Search for CP-odd ttH production with H to bb in ATLAS

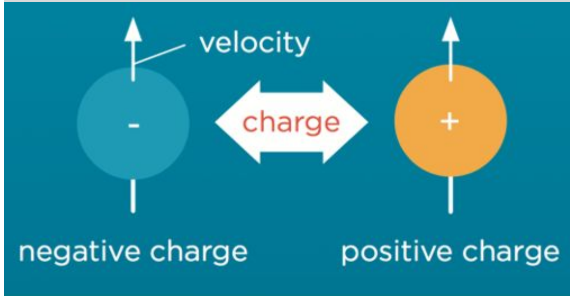
IDPASC workshop, June 25th 2020

Ana Luísa Carvalho
Ricardo Gonçalo, Patricia Conde Muiño

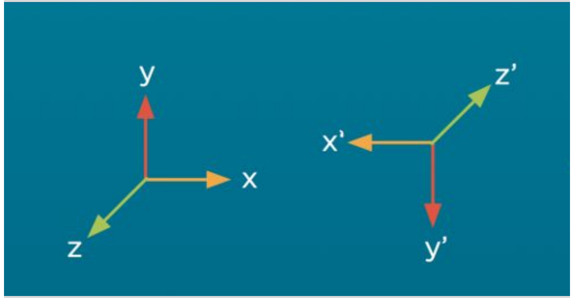
CERN/FIS-PAR/0002/2019
SFRH/BD/143435/2019

Motivation

Charge symmetry



Parity symmetry

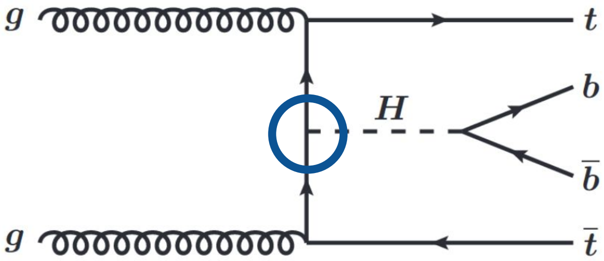


- According to the Standard Model of Particle Physics (SM), the Higgs boson is a CP-even particle, meaning that its properties are invariant under combined charge-parity (CP) transformations
- A pure CP-odd Higgs boson has been excluded at 99.98% confidence level

The presence of CP-odd components in the Higgs boson couplings is predicted by many beyond the SM models and its discovery would be a clear sign of new physics phenomena

Analysis overview

This analysis looks for CP-odd components in the interaction between the Higgs boson and top quarks



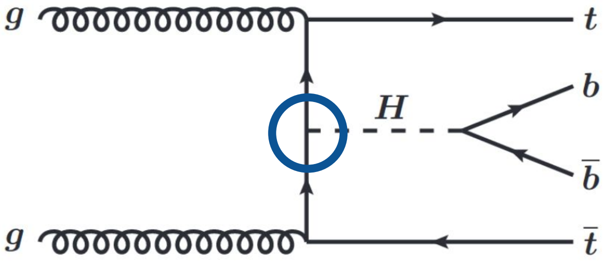
$$\mathcal{L}_{tth} = y_t \bar{\psi}_t \kappa'_t (\cos\alpha + i\gamma_5 \sin\alpha)\psi_t h$$

Pure CP-even

Pure CP-odd

- α is the CP-mixing angle that is predicted to be equal to 0 in the SM
- Event yields and shape of distributions of CP-sensitive variables are used as input to a likelihood fit to set limits on the value of α

Experimental signature

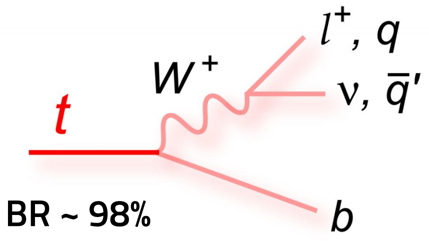


Both tops decay leptonically ⇒ **Dileptonic channel**

- 2 charged leptons
- 4 jets containing B-hadrons (b-jets)

One top decays leptonically and the other hadronically ⇒ **Semileptonic channel**

