Improvements in the measurement of VBF and ttH production with H→$\tau\tau$ in ATLAS

Topical Talks - Higgs Hunting 2024

Enrique Valiente Moreno (IFIC, CSIC-UV), on behalf of the ATLAS Collaboration

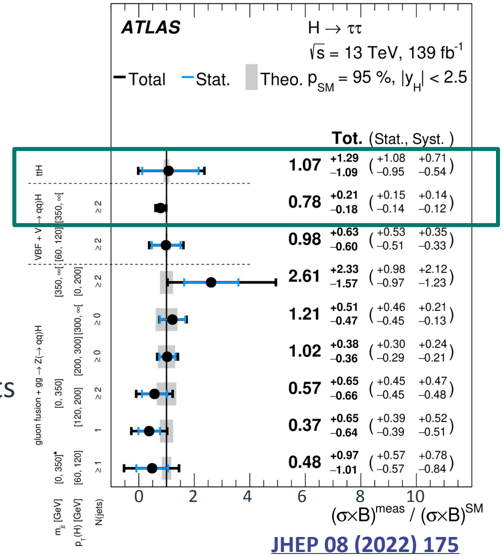
23/09/2024

Introduction - H→$\tau\tau$ analysis

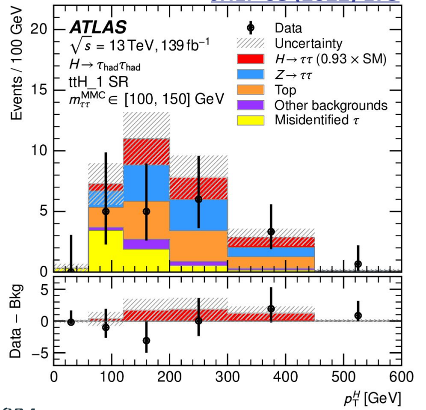
- Largest leptonic branching ratio of 6.3%. Unique opportunity to study Yukawa mechanism
- Reduced SM backgrounds requiring final states with hadronic or leptonic τ-lepton decays
  - **Higgs boson mass peak** and **p_T(H)** easily built from di-τ system
- High purity final states + sizable branching ratio give a **powerful channel for measuring Higgs boson production (ggF, VBF, VH and ttH)**

Previous H→$\tau\tau$ Run-2 results (JHEP 08 (2022) 175)

- 9 Parameter of Interest (PoI) measurement using Simplified template cross-sections **(STXS)**
  - Phase space partitions using kinematic properties of the Higgs boson and associated objects
- Inclusive measurement for **ttH** and **VBF**
  - H→$\tau\tau$ provided most precise **VBF** production measurement (Nature 607 (2022) 52)
  - **ttH** limited statistically, but sensitivity for high p_T(H) bins



JHEP 08 (2022) 175

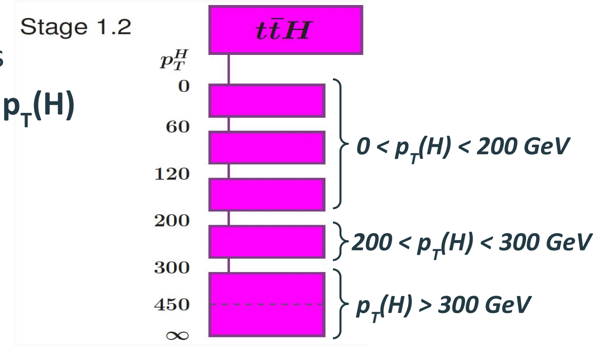


New H→$\tau\tau$ Run-2 legacy analysis (HIGG-2022-07)

- Based on previous round and focused on VBF and ttH
- Improvements via new MVA techniques and strategy for statistical fit
- **First unfolded fiducial differential cross-section measurement of H→$\tau\tau$** in a VBF enhanced phase space

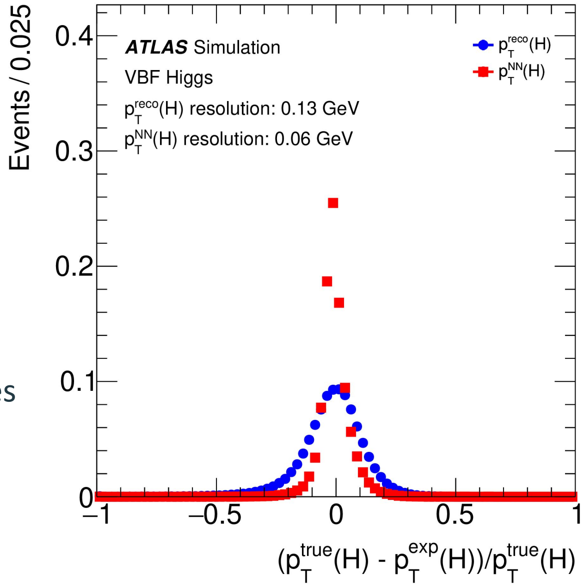
$H \rightarrow \tau\tau$ *legacy* analysis - Improvements in ttH

- Higgs boson coupling to heaviest fermions ⟶ Highly sensitive process to new physics
- Due to limited data statistics, ttH measurement **only extended to three STXS bins in p_T(H)**



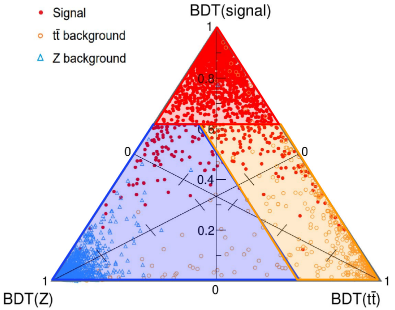
- **Previously**: two BDTs trained to separate ttH from main backgrounds, **Z($\tau\tau$) and tt**
  - Signal and Control regions (SR, CR) defined combining both scores
  - BDT trainings inclusive in **p_T(H)**

- **NEW p_T(H)** reconstruction via **Neural Networks (NN)**
  - Reduces event migration between STXS bins and improves resolution compared to the **previous method**
  - **Input variables for NN:**
    - **$\Delta R_{\tau\tau}$, $\Delta\phi_{\tau\tau}$**: angular distances between the two τ-leptons
    - **E_T^{miss}** : missing transverse energy
    - **$p_T^{\tau\tau}$**: built from the four-momenta of the two τ-leptons and **E_T^{miss}**
    - **$m_{\tau\tau}^{coll}$** : di-τ invariant mass in the collinear approximation
  - Trained using ggF events, performance checked for other production modes

