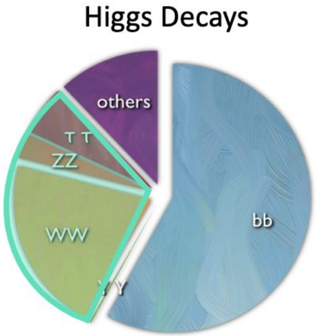
Top quark to Higgs boson Yukawa coupling measurement using the multilepton final states with the ATLAS detector at the LHC

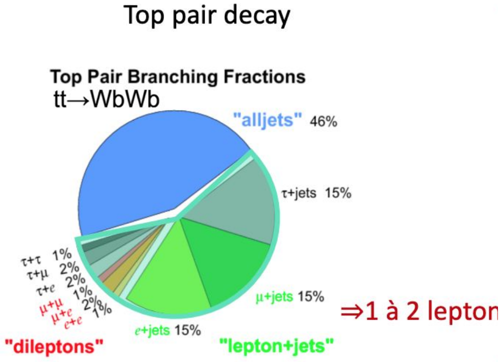
Speaker: Xuan Yang
On behalf of the ATLAS collaboration
Supervisors: Johann Collot/Cunfeng Feng/Marine Kuna

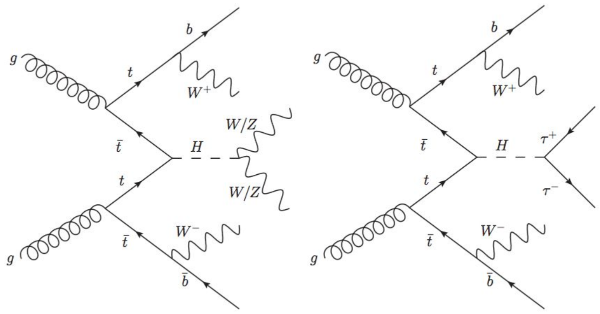
Higgs Hunting 2019
Paris/Orsay
2019.7.30

Motivation

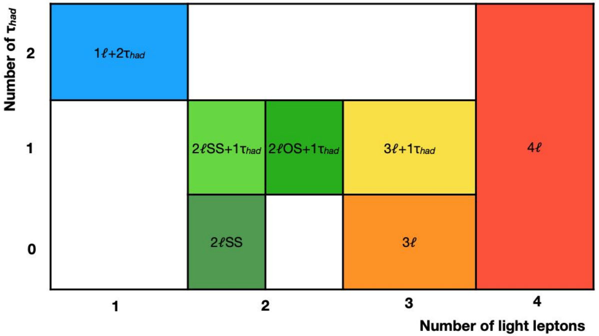
- Direct measurement of the Yukawa coupling between top quark and Higgs boson at tree level via ttH cross section measurement
- ttH prodution cross section at 13 TeV is ~1 % of total Higgs production cross section
- Multilepton signatures at decay modes with >= 1leptons (WW, ττ, ZZ) from both top quark and Higgs boson decays



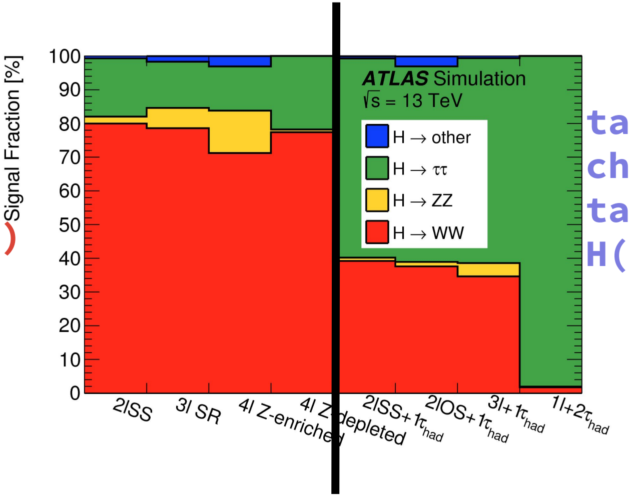




Overview



0-tau: target at H(WW)



tau channels: target at H(ττ)

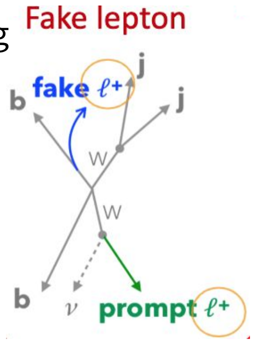
What we have in multilepton:

