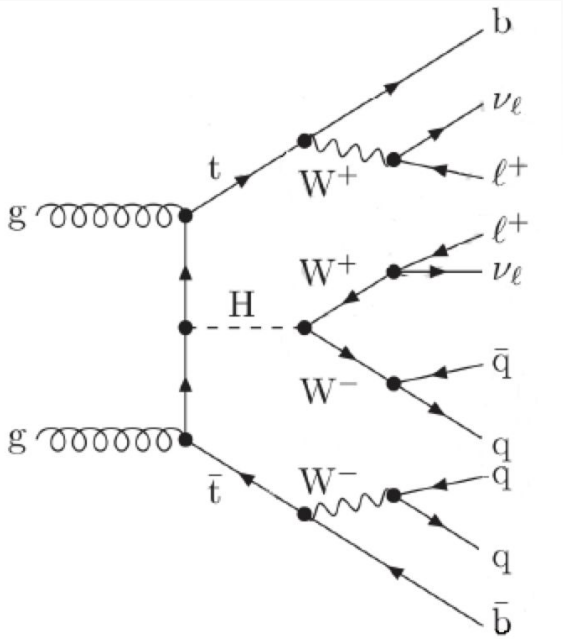
Joker talk: differential ttH production in the multilepton final state using Run 2 dataset

Angela Taliercio on behalf of the CMS collaboration

November 7 2024

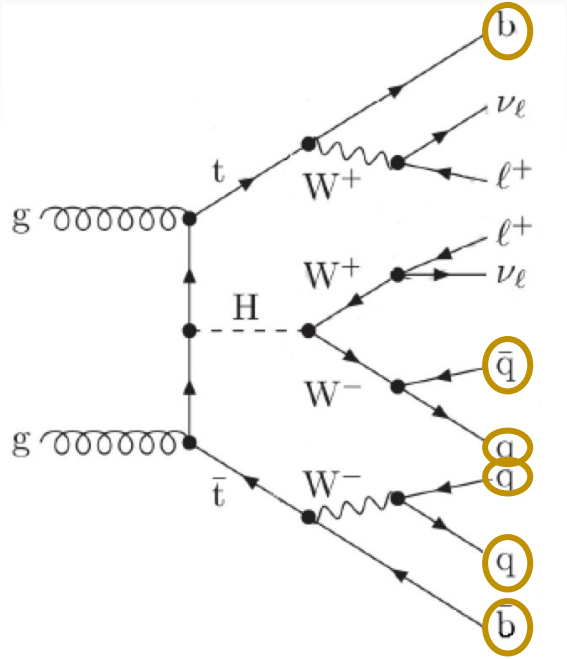
ttH multilepton: introduction

- Top Yukawa coupling
  - Top is heaviest quark. yt~1
  - Sensitive to BSM effects
- Target Higgs decay modes:
  - H→ WW → lvlv / lvqq
  - H→ ττ → hadrons/leptons+v's



ttH multilepton: introduction

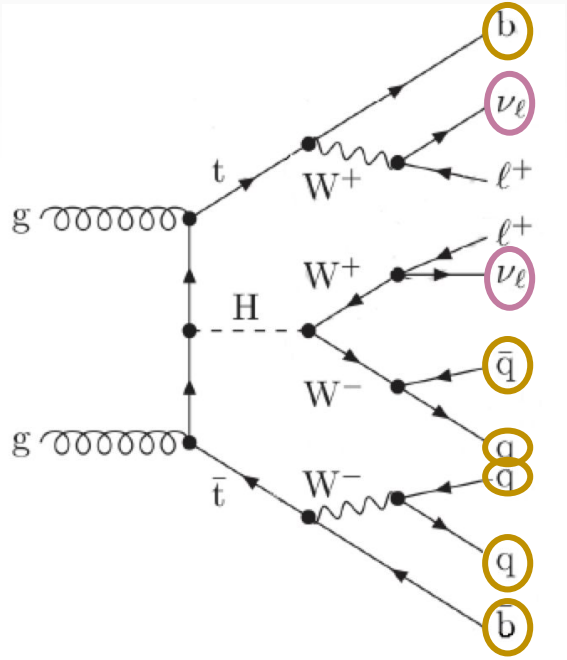
- Top Yukawa coupling
  - Top is heaviest quark. yt~1
  - Sensitive to BSM effects
- Target Higgs decay modes:
  - H→ WW → lvlv / lvqq
  - H→ ττ → hadrons/leptons+ν's