- 1 charged lepton
- 6 jets: 4 b-jets



Higgs boson has a high transverse momentum (p_T~300 GeV) ⇒ Two b-quarks are more efficiently reconstructed using a single jet with a large radius parameter ⇒ **Boosted channel**

Event selection

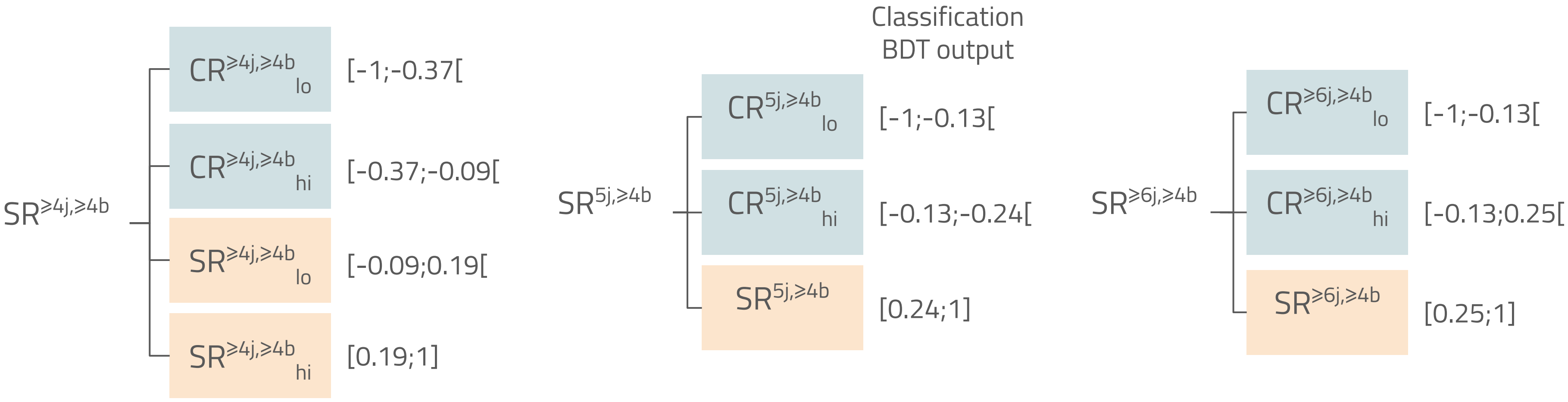
Trigger: ⩾ 1 isolated lepton with pT ≳ 20 GeV

In order to take advantage of the higher jet and b-jet multiplicities of the ttH signal process, events are classified into non-overlapping analysis categories

Regions	No. leptons	No. jets	No. b-tagged jets	No. Higgs candidates
$\mathrm{CR}^{3j,3b}_{hi}$	= 2 (dileptonic channel)	= 3	= 3	-
$\mathrm{CR}^{\geqslant 4j,3b}_{lo}$		⩾ 4		-
$\mathrm{CR}^{\geqslant 4j,3b}_{hi}$				-
$\mathrm{SR}^{\geqslant 4j,\geqslant 4b}$			⩾ 4	-
$\mathrm{SR}^{5j,\geqslant 4b}$	= 1 (semileptonic channel)	= 5	⩾ 4	-
$\mathrm{SR}^{\geqslant 6j,\geqslant 4b}$		⩾ 6		-
$\mathrm{SR}_{boosted}$		⩾ 4	⩾ 3	⩾ 1

Event categorization

Resolved inclusive signal regions are further divided according to the output of a boosted decision tree (BDT) trained to separate signal (ttH) from background (tt+b)



Classification BDT output

$\mathrm{SR}^{\geq 4j,\geq 4b}$
- $\mathrm{CR}^{\geq 4j,\geq 4b}_{\mathrm{lo}}$ [-1;-0.37[
- $\mathrm{CR}^{\geq 4j,\geq 4b}_{\mathrm{hi}}$ [-0.37;-0.09[
- $\mathrm{SR}^{\geq 4j,\geq 4b}_{\mathrm{lo}}$ [-0.09;0.19[
- $\mathrm{SR}^{\geq 4j,\geq 4b}_{\mathrm{hi}}$ [0.19;1]