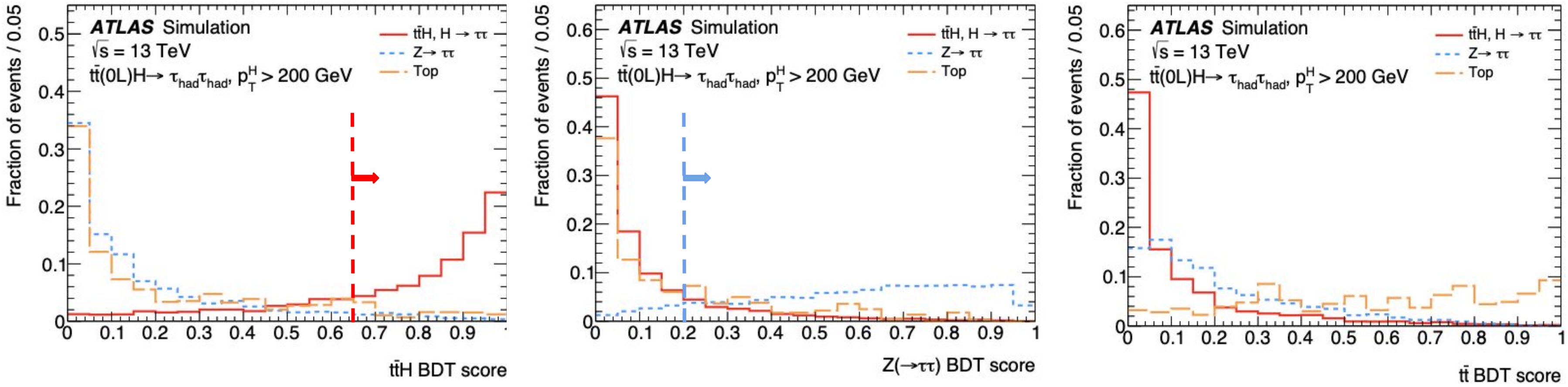


H→ττ *legacy* analysis - Improvements in ttH

- **Legacy analysis:** Multiclass classifiers trained using simultaneously the three processes **(ttH, Z(ττ) and tt)**
  - SRs and CRs defined combining **the three different scores** produced
  - **Two trainings** performed for **p_T(H) < 200 GeV and p_T(H) > 200 GeV**
    - The relative contribution from main backgrounds varies as a function of p_T(H)
    - Taking advantage of **new p_T(H) reconstruction method**

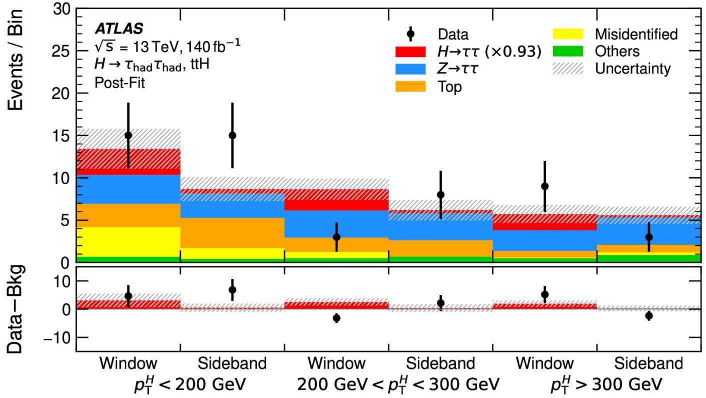


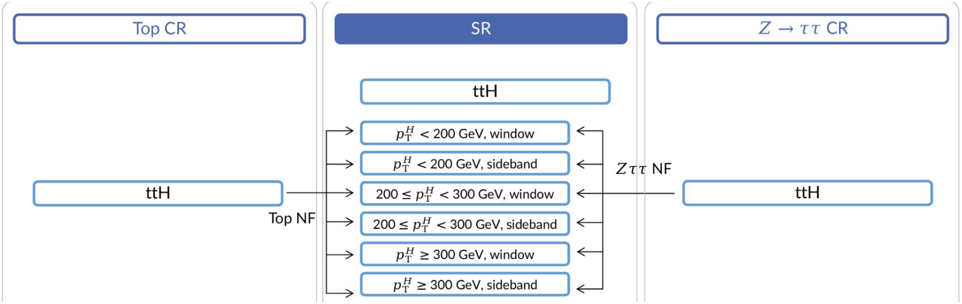
- **MC distributions of BDT scores** for **p_T(H) > 200 GeV**

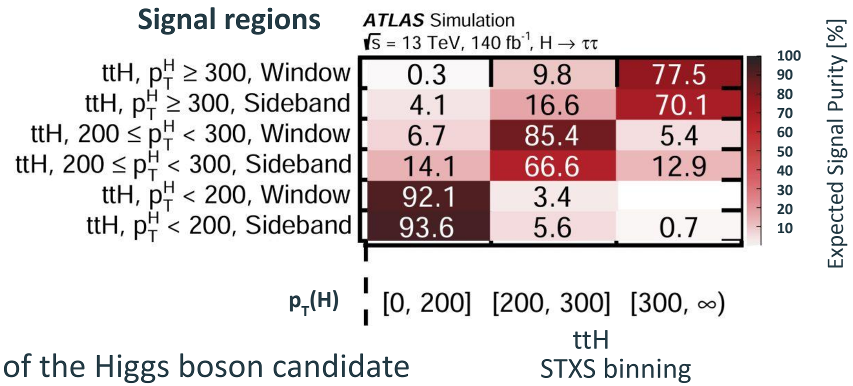


H→ττ *legacy* analysis - Improvements in ttH

- 6 SRs defined, based on the multiclass scores and the **di-τ system invariant mass** (using the Missing Mass Calculator, MMC)
  - "Window" regions with MMC around the SM Higgs mass and "Sideband"
- Inclusive CRs also defined for **Z(ττ) and tt** in order to constraint the main background sources



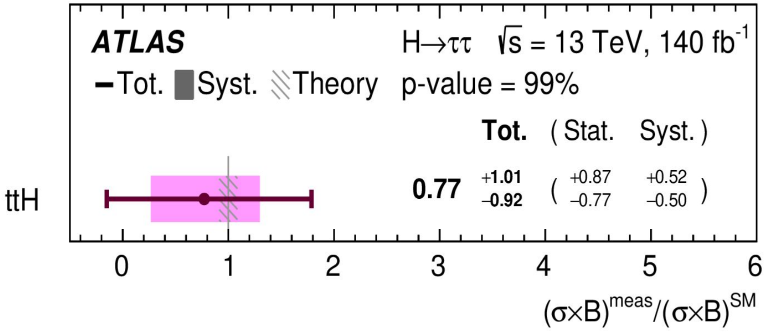




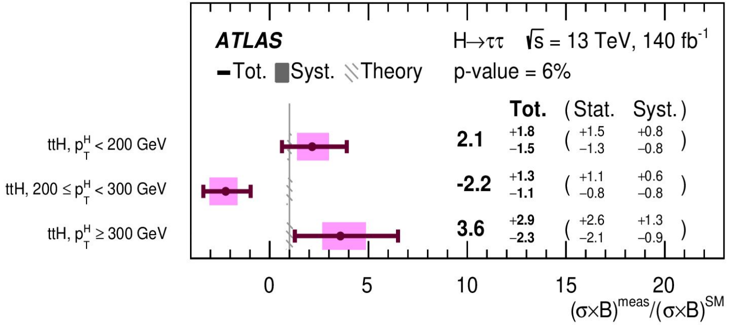
- Highly pure SRs obtained in ttH(ττ) events for each STXS bin in p_T of the Higgs boson candidate
- **Improved by 25% expected** sensitivity in the PoI $\mu = (\sigma \times \mathrm{BR}_{\tau\tau})/(\sigma \times \mathrm{BR}_{\tau\tau})_{SM}$ of ($1^{+0.92}_{-0.79}$) compared to the previous result ($1^{+1.24}_{-1.06}$)