- Very crowded final state (lots of jets coming both from the Higgs and from the tops)

ttH multilepton: introduction

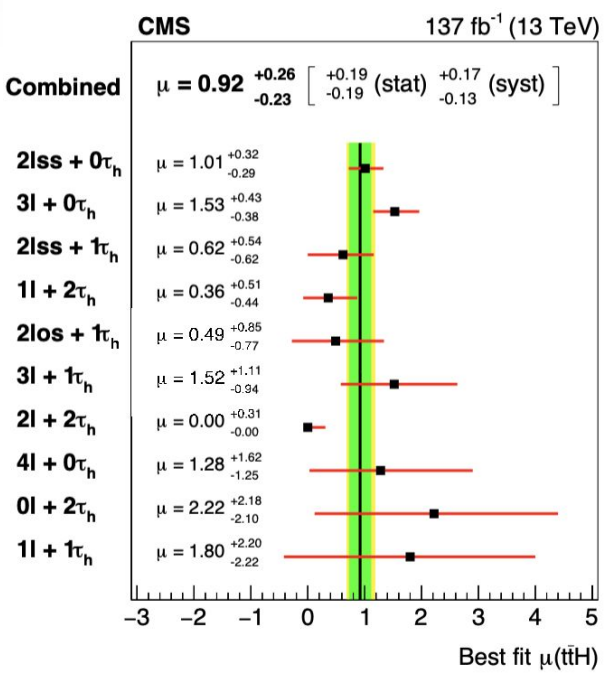
- Top Yukawa coupling
  - Top is heaviest quark. yt~1
  - Sensitive to BSM effects
- Target Higgs decay modes:
  - H→ WW → lvlv / lvqq
  - H→ ττ → hadrons/leptons+ν's



- Very crowded final state (lots of jets coming both from the Higgs and from the tops)
- Challenging final state for the presence of neutrions
  - Cannot reconstruct without ambiguity the full process

ttH multilepton: previous measurement

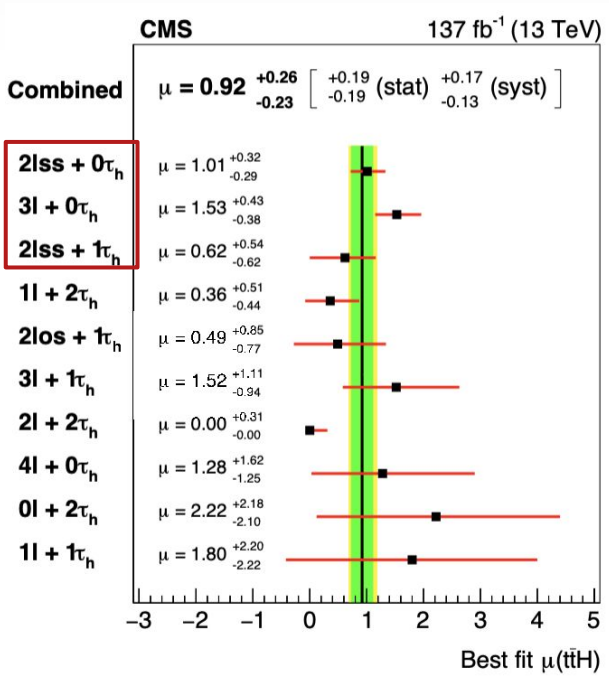
- Run 2 inclusive measurement (HIG-19-008):
  - Signal strength: 0.92 ± 0.19 (stat)+0.17-0.13 (syst)
    - Significance: 4.7σ (5.7σ exp)
  - Categorize events by number of leptons and hadronic taus
    - 3 most sensitive channels: 2lss, 3l, 2lss+1τ_h
- Run 2 CP Measurement (HIG-21-006)



Differential ttH

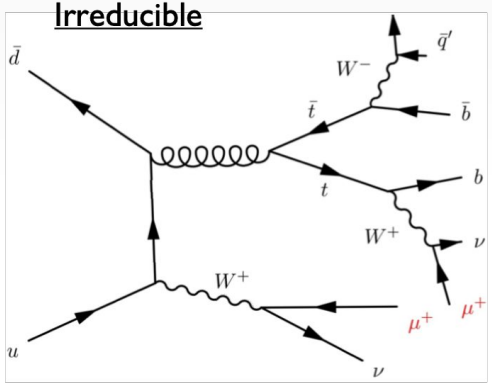
- Run 2 inclusive measurement (HIG-19-008):
  - Signal strength: 0.92 ± 0.19 (stat)+0.17-0.13 (syst)
    - Significance: 4.7σ (5.7σ exp)
  - Categorize events by number of leptons and hadronic taus
    - 3 most sensitive channels: 2lss, 3l, 2lss+1τ_h
- Run 2 CP Measurement (HIG-21-006)