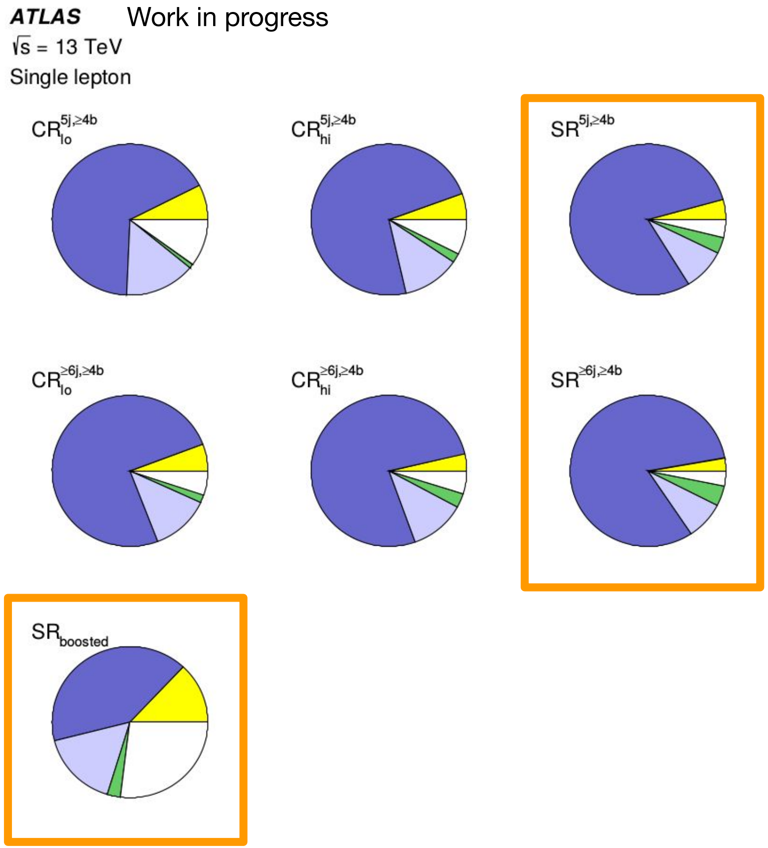
$\mathrm{SR}^{5j,\geq 4b}$
- $\mathrm{CR}^{5j,\geq 4b}_{\mathrm{lo}}$ [-1;-0.13[
- $\mathrm{CR}^{5j,\geq 4b}_{\mathrm{hi}}$ [-0.13;-0.24[
- $\mathrm{SR}^{5j,\geq 4b}$ [0.24;1]

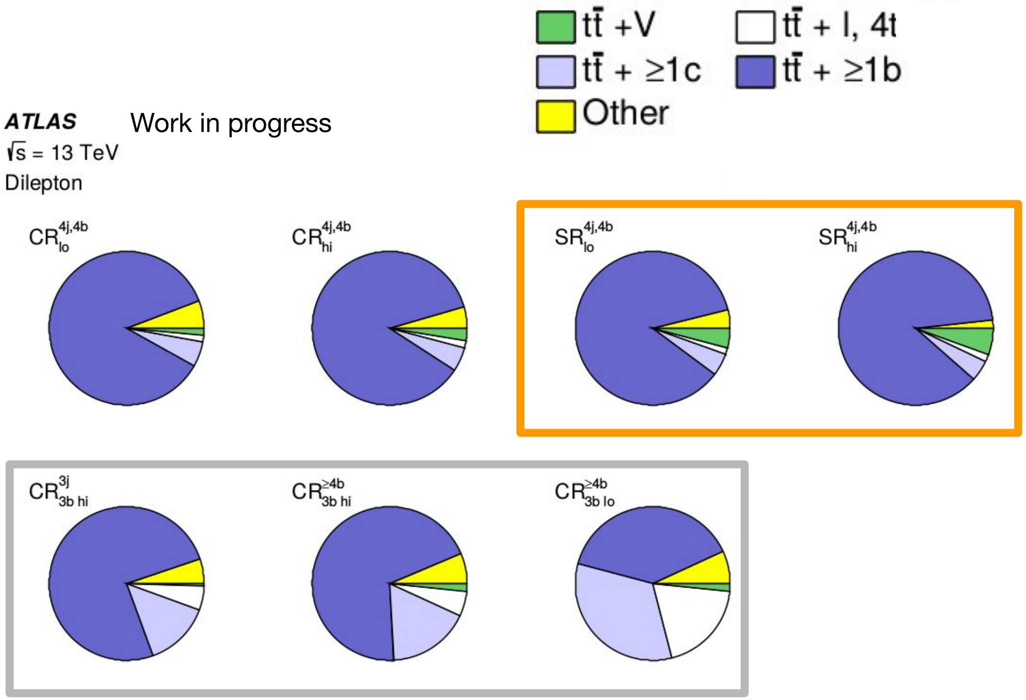
$\mathrm{SR}^{\geq 6j,\geq 4b}$
- $\mathrm{CR}^{\geq 6j,\geq 4b}_{\mathrm{lo}}$ [-1;-0.13[
- $\mathrm{CR}^{\geq 6j,\geq 4b}_{\mathrm{hi}}$ [-0.13;0.25[
- $\mathrm{SR}^{\geq 6j,\geq 4b}$ [0.25;1]