$H\rightarrow\tau\tau$ *legacy* analysis - Results for ttH

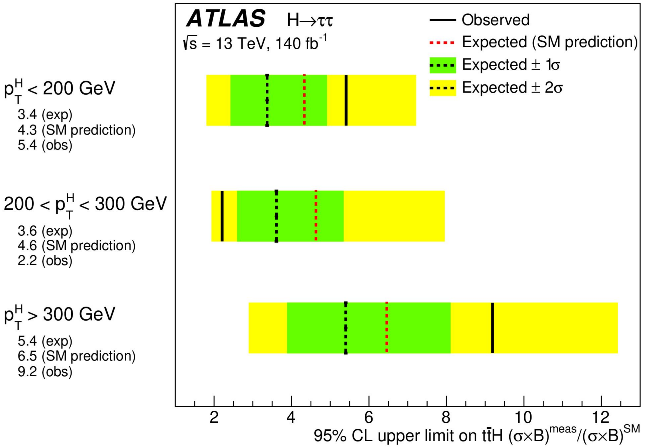
ttH measurement from 4-PoI fit



ttH measurement from 18-PoI fit

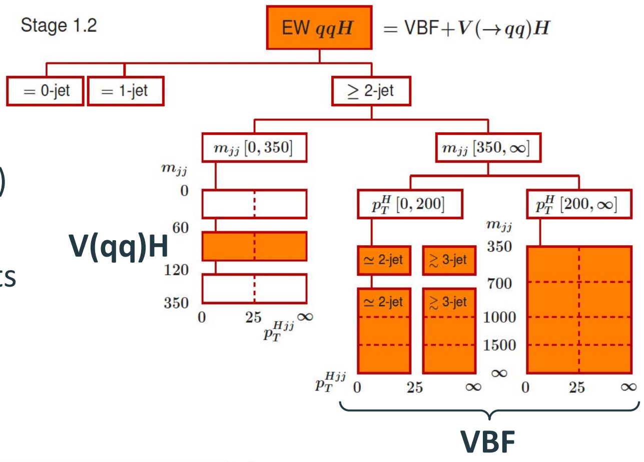


- Binned maximum likelihood fits performed using MMC as discriminant to measure the PoIs
- **4-PoI fit:** dedicated PoI for each production mode
- **18-PoI fit**: no significant deviations from the SM in ttH STXS bins
  - Limited sensitivity obtained in the fit due to poor statistics.
  - **Upper exclusion limits at 95% CL** were computed:
    - Expected (μ=0): ranging between **~3-5xSM prediction**
    - Expected injecting μ=1: **~4-6xSM prediction**
    - Observed: **~2-9xSM prediction**



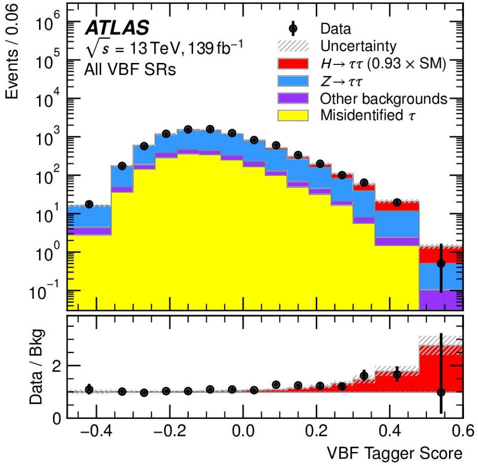
H→ττ *legacy* analysis - Improvements in VBF

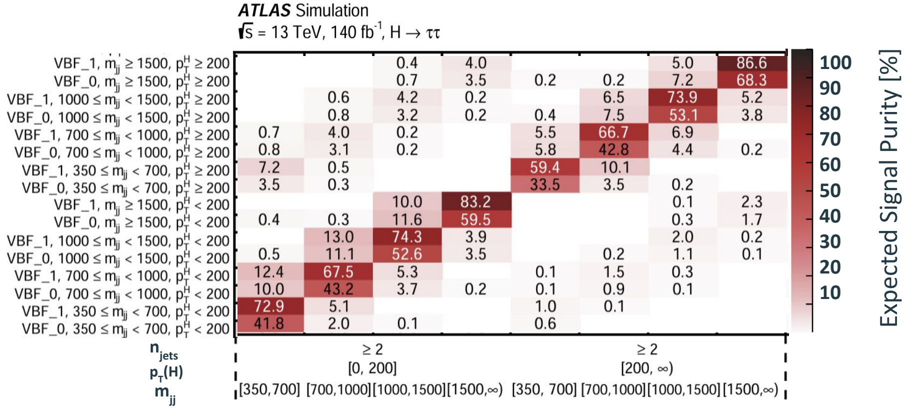
- Aim to extend the VBF measurement to more refined regions of phase space through new STXS bins
  - From inclusive measurement to 8 STXS bin measurement in p_T(H) and m_{jj} (invariant mass of di-jet system associated to Higgs boson production)
- **Event selection** strategy: **same BDT** from previous round (VBF vs ggF and Z(ττ)+jets) used to define VBF_1 (signal enriched) and VBF_0 SRs, optimizing cuts to enhance VBF sensitivity in each STXS bin