- We select only the best 3 final states
- Differential measurement using full Run 2 data
  - Custom Higgs p_T bins
  - mttH bins



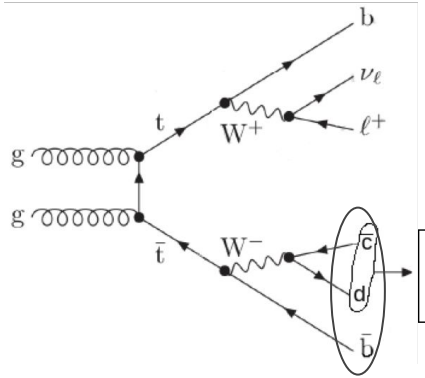
Major Backgrounds

Irreducible



- ttW, ttZ are dominant
- Modeled in MC, normalization left floating in fit
  - Control regions to constrain
  - Validation regions to validate modeling

Reducible

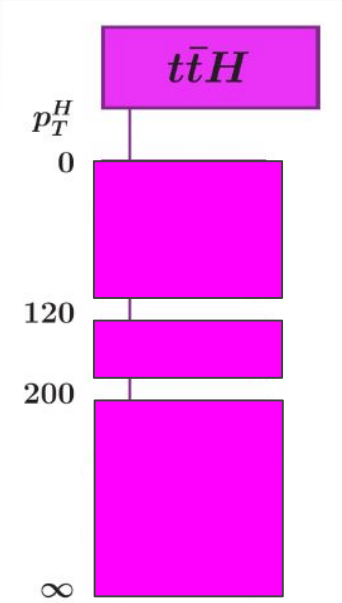


Fake or non-prompt l^+

- Non-prompt lepton or jet fakes a prompt lepton
- Mostly tt and W+jets in 2lss
- Estimated with data-driven methods

Analysis Strategy

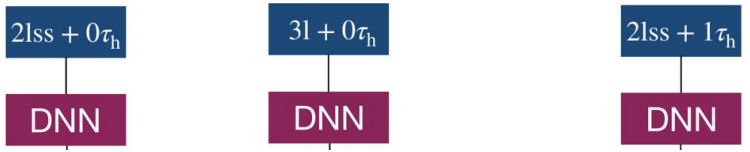
- Pick three best channels from inclusive ttH analysis to do differential measurement
  - Custom Higgs p_T binning for optimized sensitivity (3 bins)
  - Custom Higgs m_{ttH} binning (3 bins)
- Estimate the ttW, ttZ, WZ, ZZ background from control regions:
  - 3l and 4l control region
- DNN multiclassifier, to maximise the sensitivity and the background modelling
  - Higgs p_T regression with DNN
- Fit all simultaneously all the classes of the DNN and the control regions



custom binning

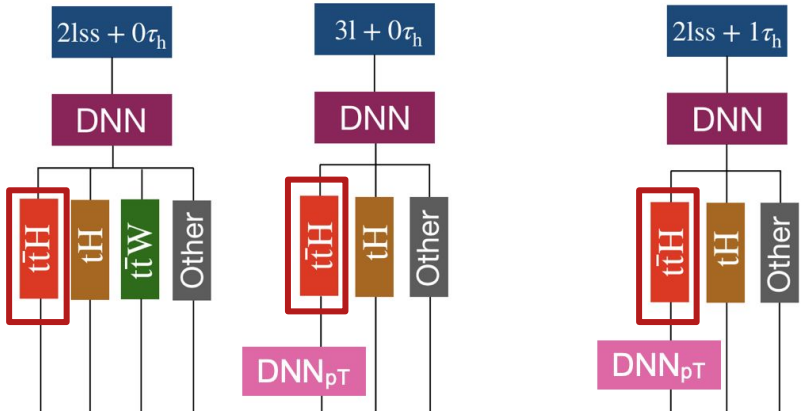
Event Categorization & Fit Strategy

- In each channel, train a DNN Classifier to distinguish events between ttH low Higgs p_T, ttH high Higgs p_T, tH and background