$\mathrm{CR}^{3j,3b}_{\mathrm{hi}}$ $\mathrm{CR}^{\geq 4j,3b}_{\mathrm{lo}}$ $\mathrm{CR}^{\geq 4j,3b}_{\mathrm{hi}}$ $\mathrm{SR}_{\mathrm{boosted}}$

14 analysis regions:
- 9 control regions (CR’s)
- 5 signal regions (SR’s)

Background composition





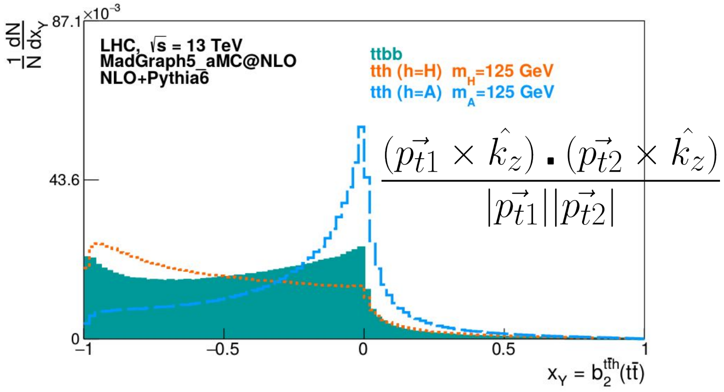
Used to control tt+≥1c normalization

Main background is tt+≥1b in all regions

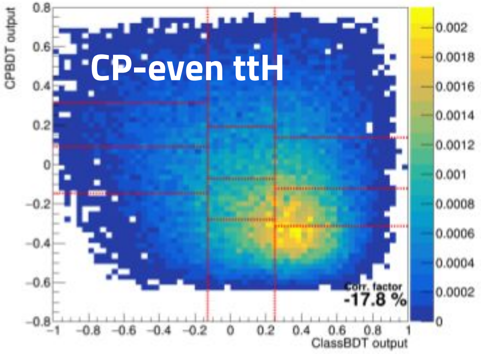
Sensitivity to CP structure

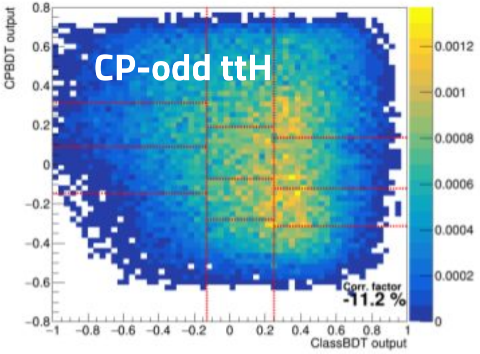
The output of a BDT trained to separate CP-even ttH events from CP-odd ones is used as the final discriminant