BDT score in all VBF SRs merged

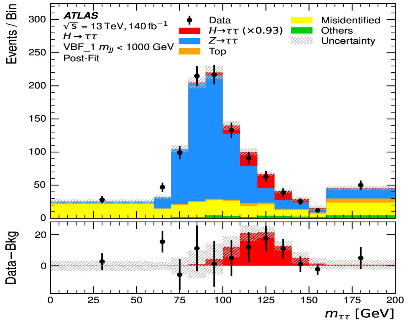
JHEP 08 (2022) 175

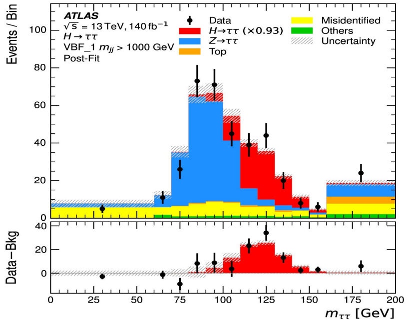




STXS binning

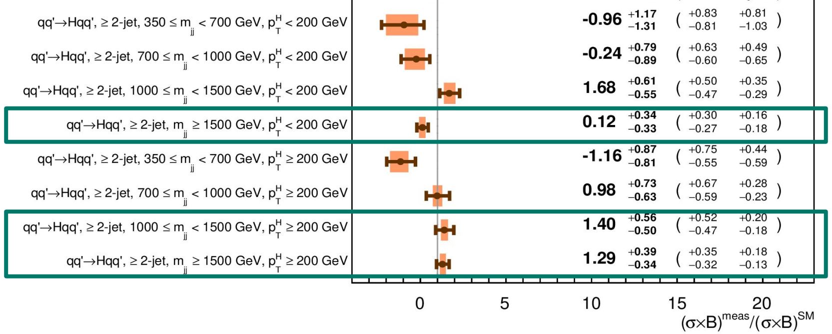
H→ττ *legacy* analysis - VBF results



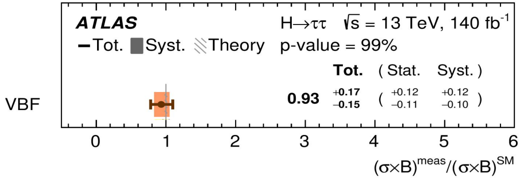


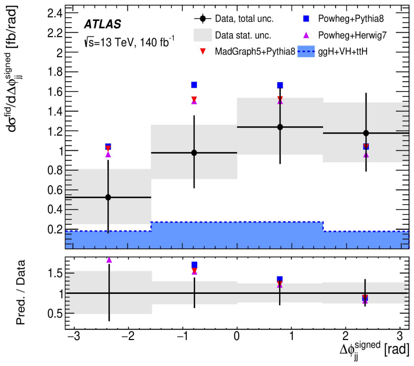
- No significant deviations from the SM
- High precision results in the high-$p_T(H)$ and/or high-m_{jj} regions due to the reduced SM backgrounds
  - First measurement for the higher-$p_T(H)$, and the most precise one in the lower-$p_T(H)$ region
- **Differential cross-section was measured for the first time in H→ττ in a fiducial phase space optimised for VBF production**
- Measured in bins of 4 variables sensitive to VBF kinematics: $\Delta\phi_{jj},\ p_T^{j_0},\ p_T^H,\ \Delta\phi_{jj}\ \text{vs}\ p_T^H$
- Good agreement between measured cross-sections and SM expectations

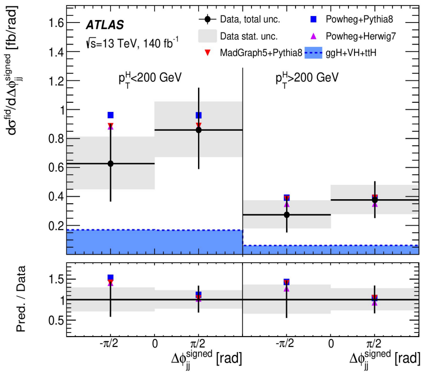
VBF measurements from 18-PoI fit



VBF measurement from 4-PoI fit







Summary

- Presented latest improvements in the measurements of VBF and ttH production processes in the $H \to \tau\tau$ channel in the ATLAS experiment
- STXS: Improvements mostly driven by new MVA techniques and strategy for statistical fit
  - Improved resolution in reconstructed $p_T(H)$ thanks to new developed NN-based approach
  - for ttH process, using a multiclass identifier compared to the previous binomial BDT
  - for VBF optimizing cuts on MVA score for each of the STXS bins
- Results mainly limited by size of analyzed data sample
  - ttH measurement improved by 25% with respect to previous analysis $\mu_{t\bar{t}H} = 1.06^{+1.28}_{-1.08}$
  - Reached most precise measurement of VBF Higgs boson production $\mu_{VBF} = 0.98^{+0.17}_{-0.15}$
  - Higher precision achieved for VBF in higher $p_T(H)$ and higher m_{jj} regions
- Obtained first unfolded differential cross-section measurements for VBF in $H \to \tau\tau$

Acknowledgements

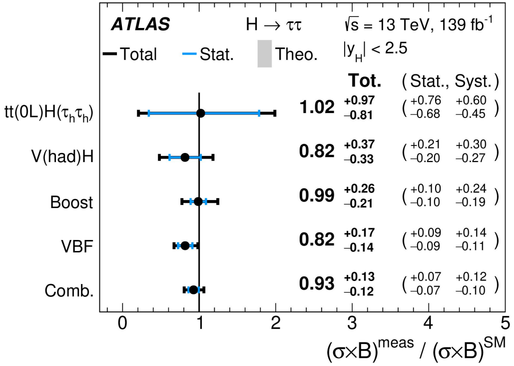
- The author's work is supported by:
    - FPU2021 grant funded by MICIU/AEI /10.13039/501100011033 and by the FSE+
    - Grants PID2021-124912NB-I00 andPID2021-125069OB-100 funded by MCIN/AEI/10.13039/501100011033
    - Project ASFAE/2022/010 funded by MCIN, by the European Union NextGenerationEU (PRTR-C17.I01) and Generalitat Valenciana

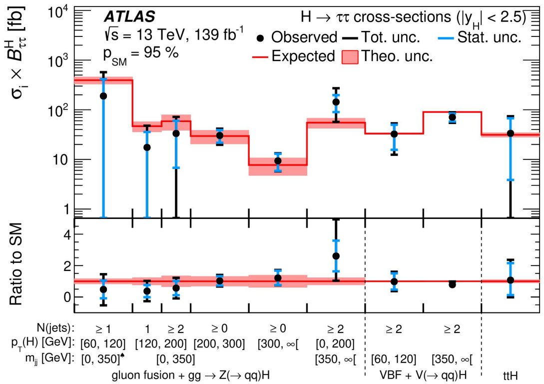
Additional material

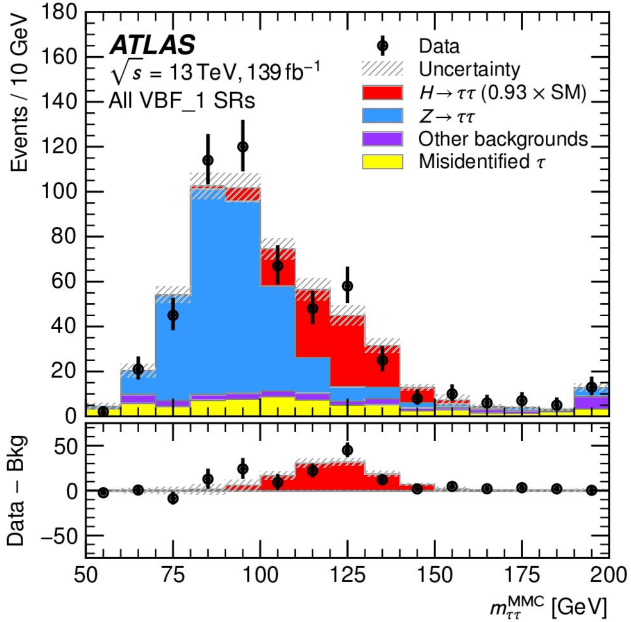
Previous analysis - Highlights

- Measured global μ for **pp→H→$\tau\tau$**, by production process (ggF, VBF, VH and ttH) and in 9 STXS bins (according to sensitivity)
  - ggF separated in p_T^H (expected better purity at high p_T^H)
  - VBF, VH and ttH are split using BDTs
- Di-tau system mass distribution considered in the statistical fit
- Achieved a relative uncertainty on global μ of 14%

