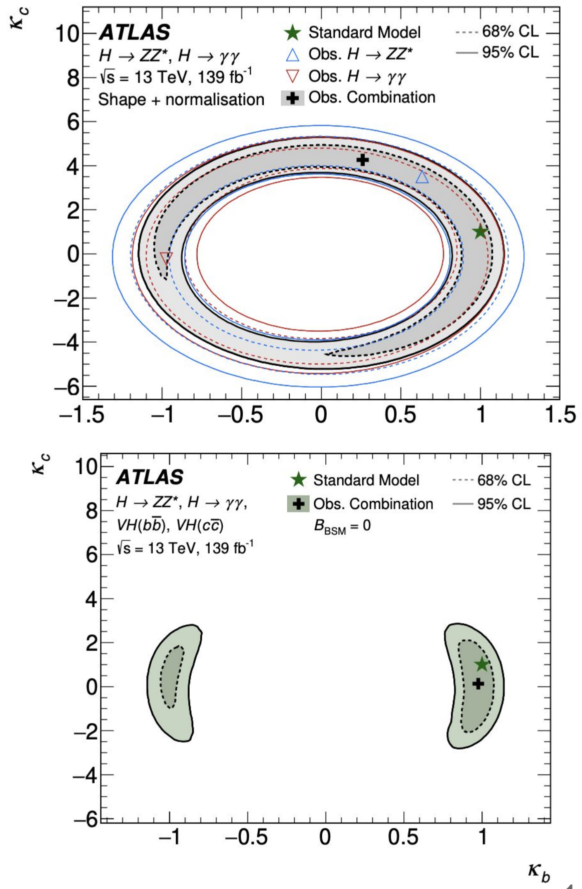
Differential measurements

Key points for a successful **combination**

- The bin-edges harmonization (across channels and, experiments)
- Common definition of the unfolded observables
- Definition of a common fiducial phasespace for possible ATLAS-CMS per channel combination

Combination of the measurements can also be limiting for some **interpretation**

- Use fiducial measurements from multiple channel for a combined constraint on BSM parameters used (e.g. k_b k_c using the p_TH spectra)
  - Need to know the impact of BSM parameters on a given observable in each fiducial phase space

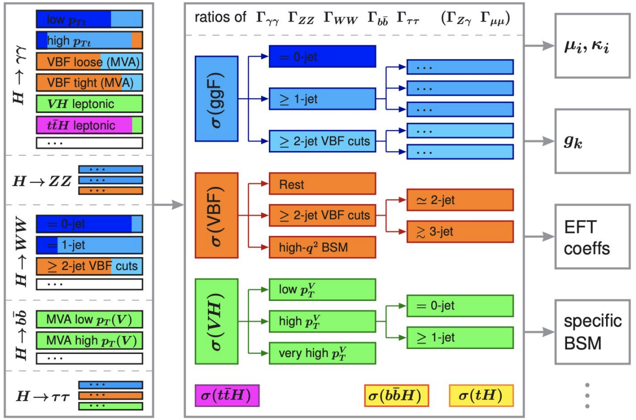


arXiv:2207.00092

Simplified Template XS

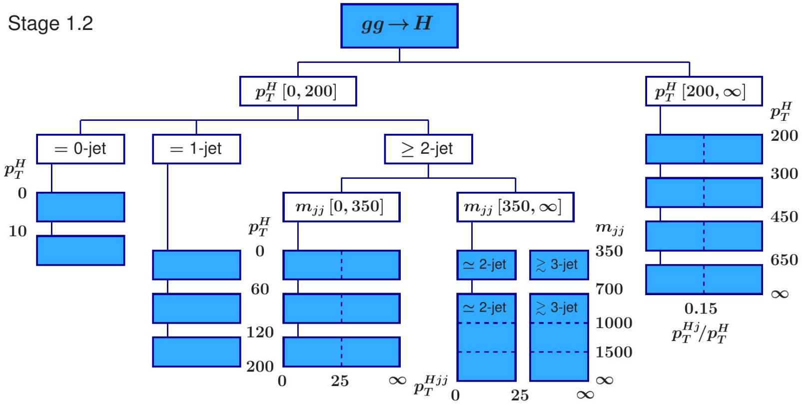
Natural evolution of σ x BR Run1 measurements