- Different BDT’s are trained for the inclusive 4-jet, 5-jet and ≥6-jet regions
- Input variables are chosen to optimize CP-even/odd discrimination in each category
    - Lab-frame observables probe boost of Higgs and top-quark systems
    - Angular variables probe helicity and spin correlations

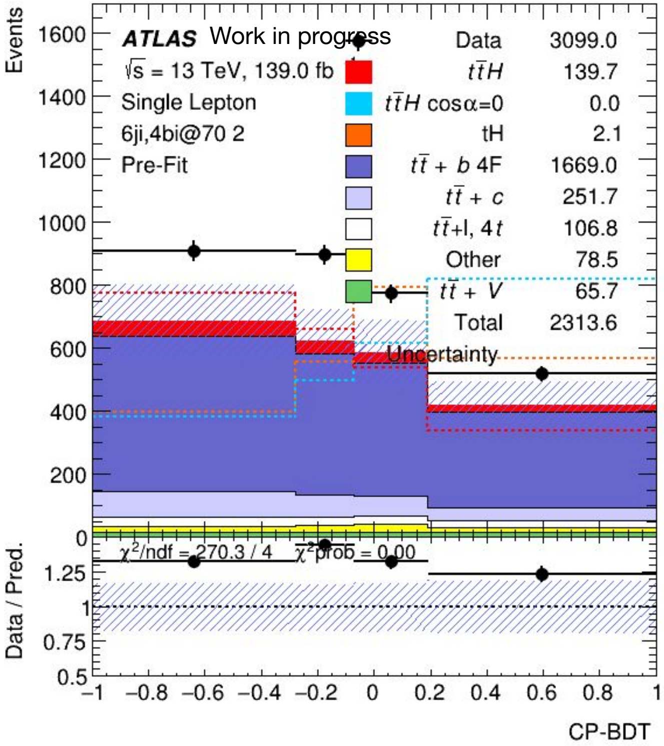


$$\frac{(\vec{p_{t1}} \times \hat{k_z}) \cdot (\vec{p_{t2}} \times \hat{k_z})}{|\vec{p_{t1}}||\vec{p_{t2}}|}$$





Semileptonic resolved analysis region



CP-even

CP-odd

- Example of distribution used as input to the likelihood fit
- Yield and shape of MC samples is allowed to change to best fit the data
- Data/MC agreement is not great pre-fit

- CP-even ttH shifted towards low CP-BDT values
- CP-odd ttH shifted towards high CP-BDT values

Statistical analysis

A binned likelihood fit is performed to extract the value of the CP-mixing angle (α)

- Likelihood function is constructed as a product of Poisson probability terms over all bins
- Normalization factor of tt+⩾1b background is free-floated
- A gaussian prior is used for the other nuisance parameters
- Measured best-fit value of α is the one that maximizes the likelihood function