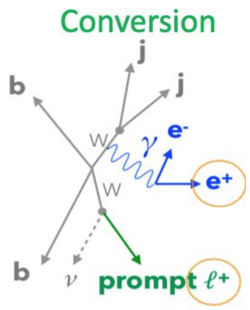
- 2/3/4 leptons (e/μ)
- 1 or 2 hadronic taus (τ_had)
- several jets (usually >=4)
- in which some are b-tagged (usually >=1)

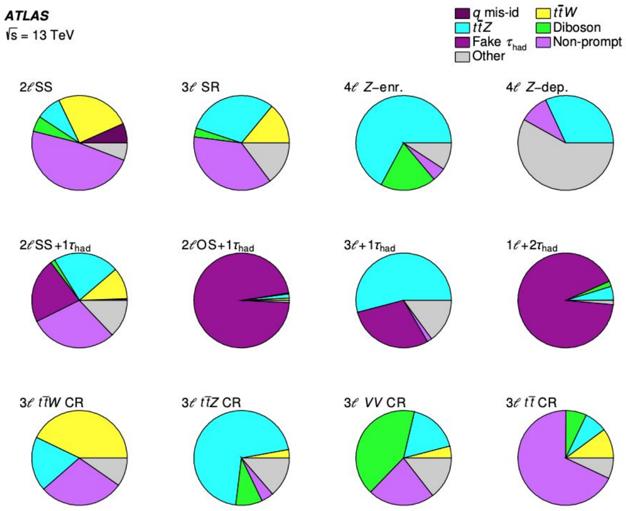
The selections among different sub-channels are **orthogonal at object level** to avoid overlaps and to allow combination. Each channel has its own requirement on the objects (i.e. tight lepton or loose lepton) to get maximum statistics and sensitivity.

Backgrounds

- Reducible backgrounds: can be better estimated by improving the estimate method
  - Non-prompt(fake) leptons
  - Fake taus
  - Charge mis-id (charge flip)







- Irreducible background: multilepton final state with prompt leptons possible. Prompt contribution is from simulation
  - ttW
  - ttZ
  - Diboson

In most of the channels, non-prompt and fake taus are the main backgrounds. It is essential to get a good estimate of these backgrounds.

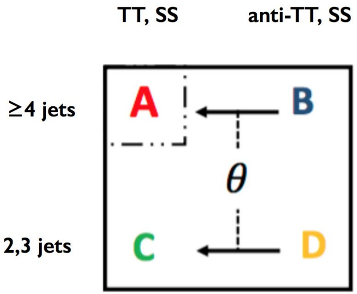
Non-prompt leptons (e/μ)

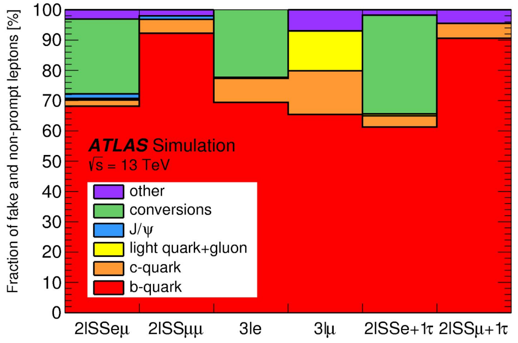
Non-prompt leptons are hard to estimate from simulation.

Therefore we developed two data-driven methods:

1. Fake factor: ABCD method, split by jet multiplicity and tight/anti-tight identified leptons. Fake factor is parametrized as a function of pT.

2. Matrix method: The idea is similar, estimating fake leptons from looser lepton ID region (looser than signal region). Elements in the matrix is the efficiency of a real/non-prompt lepton to pass tight/anti-tight ID, which are measured in data in a dedicated control region.