- complete framework: multiple analyses contributes in different phasespace
- Back-compatible for inclusive measurements and interpretations
- Key player to constraint more complex BSM models (e.g. EFT)



Trade-off between “measurable” & “interesting” (see high-pT) phase space

- Reduced model dependence in the measurement
- Same scheme used to develop th. syst. uncertainties prescriptions (within the LHCHWG)



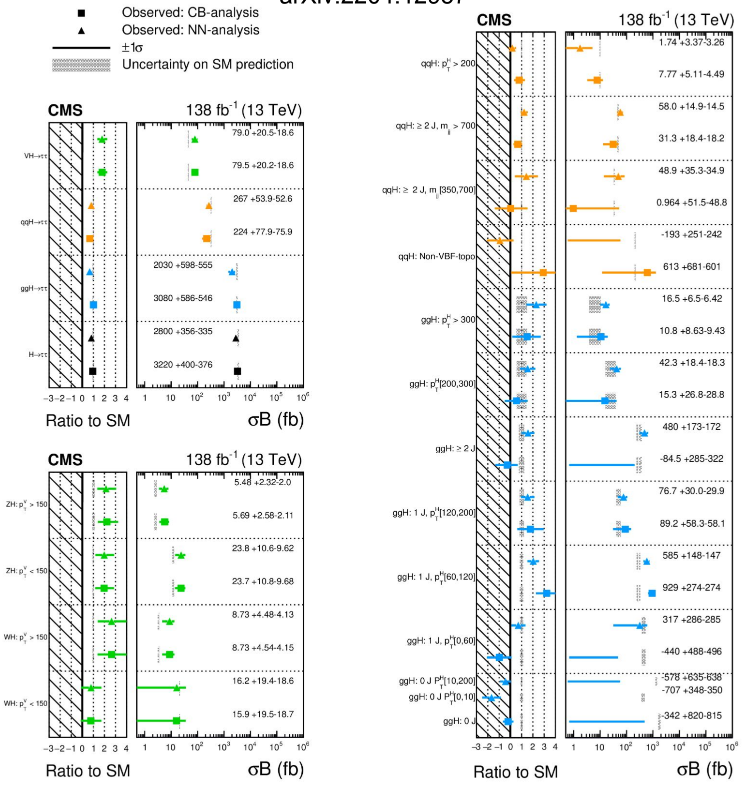
Simplified Template XS

Most of the analyses with (decent) sensitivity to the Higgs boson production published STXS measurements

- Complexity of the analyses increased and designed to explicitly measure each bin of STXS

- New challenges faced wrt Run1
  - Bkg and syst uncertainty evaluation in multiple analysis regions
  - Pruning processes not contributing to experimental channels
- Study merging scheme: both for single analyses and combinations