JHEP 08 (2022) 175





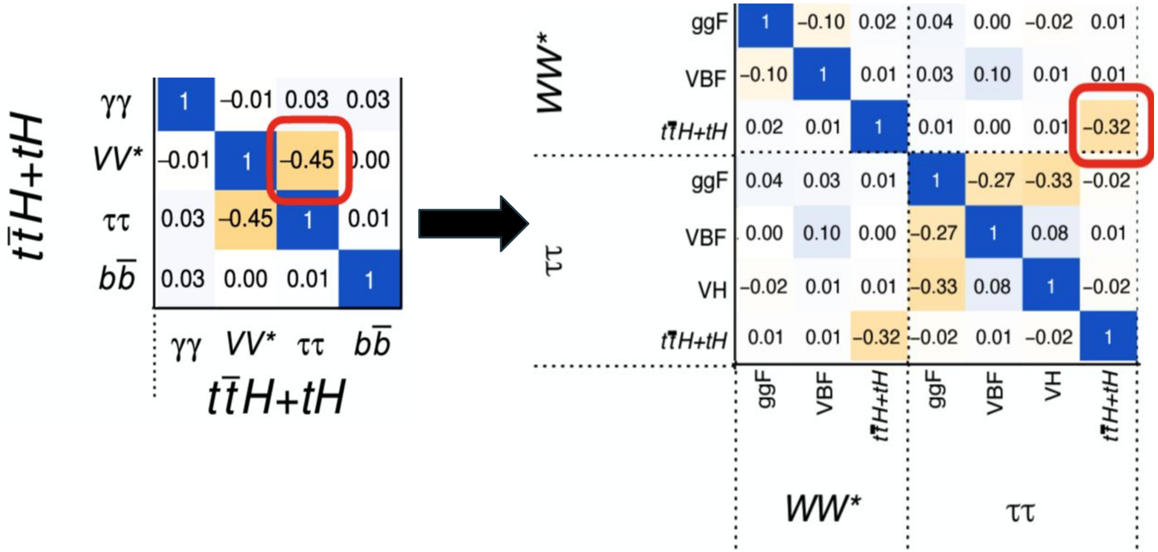


→Measured μ_{VBF} with uncertainty of +22%, -19%
→From Higgs combination has been seen that **$\tau\tau$ channel** is one of the most powerful and sensible to VBF production mode

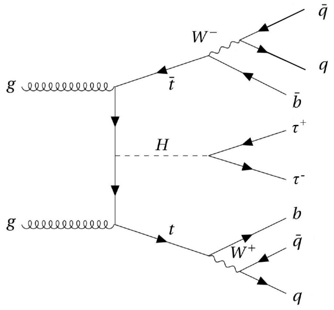
Previous analysis - Highlights

- General ATLAS Higgs combination: reduced ttH(WW) and ttH($\tau\tau$) anti-correlation **from -45% to -32%.** Attempted improvement in new legacy analysis

Nature 607 (2022) 52



ttH event selection



Signal topology: fully hadronic top-antitop decays

- 6 jets, 2 of them coming from b-quark hadronization
- Only considering hadronically decaying tau-leptons

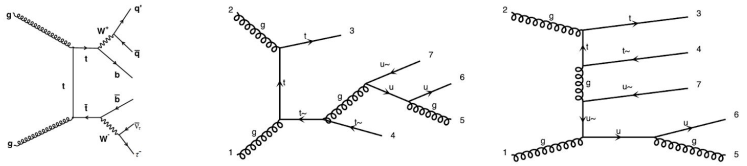
Event selection criteria:

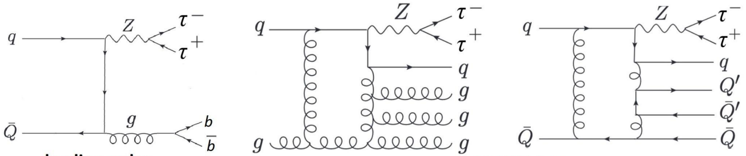
	Number	p_T	$\|\eta\|$	Other
τ_{had}	≥ 2	> 20 GeV	< 2.5	opposite charge, medium RNN identification
Jets	$\geq 5(6)$	> 20 GeV	< 4.4	anti-κ_t, $R = 0.4$
b-jets	$\geq 2(1)$	> 20 GeV	< 2.5	70% efficiency working point

MMC: (~) invariant mass of the di-τ system (missing mass calculator) →

- **Plays an important role**: first separation of signal against main background can be made applying cuts on this variable
- Signal-to-background enhanced regions can be defined out of this variable

Higgs $m_{\tau\tau}^{MMC}$ window
(100 GeV$< m_{\tau\tau}^{MMC} < 150$ GeV)
Signal enhanced region.
High $m_{\tau\tau}^{MMC}$ sideband
($m_{\tau\tau}^{MMC} > 150$ GeV or $m_{\tau\tau}^{MMC} = 0$)
Dominated by $t\bar{t}$ events.
Low $m_{\tau\tau}^{MMC}$ sideband
($0 < m_{\tau\tau}^{MMC} < 100$ GeV)
Dominated by $Z(\tau\tau)$ production.