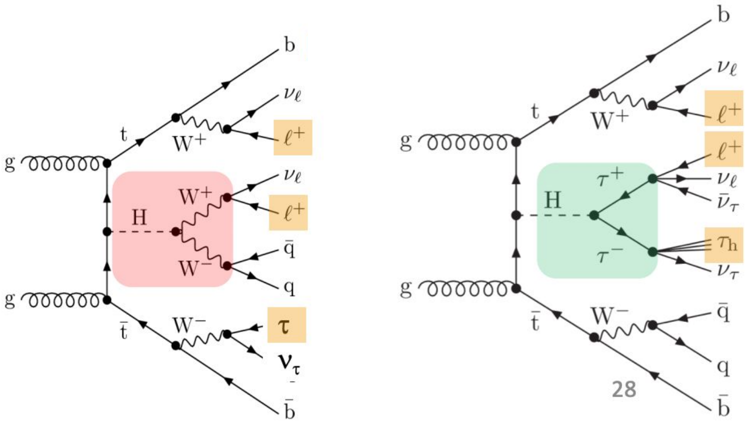


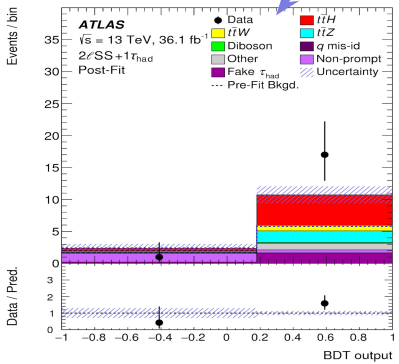


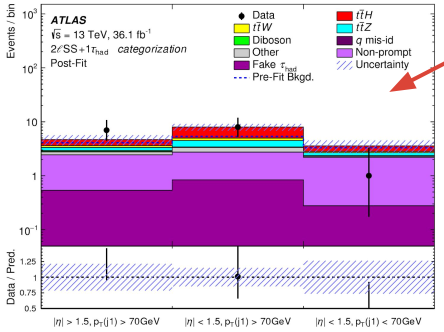
$$\begin{pmatrix} N^{TT} \\ N^{T\not T} \\ N^{\not T T} \\ N^{\not T\not T} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\not\varepsilon_{r,2} & \varepsilon_{r,1}\not\varepsilon_{f,2} & \varepsilon_{f,1}\not\varepsilon_{r,2} & \varepsilon_{f,1}\not\varepsilon_{f,2} \\ \not\varepsilon_{r,1}\varepsilon_{r,2} & \not\varepsilon_{r,1}\varepsilon_{f,2} & \not\varepsilon_{f,1}\varepsilon_{r,2} & \not\varepsilon_{f,1}\varepsilon_{f,2} \\ \not\varepsilon_{r,1}\not\varepsilon_{r,2} & \not\varepsilon_{r,1}\not\varepsilon_{f,2} & \not\varepsilon_{f,1}\not\varepsilon_{r,2} & \not\varepsilon_{f,1}\not\varepsilon_{f,2} \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}$$

2lss+1tau

- The most sensitive tau channel
- Backgrounds from both fake taus and fake leptons
- Statistically limited
- A Boosted Decision Tree (BDT) is trained to extract signal



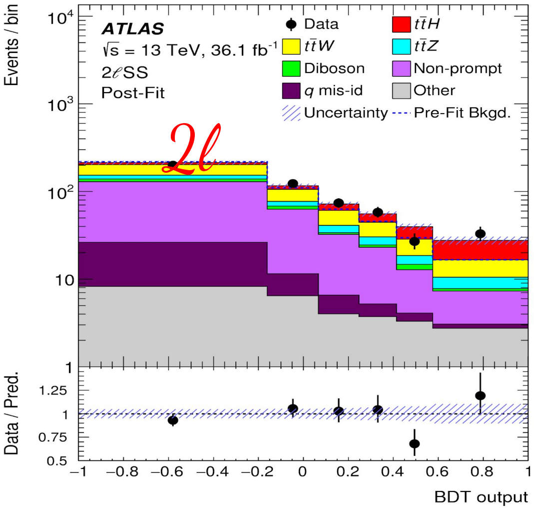




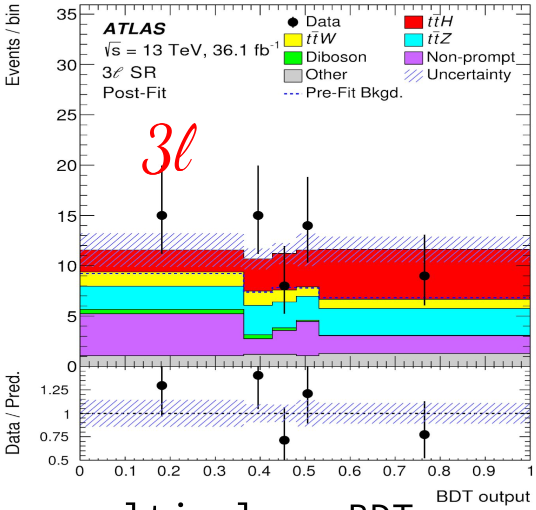
- Beside BDT, a cut-based analysis is also performed as a cross-check and alternative approach
- Categories set by cutting on two variables based on SR
  - Var1: maximum η of two leptons
  - Var2: leading jet pT
- It gives similar sensitivity to the BDT analysis

