Measurements of Higgs boson production with top quarks with the ATLAS detector

ATLAS EXPERIMENT

ICHEP 2024

Filip Nechanský

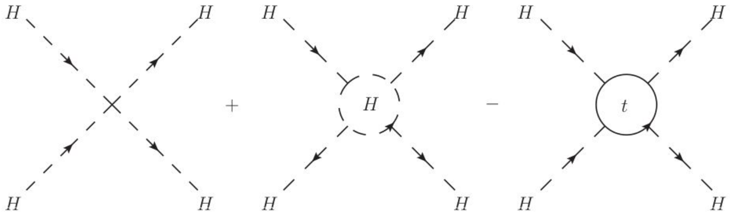
FZU Institute of Physics of the Czech Academy of Sciences

Co-funded by the European Union

PHYSICS FOR FUTURE

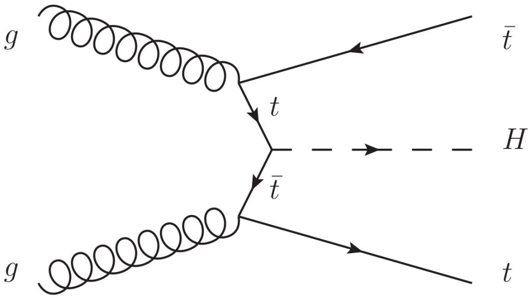
Top to Higgs coupling

- Top-Higgs strong Yukawa coupling $y_t \sim 1$
- **Consequence**:
  - Significant correction to Higgs self-coupling



  - Major contributor to fermion loops in general

- Can be measured indirectly (e.g. H->yy)
  caveat: assumes no new physics
- H→tt not accessible due to $m_H << 2m_t$
- **Direct measurement!** E.g. in tH or **ttH**:



ttH in nutshell

- Production channel with low cross-section
- Both Higgs and top have diverse decay channels,
  - Impossible to cover in single analysis
  - Relies on most objects ATLAS reconstructs and identifies:
    e, μ, γ, hadronically decaying τ-jets, b-jets, missing energy …

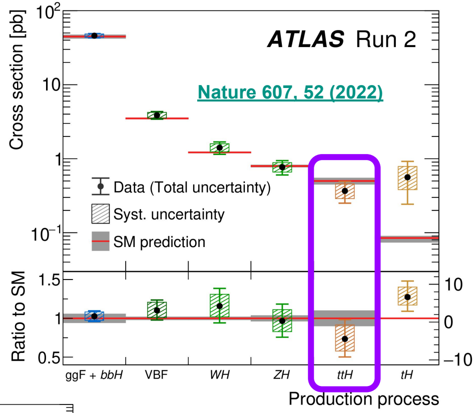
- Analysis split based on Higgs-decay,
  each with unique challenges:

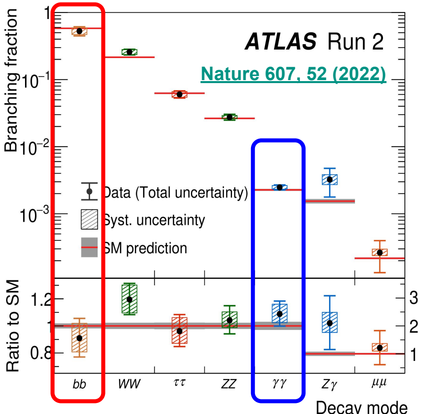
difficult background in ttH(bb)

vs

low statistics of ttH(yy)

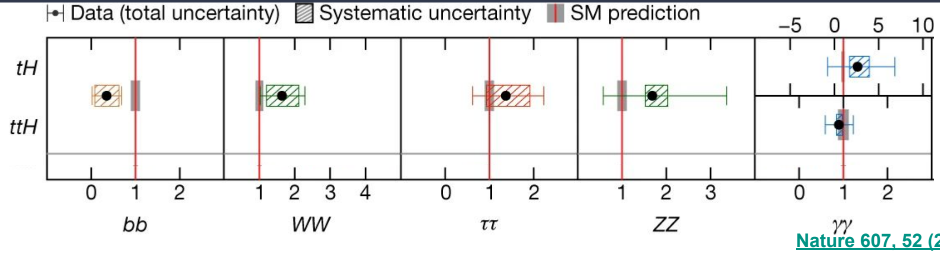
- Observation by both CMS (Phys. Rev. Lett. 120 (2018), 231801)
  and ATLAS (Phys. Lett. B 784 (2018) 173)
  (partial Run 2)