Event Categorization & Fit Strategy

- In each channel, train a DNN Classifier to distinguish events between ttH low Higgs p_T, ttH high Higgs p_T, tH and background

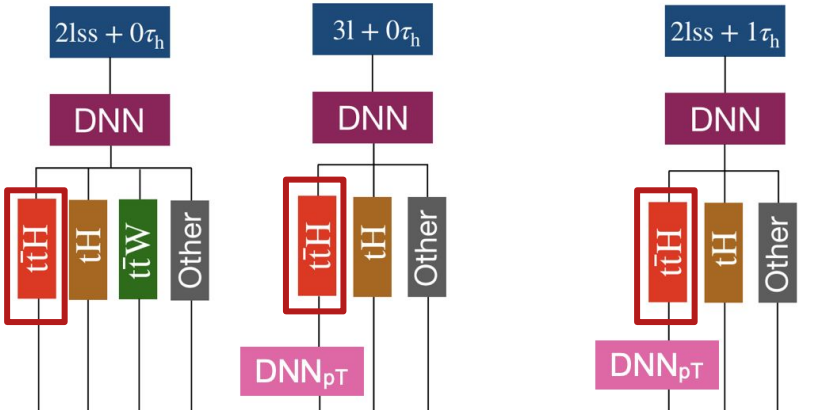


- For ttH like events
  - Train DNN for Higgs p_T regression
- For tH, ttW and other backgrounds
  - Split by lepton flavor, number of b jets when statistics allow

: DNN Classifier outputs two ttH scores: H p_T < 300, H p_T > 300 Ge

DNN Classifier

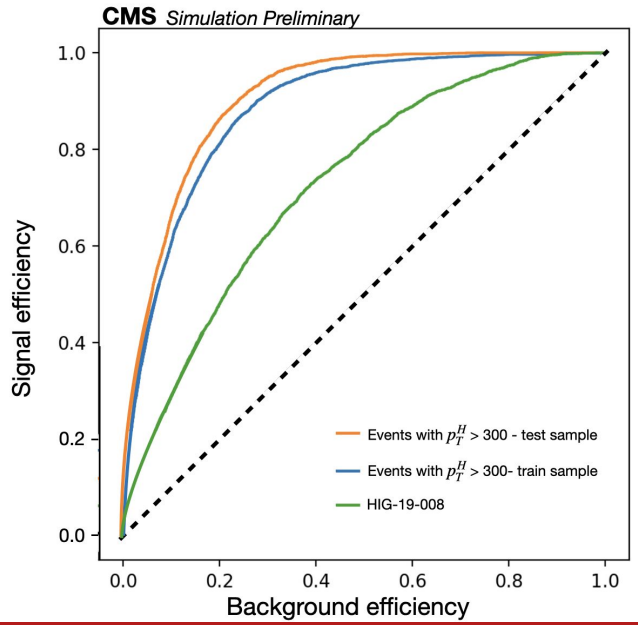
- Improved upon inclusive analysis by training separately high and low Higgs pT events
  - BSM effects more sensitive at higher Higgs pT



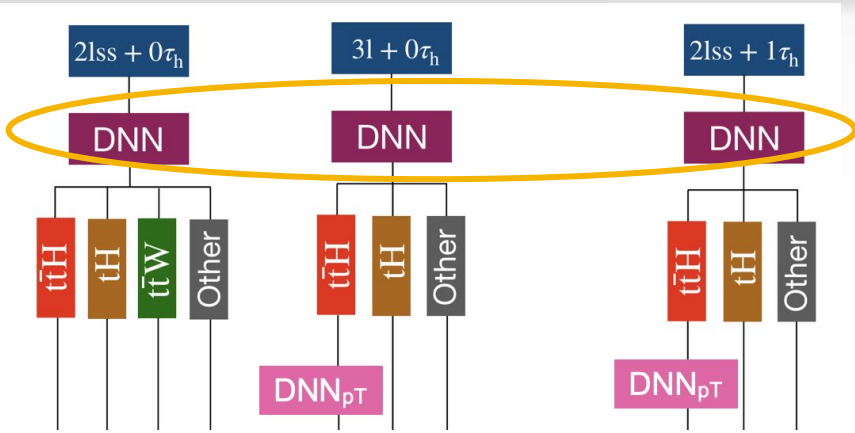
: DNN Classifier outputs two ttH scores: H p_T < 300, H p_T > 300 (