leading order

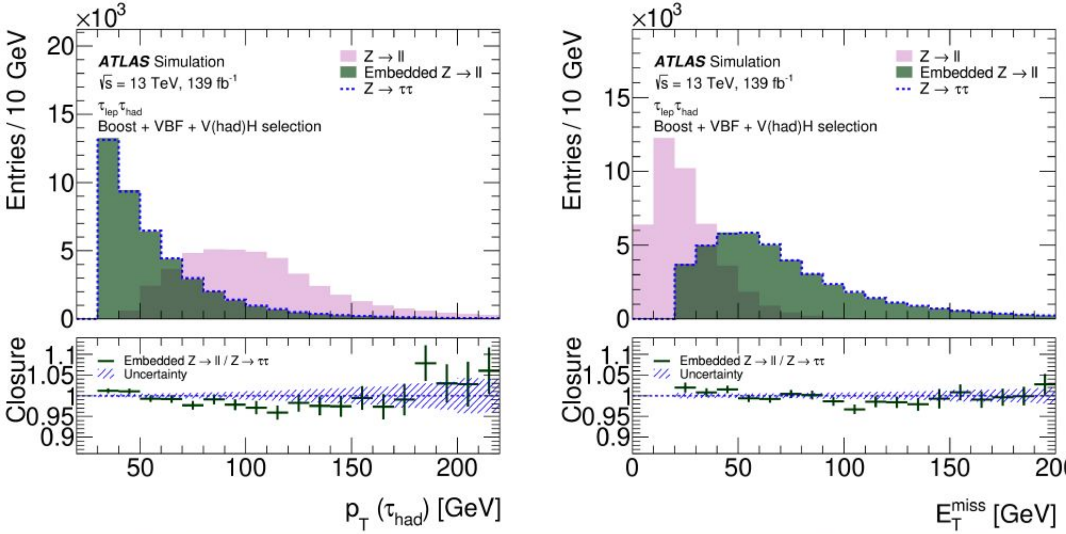
+3 QCD jets

Input variables for MVA trainings

	Variable	VBF	ttH multiclass
Jet properties	Invariant mass of the two leading jets	•	
	$p_T(jj)$	•	
	Product of η of the two leading jets	•	
	Sub-leading jet p_T	•	
	η of the 5 leading jets		•
	Scalar sum of all jets p_T		•
	Scalar sum of all b-tagged jets p_T		•
	Best W-boson candidate dijet invariant mass		•
	Best top-quark candidate three-jet invariant mass		•
Angular distances	$\Delta\phi$ between the two leading jets	•	
	$\Delta\eta$ between the two leading jets	•	
	Minimum ΔR between two jets		•
	Minimum ΔR between a b-tagged jet and a $\tau_{\text{had-vis}}$		•
	$\lvert\Delta\eta(\tau,\tau)\rvert$		•
	$\Delta R(\tau,\tau)$		•
τ-lepton properties	$p_T(\tau\tau)$		•
	Sub-leading τ p_T		•
	Leading τ η		•
H candidate plus jets system	$p_T(Hjj)$	•	
$\vec{p}_T^{\text{ miss}}$	Missing transverse energy E_T^{miss}		•
	Smallest $\Delta\phi(\tau, \vec{p}_T^{\text{ miss}})$		•

Background Estimation

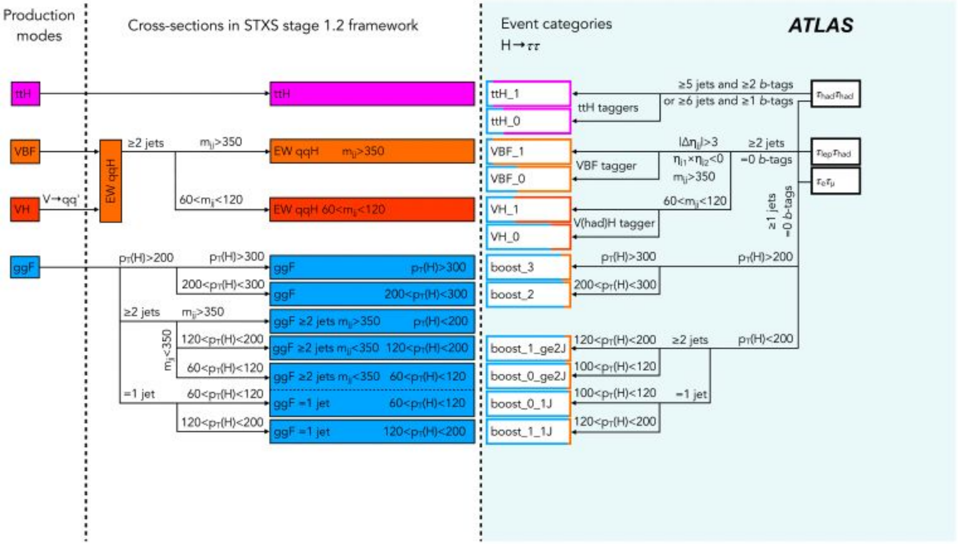
- For this new round, background estimation strategy is mostly inherited from previous analysis
- **Z→$\tau\tau$ is the largest background.** Object-level-embedding is employed to build control regions out of Z→ℓℓ events.
  - Kinematic cuts are applied on embedded τ objects that are created by splitting the ℓ into a visible and a neutrino component
  - Each signal region has single bin Z control region to constrain the Z contribution. MC is used to model the $m_{\tau\tau}$ contribution in the SR.
- **Fake leptons background** estimation: same Matrix Method (MM) in the $\tau_e\tau_\mu$ channel as before, and a fake factor derived for $\tau_\ell\tau_h$ and $\tau_h\tau_h$ channels.



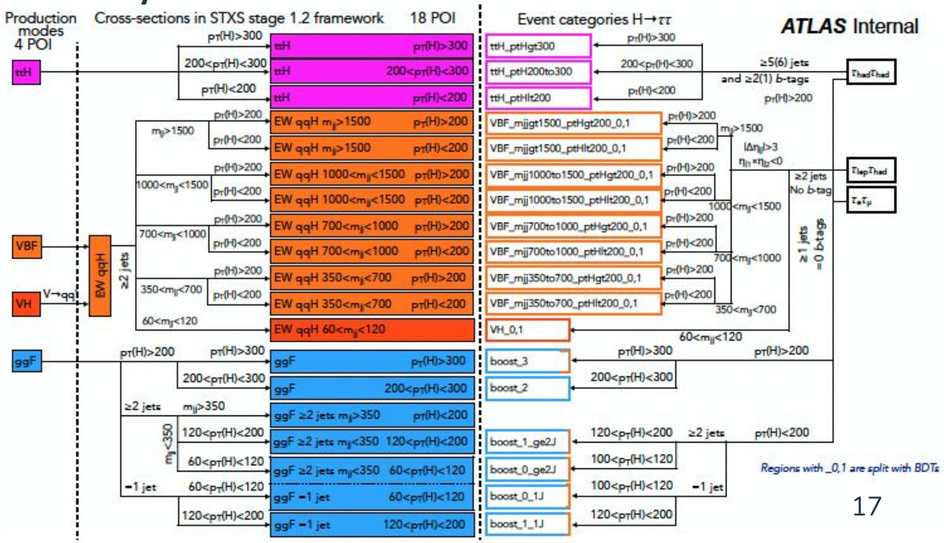
STXS Statistical fit - Fit setup

- Binned maximum likelihood fit using TRExFitter. A total of 3 differents fits are carried out, for different combination of parameters of interest (PoI):
  - **1-PoI** fit: inclusive $H \rightarrow \tau\tau$ production combining main Higgs production modes
  - **4-PoI** fit: dedicated PoI for each production mode i.e ggH, VBFH, VH, ttH. Targets $\sigma_i x BR_{\tau\tau}$ for each mode
  - **18-PoI** fit: nominal STXS fit, with setup optimized to measure the maximum number of possible STXS bins