Signal extraction

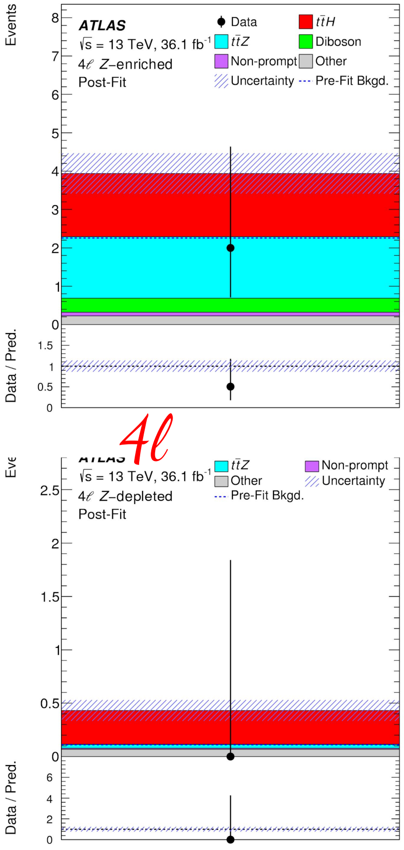
For non-tau channels



combination of signal vs. ttbar and signal vs. ttV BDT



multi-class BDT (ttH, ttW, ttZ, ttbar, diboson). Fit the signal output



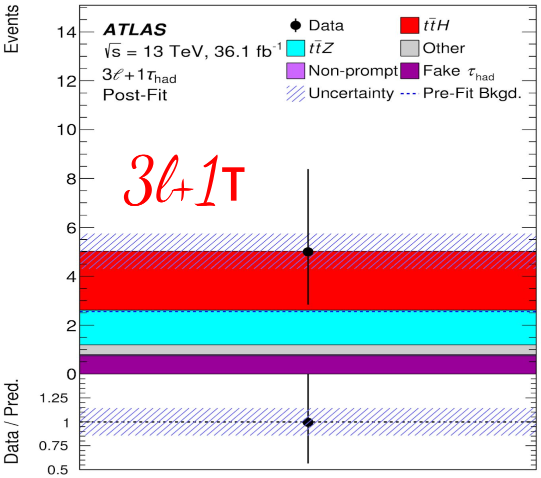
split into Z-enriched and z-depleted regions

These variables are used in the final simultaneous fit

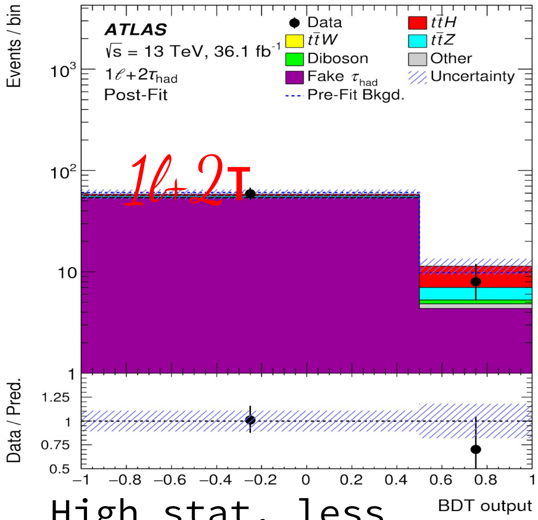
Signal extraction

For tau channels (except 2lss1tau)

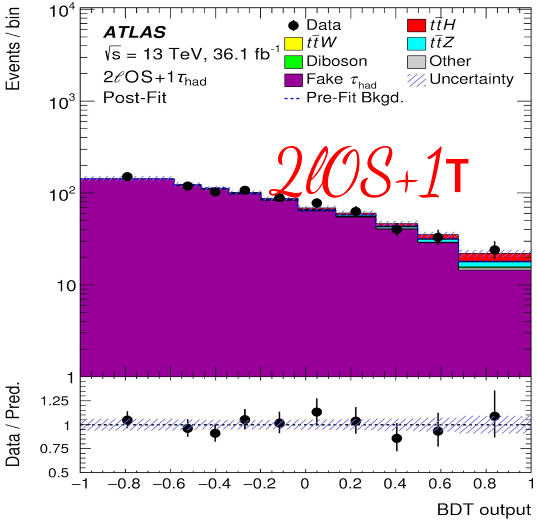
These variables are used in the final simultaneous fit



Very low stat, but very pure.
Cut-and-count



High stat, less pure.
BDT (signal vs ttbar). The lepton is 99% prompt, with 1 or 2 fake taus.