DNN Classifier Performance

- New classifier perform significantly better at separating higher Higgs p_T & tHq signal from others

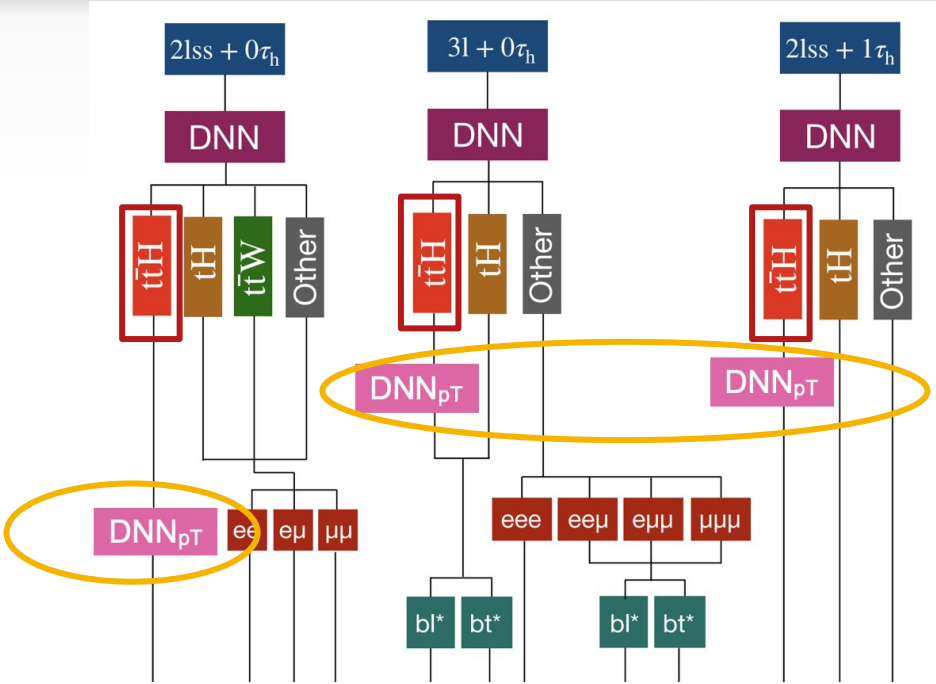


2lss



Higgs p_T Regression DNN

- Higgs p_T Regression DNN
  - New DNN developed for differential analysis



: DNN Classifier outputs two ttH scores: H p_T < 300, H p_T > 300

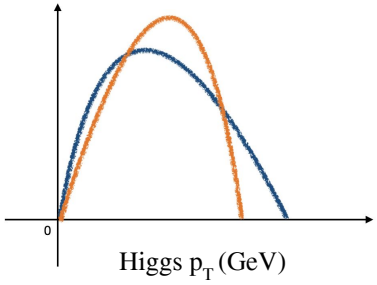
DNN Loss Function

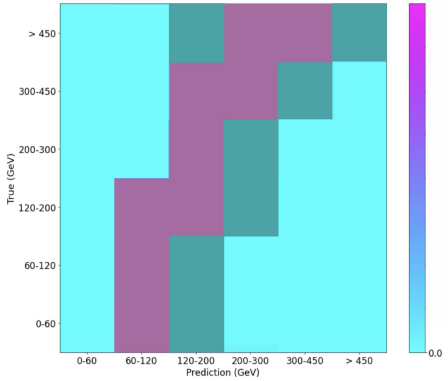
- Loss Function: Quantity that DNN seeks to minimize during training

  Common Choice: Mean-Squared-Error

$$L = \frac{1}{N} \sum \left((y_t - y_p)^2 \right)$$

- When using standard loss functions (like MSE), predictions cluster around center of true distribution.





DNN Loss Function

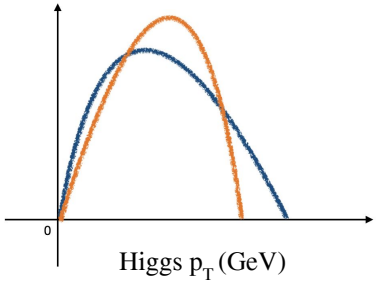
- Loss Function: Quantity that DNN seeks to minimize during training