ttH measurements overview

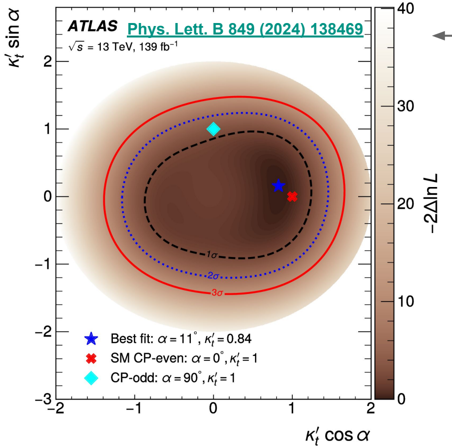
Latest combination in [Nature 607, 52 (2022)]



Nature 607, 52 (2022) (modified)

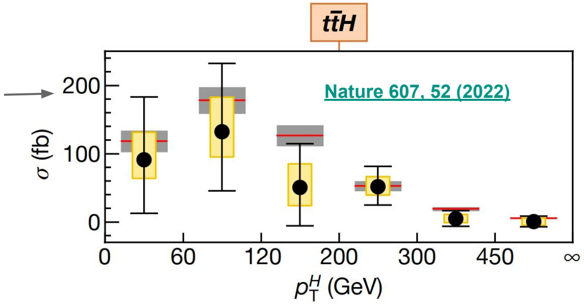
Individual channels:

- ttH(bb) → this talk!
- ttH multi-lepton [Phys. Rev. D 97 (2018) 072003]
- ttH(ZZ) [Eur. Phys. J. C 80 (2020) 957]
- ttH(yy) [JHEP 07 (2023) 088]



← CP-odd measurements,
more details on Higgs CP in [Matthew Basso's talk]

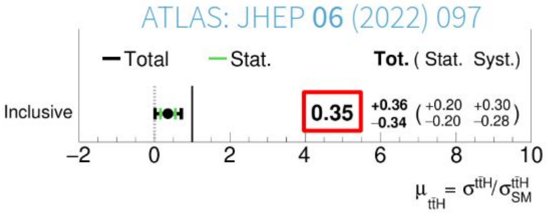
p_T(H) differential (STXS [10.23731/CYRM-2017-002]), → see [talk by Xiao Yang] for STXS overview



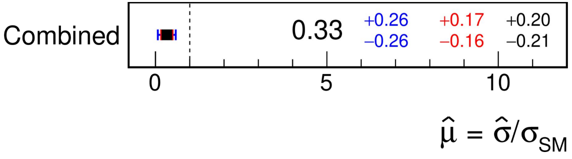
ttH(bb) with full Run 2 data

JHEP 06 (2022) 97

- Main challenge: tt+bb background
  - How to separate the signal?
  - What are the modeling uncertainties? (heavy quarks + large number of jets!)
- Expected sensitivity 2.7σ, observed 1σ:

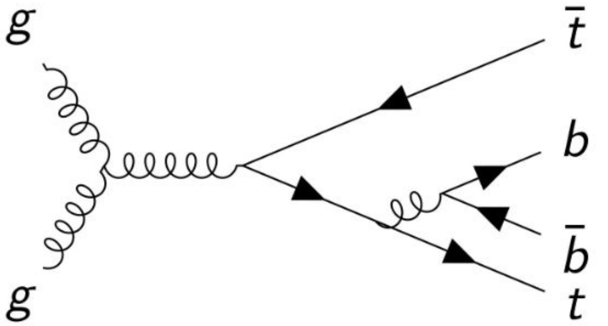


arXiv:2407.10896



- Also measured $p_T(H)$ in STXS scheme

$t\bar{t} + b\bar{b}$ background

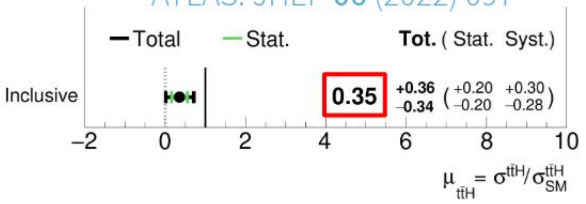


ttH(bb): We can do better!