Previous analysis

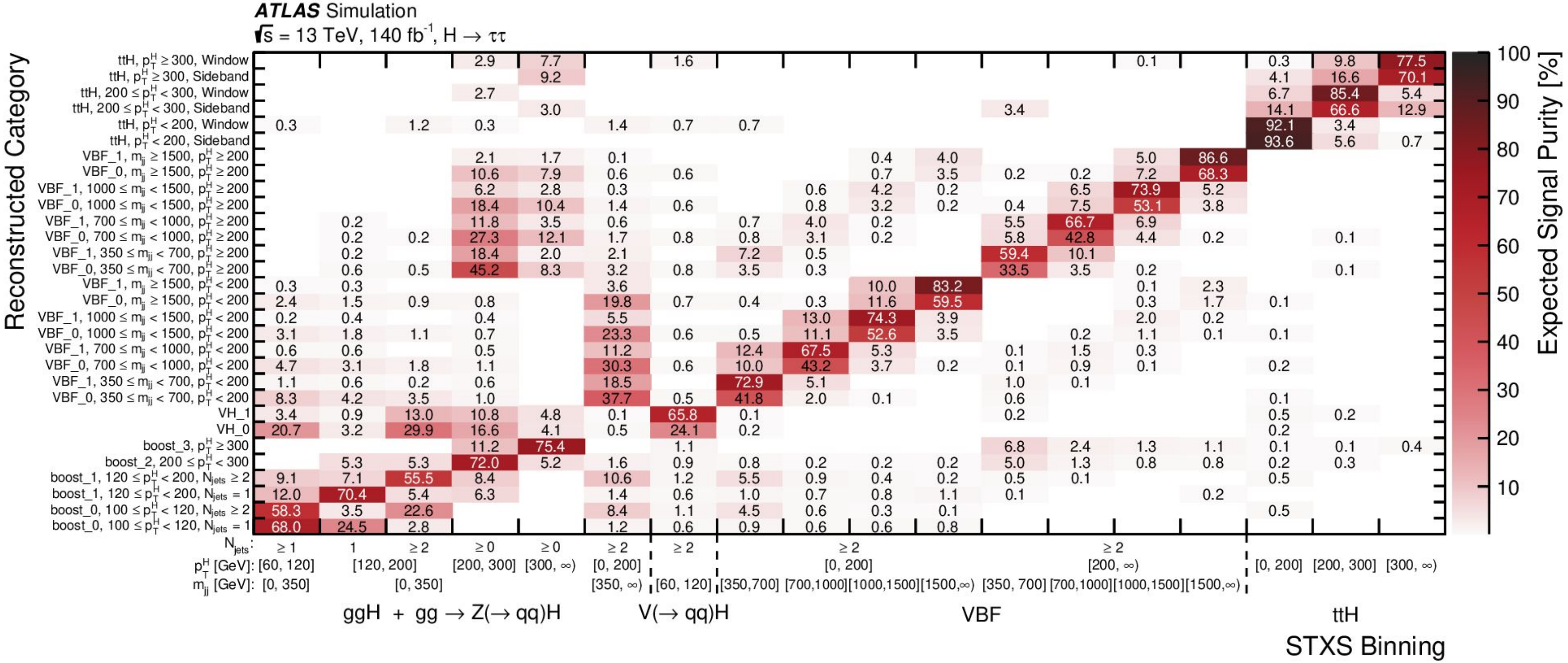


New analysis



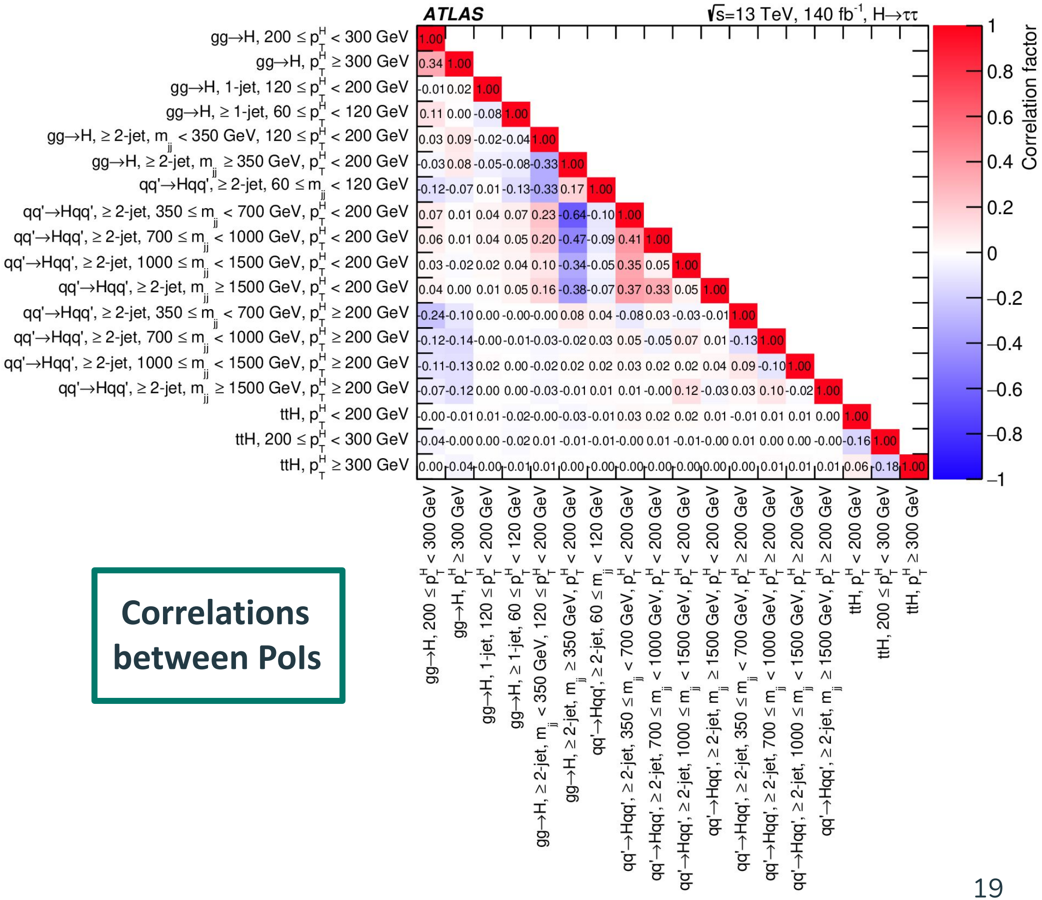
STXS - 18 PoI fit

Pre-fit signal purity in each reconstructed category (per bin)