arXiv:2204.12957



Simplified Template XS

Increased complexity in the input analyses

→ n. of obj. in the likelihood goes ~n^2

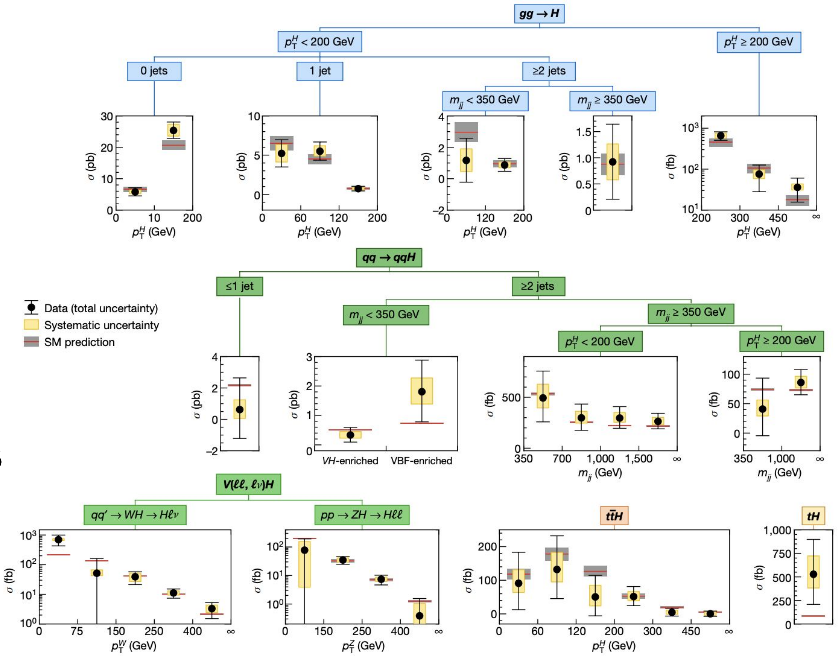
→ fully projected in the combined model

Models with >~3k nuisance parameters

Key aspects for the combination

- Check analysis overlap
- Check the meaning of each systematic uncertainty coherently used across analyses
- Nuisance parameters pruning
- Proper handling of the input analyses normalization factors and TH syst uncertainty on the backgrounds
- Dealing with nuisance parameters constraints & pulls

Nature 607, pages 52-59 (2022)



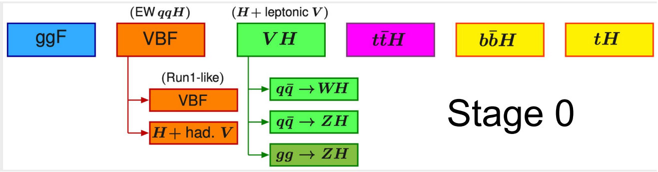
Simplified Template XS

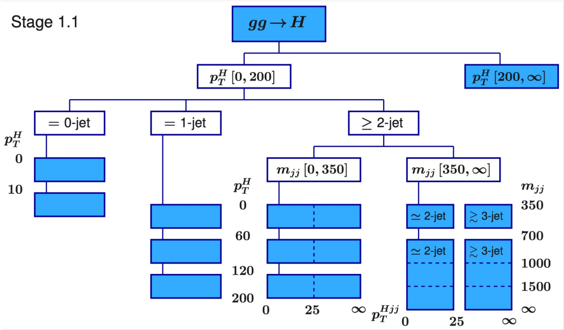
A caveat for combinations is represented by the evolution of the STXS bin definition

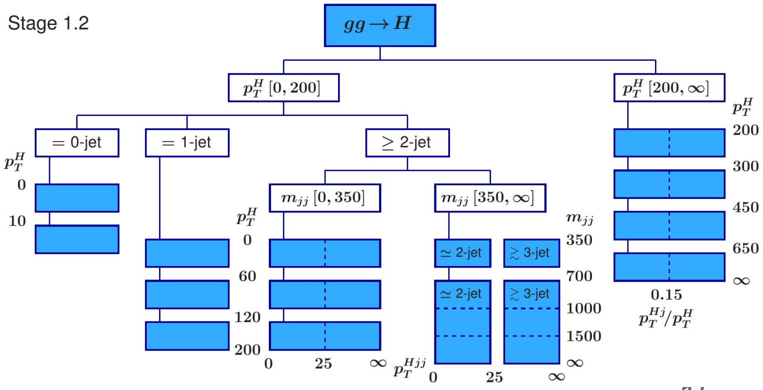
- Different “Stages” (Stage0, 1,1.1,1.2…)
- Redefinition of STXS Stage 1 bins implies: Update analyses, uncertainty schemes, BSM parametrizations each time …

In Run3 LHC is running at √s=13.6 TeV - Run2 @ 13 TeV

- Cannot “combine” cross sections
- Consider to measure also STXS → STμ (?)
  - If so, th uncertainties on the predictions fully back in our measurements