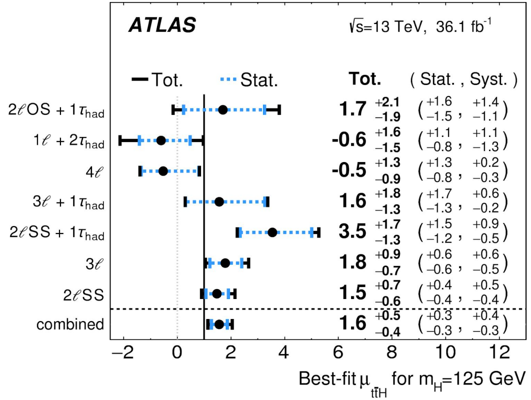


High stat, less pure.
BDT (signal vs ttbar). Dominated by fake taus.

Results

Observed: 4.1σ
Expected: 2.8σ

$\mu_{ttH}\,(\mathrm{ML}) = 1.56^{+0.49}_{-0.42}$



Evidence of ttH production with only ttH multilepton analysis using 36.1/fb data.

The results from CMS is 3.2σ observed and 2.8σ expected.

Combining with other ttH analysis (bb, ZZ, γγ), we observed 5.8σ and 4.9σ expected.

References

1. ATLAS results: https://arxiv.org/abs/1712.08891
2. CMS results: https://arxiv.org/abs/1803.05485
3. ATLAS combination results: https://arxiv.org/abs/1806.00425
4. CMS combination results: https://arxiv.org/abs/1804.02610

Thanks For Your Attention!

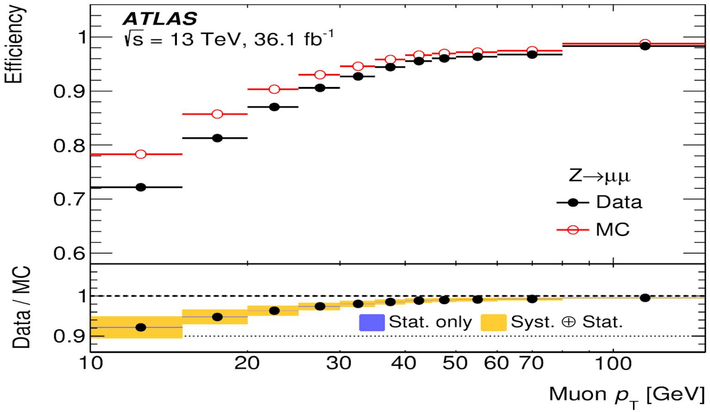
Non-prompt leptons (e/μ)

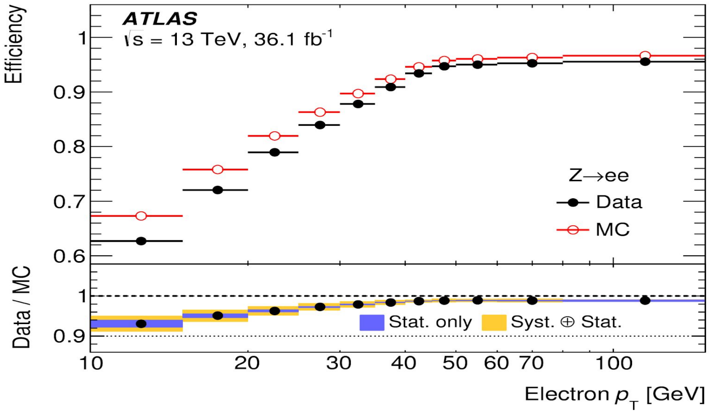
- The most important background: non-prompt leptons from semi-leptonic b decay
- Implement a new variable to reject non-prompt leptons -> PromptLeptonIso(PLI)

PLI is a BDT trained with:

- lepton and overlapping track jets properties
- lepton track/calorimeter isolation variables

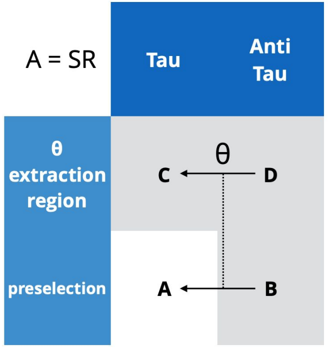
Scale factors (ratio of efficiency in data and in MC) to be used, measured from Z(ll) events. Maximum 0.95 at low pT.