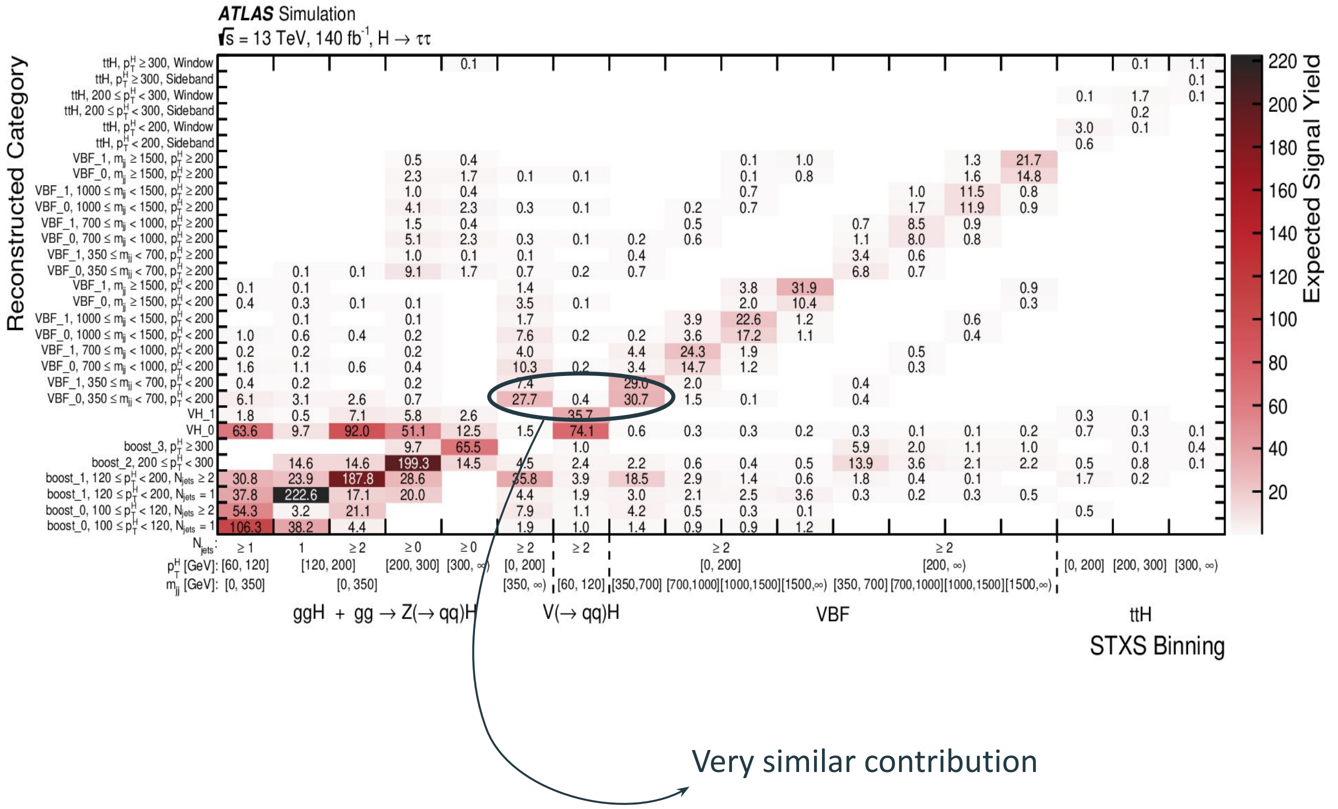
STXS - 18 PoI fit

- 18 PoI fit corresponding to the different STXS regions considered
  - Obtained better precision in VBF phase space for higher p_T^H and/or m_{jj} due to reduced SM backgrounds
  - VBF cross-sections at lower m_{jj} and p_T^H<200 GeV slightly below the SM prediction
  - Significant VBF-like ggH contribution (ggH+2 jet production with m_{jj} > 350 GeV, p_T^H < 200 GeV) in reconstructed level categories targeting VBF signal (see next slide)
    - Leads to **anti-correlation** in the measurements



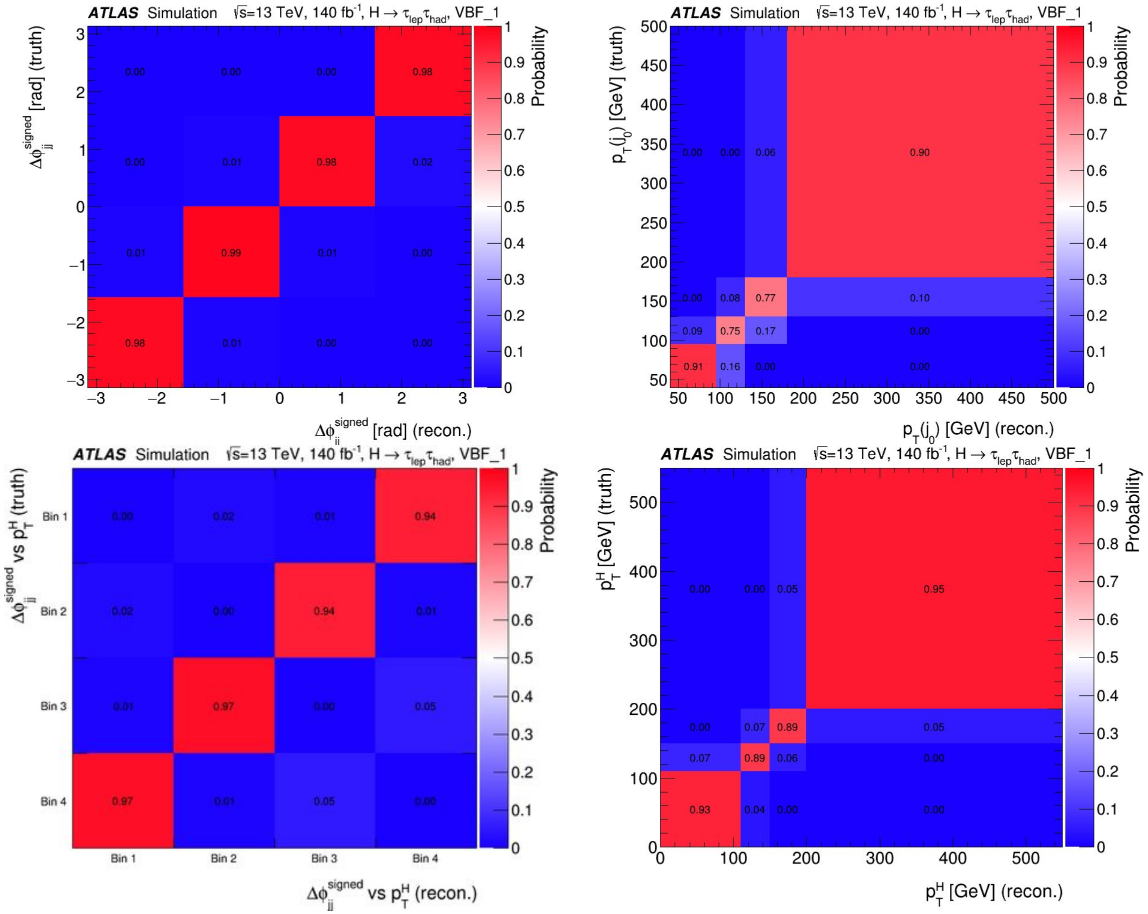
Correlations between PoIs

STXS - 18 PoI fit



- We knew that VBF-like ggH contribution was significant in the VBF 0 SRs. However, **we thought that the separation between VBF 0 and VBF 1 provided by the VBF tagger (VBF vs Ztt+ggF) was enough to isolate efficiently the VBF signal contributions.**

Unfolded differential measurement



- **Migration matrices** evaluated from MC simulations of Higgs bosons decaying to $\tau_{lep}\tau_{had}$
- Each matrix element is the probability for a signal event generated in a fiducial truth-bin to be selected in a VBF_1 reconstructed (recon.) bin in the тlepтhad channel.

18-PoI unblinded fit results

- Limited sensitivity for ttH obtained in the fit due to poor statistics, compared to other channels like H→bb. **Upper exclusion limits at 95% CL** can be also computed for the three PoIs on signal strength

H→bb (HIGG-2020-24)