Interpretations: kappa framework

Perturbation around the SM value

Factorization of production and decay

Simple parametrization of deviations

→ Easy interpretation of possible deviations

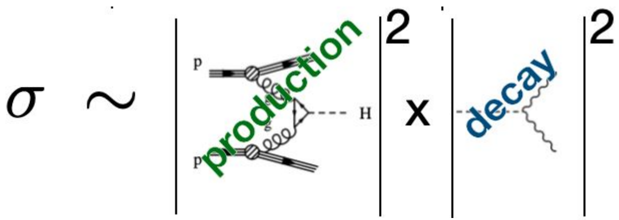
→ Easy reinterpretation on other models

Known Limitations:

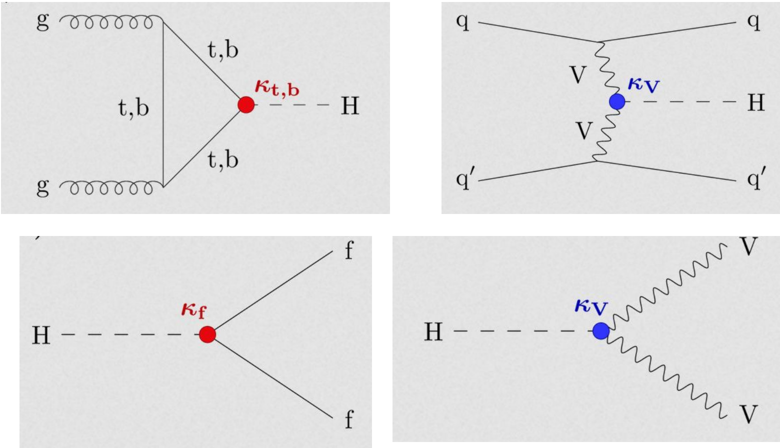
- NLO not coherently accounted in the parametrization
- Expected to break around % level precision (We are almost there!)
- Not a complete theory

Same parametrization used by both experiments

- No conceptual problems in a joint combination



$$\sigma_i \mathscr{B}_f = \left(\frac{\sigma_i \Gamma_f}{\Gamma_H}\right)_{\text{SM}} \cdot \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2}$$



Effective Field Theory

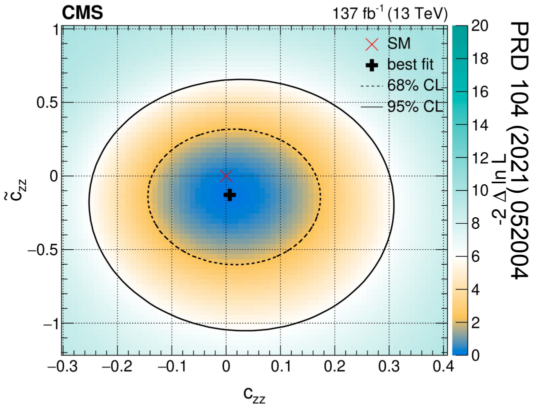
In Run2 ATLAS & CMS started to investigate potential BSM effects in the context of EFT

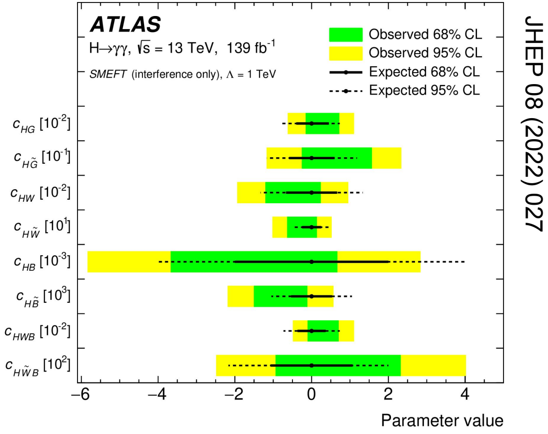
Several approaches used across different channels