$$q_\alpha = -2 \ln \frac{\mathcal{L}(\alpha)}{\mathcal{L}(\hat{\alpha})}$$

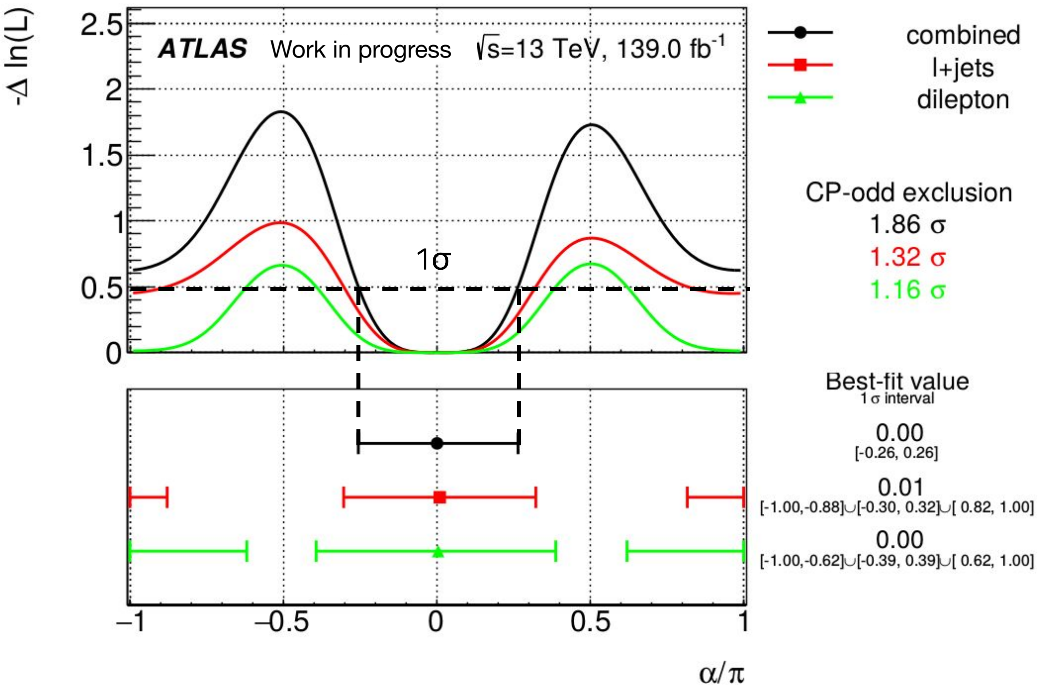
Value of α that maximizes the likelihood

Before looking at real data, it is possible to estimate the analysis sensitivity to the parameter of interest (α)
⇒ **Asimov fits**

- Perfect dataset following exactly the Monte Carlo description
- If everything is ok, this fit should return exactly the value of α that was used to generate the model

Analysis sensitivity

⇒ Asimov data set created using a CP-even (SM) Higgs boson

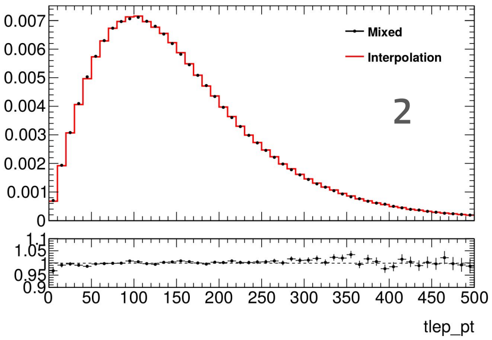


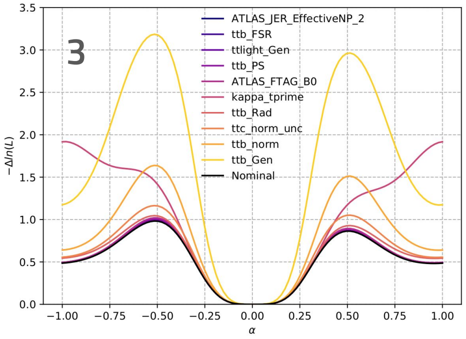
- For each value of α, the ratio of likelihood is calculated
- A larger value means a larger disagreement between data and Monte Carlo

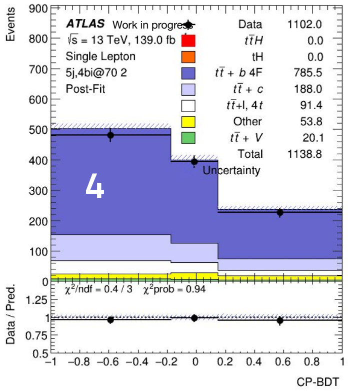
- Semileptonic channel drives the sensitivity
- Combined fit
  - ○ CP-odd exclusion of 1.86σ
  - ○ Expected uncertainty on α is 0.26π (47°)

Conclusions

- This analysis was designed and implemented by the LIP group
- Started contributing to the analysis ~6 months ago
- Contributions so far
  1. Asimov fits to estimate sensitivity and assess impact of different analysis strategies and systematic models (previous slide)
  2. Validation of the linear interpolation between CP-even/odd ttH (neglect interference)
  3. Investigation of largest systematic uncertainties and implementation of a ranking procedure
  4. Data background-only fits to assess the robustness of background model







Thank you