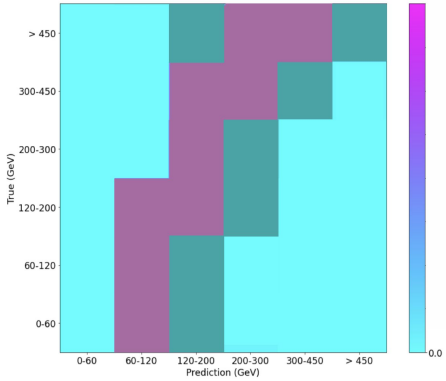
  Common Choice: Mean-Squared-Error

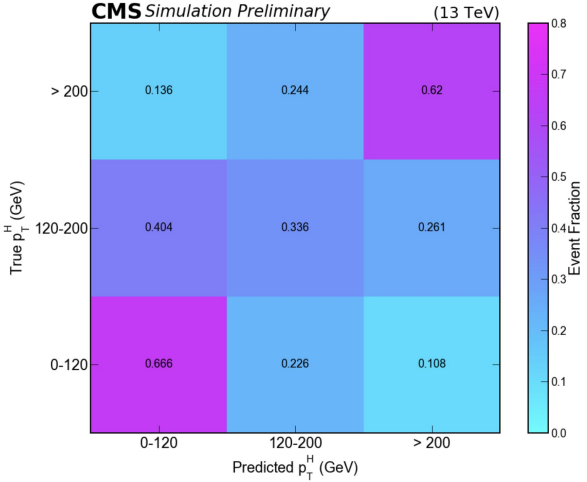
$$L = \frac{1}{N}\sum\left((y_t - y_p)^2\right) \longrightarrow L = \frac{1}{N}\sum\left((y_t - y_p)^2\right) \times \left|\left(\sigma_t^2 - \sigma_p^2\right)\right|$$

Term to penalize deviations in variance between true & prediction

- When using standard loss functions (like MSE), predictions cluster around center of true distribution.



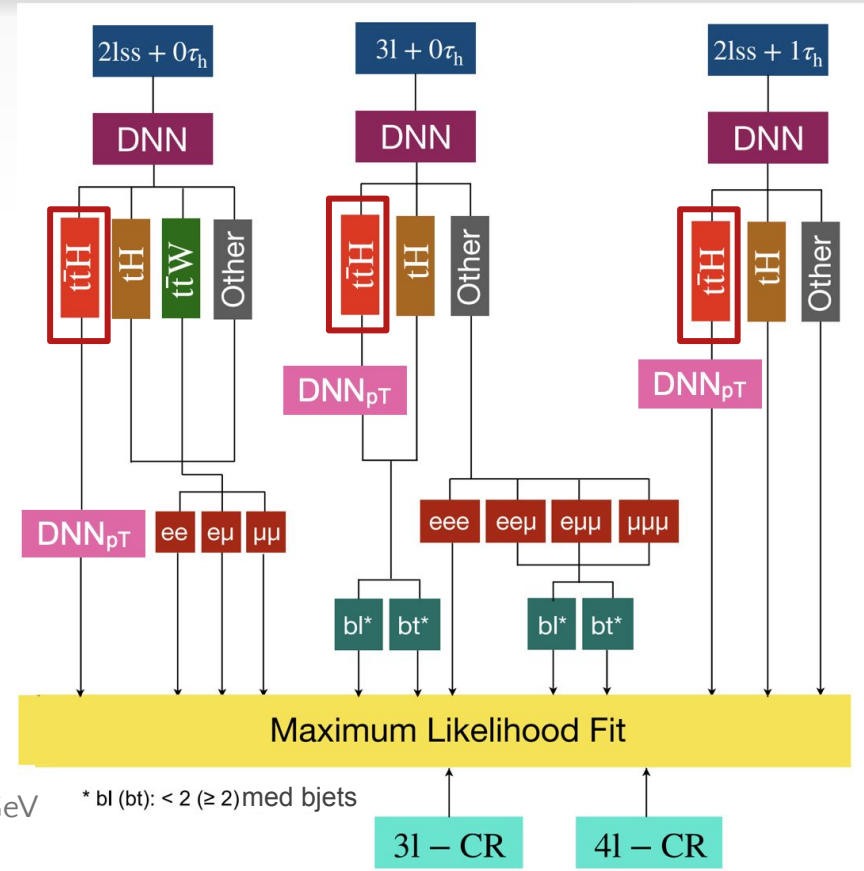




Event Categorization & Fit Strategy

- For optimized sensitivity, in each channel, train a DNN Classifier to distinguish events between ttH low Higgs p_T, ttH high Higgs p_T, tH and background