- Reinterpretation of STXS
    - Not always optimal (e.g. hard to set constraints on CP odd operators)
- Fit to dedicated observables
- Reinterpretation of differential XS measurements (of multiple observables)

Many parameters in the model:

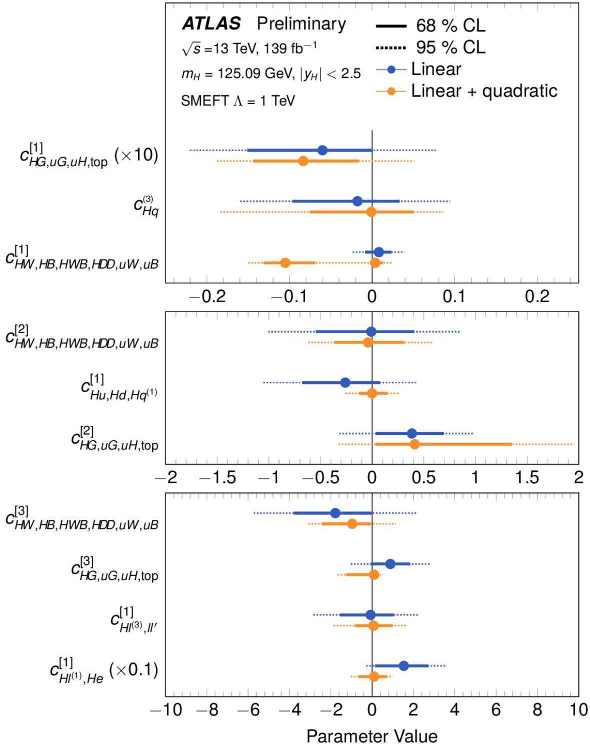
- Sensitivity only for a limited subset





Effective Field Theory

Initial effort to profit of the combined STXS measurements to constrain multiple Wilson coefficient



Needs to find a common ground for the fit definition:

- POI to fit, what to freeze, what to float.
- Linearized vs Quadratic model.
- Interpretation and validity of results.
- Common parametrization as for e.g. k_{λ}

Role of external measurements

- E.g. constraints from LEP, e.g. m_Z

Importance of the LHC EFT forum

- Discussion between theory & experimental communities
- Coordinate among the different sectors (SM, TOP)
- Background variations

Global EFT fits

First attempts to use measurements from multiple Analysis group (Top, Higgs, SM) together with LEP EWPO to constrain EFT Wilson coefficients

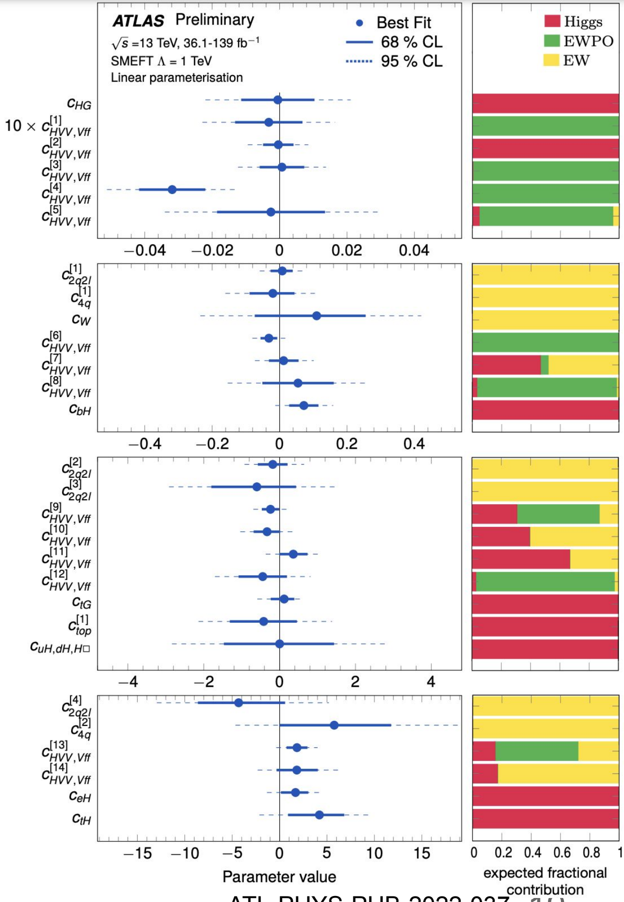
- ATLAS : Higgs+EW+EWPO
- CMS : Higgs+Top