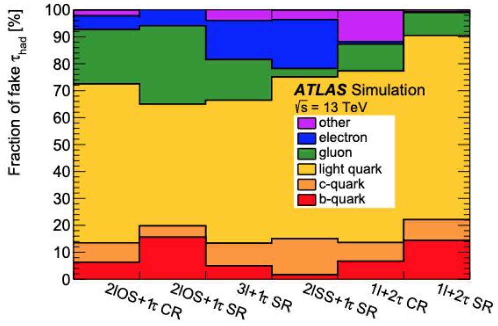




Non-prompt tau

- Fake tau is estimated from 2lOS+1tau control region, where is dominated by fake tau contribution.
- Fake factor method (ABCD method), similar to the one described for non-prompt leptons.
- Split by jet multiplicity and tau identification variable





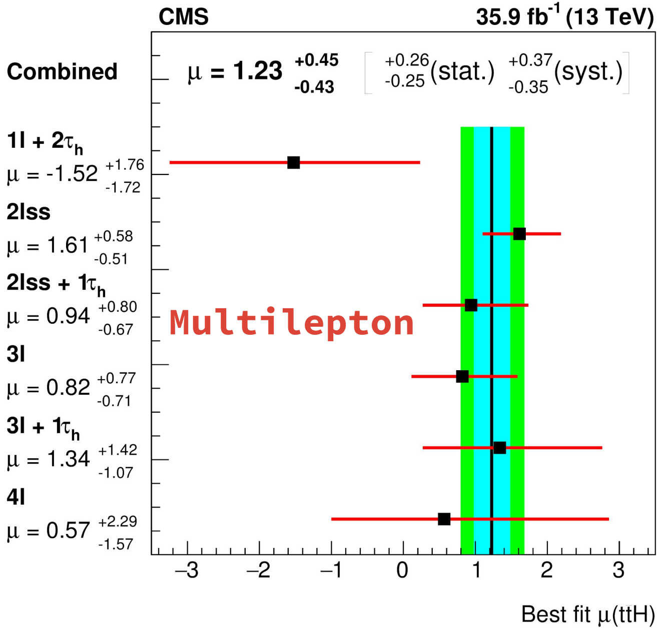
(b) Fake τ_{had} composition

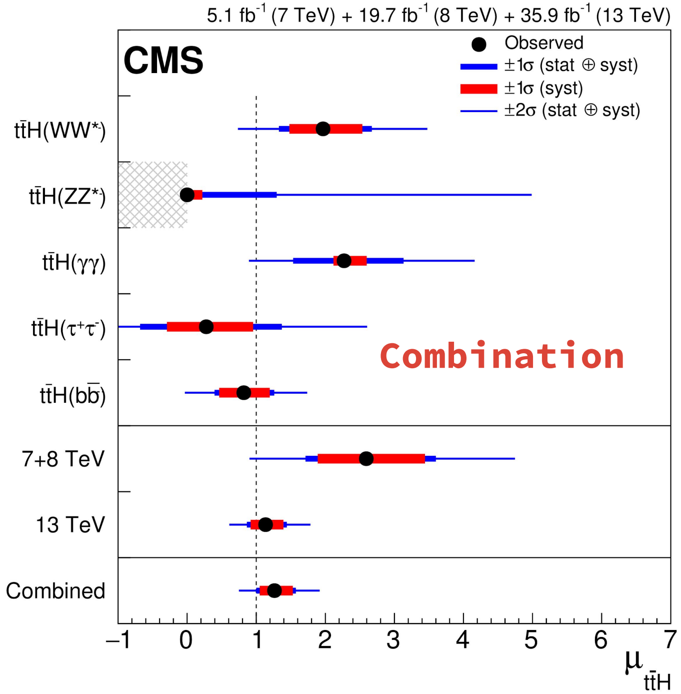
Fake tau composition is similar across channels, which allow us to just scale the factors measured in 2los1tau. A scale factor of 1.36 derived from 2los1tau CR (DD/Data) is applied to Monte carlos to get correct estimate in 3l1tau and 2lss1tau regions.

Fake tau in 1l2tau

- Fake taus in 1l2tau channel is measured from a 1l2tau CR with SS tau pair (OS in SR)
  - Jets have identical chance to be reconstructed as positve or negative charged tau
  - The estimation is taken from the SS data with small corrections from simulation samples (truth tau contribution)

CMS result





Observed: 3.2σ
Expected: 2.8σ

Observed: 5.2σ
Expected: 4.2σ