- Maximum-likelihood (ML) fit to the yields in the signal region bins for 2lss, 2lss + 1τh and 3l & 3l and 4l control regions

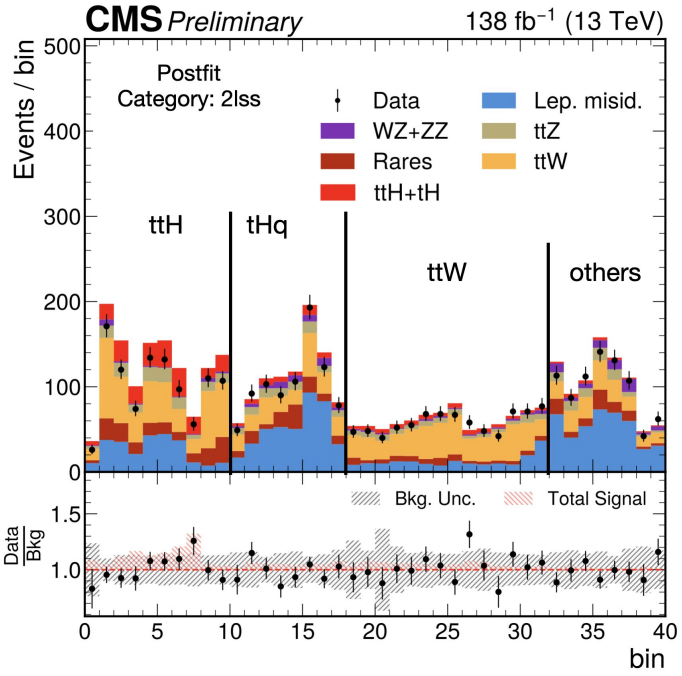


: DNN Classifier outputs two ttH scores: H p_T < 300, H p_T > 300 GeV

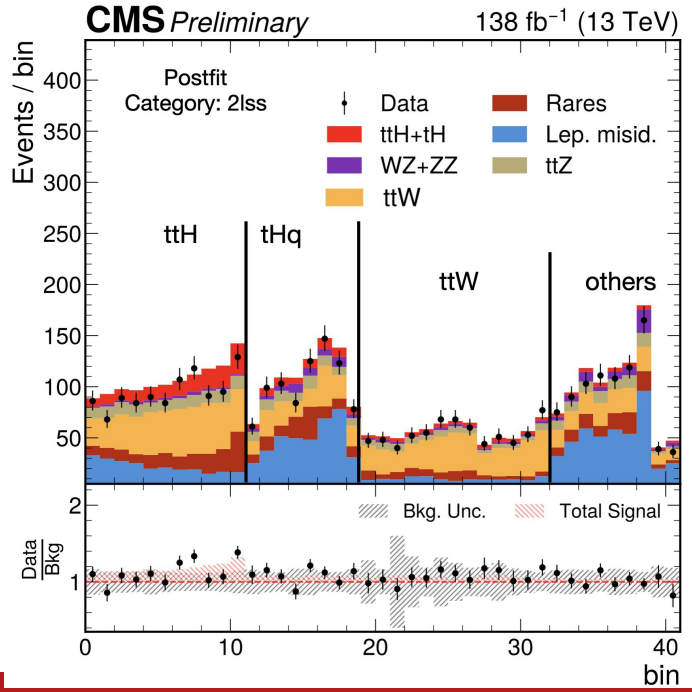
Postfit distributions

- Postfit distribution for the 2lss category for:

m_{ttH} category



Higgs p_T category

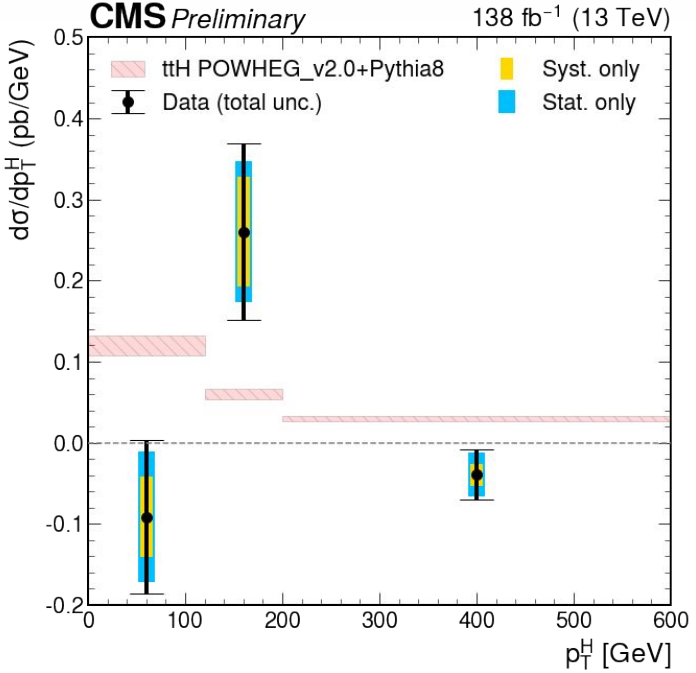


Higgs p_T Results

- We show the measurement of the cross section in
  - Custom Higgs p_T binning

class	$2\ell SS + 0\tau_h$	$2\ell SS + 1\tau_h$	$3\ell + 0\tau_h$
[] $t\bar{t}H$ _low	2×4	2×4	(< 2 bjet med., ≥ 2 bjet med.) 2×4
$t\bar{t}H$ _high	1×4	1×4	1×4
tH	8	4	(< 2 bjet med., ≥ 2 bjet med.) 2
bkg	9	5	(eee) 1 (eem: < 2 bjet med., ≥ 2 bjet med.) 2 (emm: < 2 bjet med., ≥ 2 bjet med.) 2 (mmm: < 2 bjet med., ≥ 2 bjet med.) 2
$t\bar{t}W$	13	—	—
Total	42	21	21

custom Higgs p_T binning

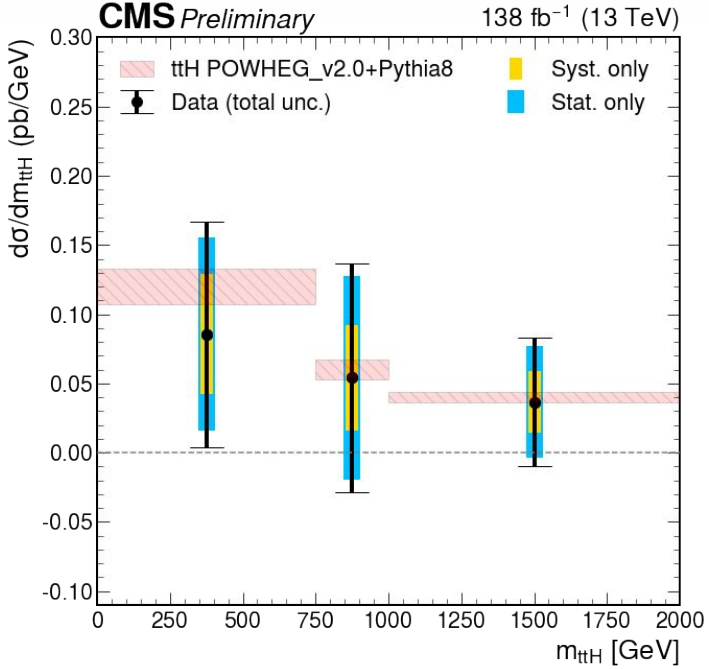


Higgs m_{ttH}

- We show the measurement of the cross section in
  - Custom m_{ttH} binning

custom mttH

class	$2\ell SS + 0\tau_h$	$2\ell SS + 1\tau_h$	$3\ell + 0\tau_h$
$t\bar{t}H$	6×3	4×3	(< 2 bjet med.) 2×3, (≥ 2 bjet med.) 2×3
tH	8	4	(< 2 bjet med., ≥ 2 bjet med.) 2
bkg	9	5	(eee) 1 (eem: < 2 bjet med., ≥ 2 bjet med.) 2 (emm: < 2 bjet med., ≥ 2 bjet med.) 2 (mmm: < 2 bjet med., ≥ 2 bjet med.) 2
$t\bar{t}W$	13	—	—
Total	42	21	21



Summary

- **Fist ttH differential measurement in ttH multilepton final state**
- Very challenging final state large → large multiplicity of jets, presence of neutrinos

- Multiclass DNN to categorize the events
  - DNN to regress the Higgs p_T

- We measured the differential cross section as a function of:
  - 3 Higgs p_T: [0,120), [120, 200), [200, inf) GeV
  - 3 m_{ttH} bins: [0,750), [750, 1000), [1000, inf) GeV

- Results are in agreement with the SM prediction