Likelihood function combinations

- Definition of parameters of interests,
- Evaluation of the effects on several analyses

Simplified approaches:

- Used by phenomenology
- Partial likelihoods / partial effects correlations
- From first studies comparable wrt full likelihood combination
    - Checked only for the linear parametrization



ATL-PHYS-PUB-2022-037

What to expect for the future …

CMS & ATLAS are finalizing the measurements on the Run 2 dataset

- A lot of couplings, STXS, BSM interpretations and differential XS results ready
- Final Run2 m_H measurements close to be ready
- Most sensitive HH channels published

→There is the potential of several combination between ATLAS - CMS experiments in the Higgs sector

- Technical exchanges already happening between the collaborations for future Run 2 combined results
- Consider discussing with LHCb (VH,H→bb/cc)

Current Limitations

Complexity of current analyses raise still issue in large combinations:

- Large memory required to build (~16Gb for CMS July 22 combination) and fit (~6Gb) the overall model
- >3k nuisance parameters
- Complexity of the models
- Fitting time is increasing more and more
  - 2h-8h for the combined workspace creation
  - 2h-8h for the minimization (Migrad)
  - 4h-12h for Hesse

Large turnaround time from the first publication to the last

- Lost of expertise in time of the combination
- Reduced possibility of re-running analyses

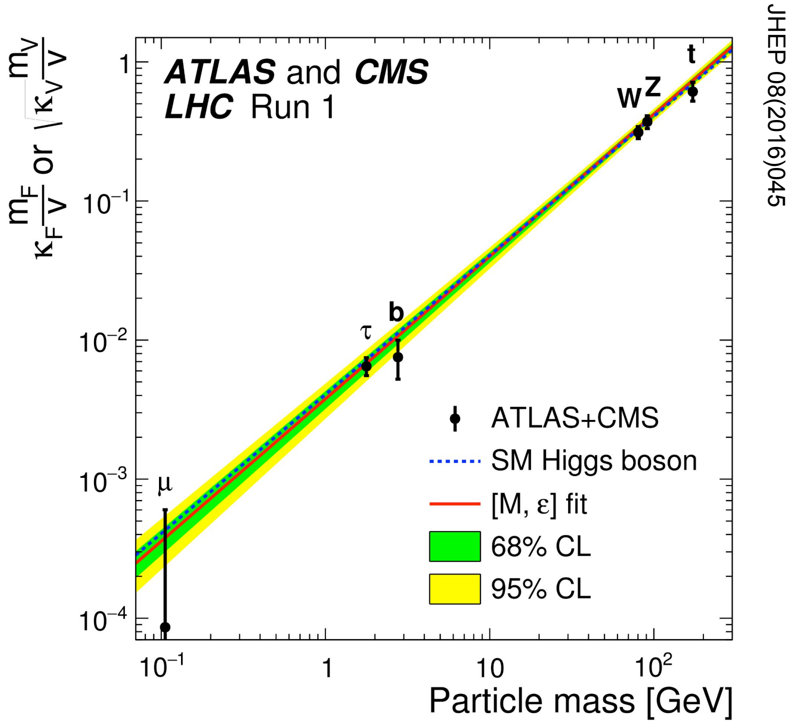
Summary

- Combinations in the Higgs sector are a powerful tool to shade the light on the Higgs boson properties

- Full engagement from the Collaborations is needed for a successful ATLAS-CMS combination programme

- Useful to learn from the experience and backpropagate recommendations to the analyses

- Combinations are very challenging



Thanks