New!

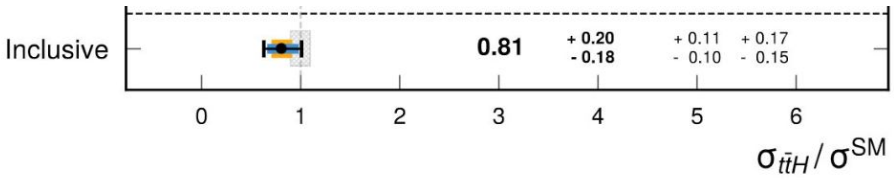
New legacy analysis of Run 2 data: **Expected sensitivity increased from 2.7σ to 5.4σ**!

ATLAS: JHEP **06** (2022) 097



~3 years of work →

today



ATLAS managed to improve precision
and reach 4.6σ observed significance in ttH(bb) alone!

… what did it take?

Problem tackled from three major directions

Improved reconstruction and particle identification

- ATLAS-wide effort
- Particle-flow jets
- Improved tagging of b-jets

State-of-the-art machine learning

- Improved signal / background classification
- Improved Higgs reconstruction

Improved modeling of background

- Consistent tt+bb systematic model
- Data-driven corrections

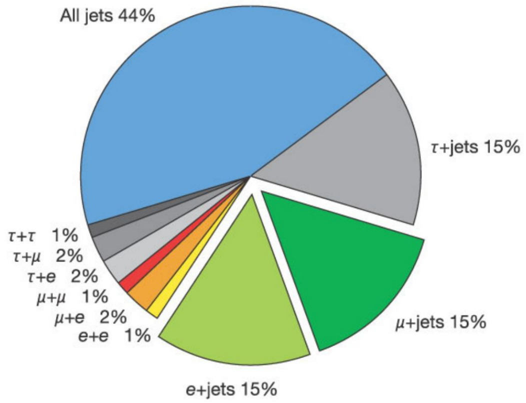
Legacy ttH(bb) analysis

Also presented as poster

Link to paper

- Analysis done for leptonic decay of ttbar
  - Single-lepton: resolved and boosted (~45% BR)
  - Dilepton: Inclusive (<9% BR)
- Goal to measure both inclusive XS and $p_T(H)$ in STXS model: [0, 60, 120, 200, 300, 450+] GeV
  - extra bin wrt to 2021 result
- Looser selection wrt 2021 result
  - Single lepton: 5 jets, 3 b-jets (85% b-tag eff.)
  - Dilepton: 2 b-jets (85% b-tag eff.) + 1b-jet (70% b-tag eff.)
  - Allows better control of tt+jets backgrounds
  - Factor 3 larger signal acceptance

ttbar branching ratio:

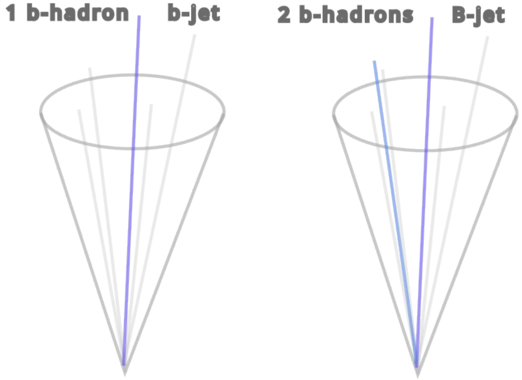


www.quantumdiaries.org - Top quark

Monte Carlo overview

- Heavy flavor classification of ttbar,
  based on hadron presence in additional jets:
  - tt+≥1b: at least one jet with b-hadron
  - tt+≥1c: not tt+≥1b
    + at least one jet with c-hadron
  - tt+light: not tt+≥1b/tt+≥1c

- tt+c and tt+light still contribute
  (looser selection + flavor tagging not perfect)



Process	Sample
$t\bar{t}H$	Powheg + Pythia8
$t\bar{t}$ + light, $t\bar{t} + \geq 1c$	Powheg + Pythia8 $t\bar{t}$ @NLO 5FS
$t\bar{t} + \geq 1b$	Powheg + Pythia8 $t\bar{t} + b\bar{b}$ @NLO 4FS
tH, $t\bar{t}V$, $t\bar{t}t\bar{t}^{*}$	aMC@NLO + Pythia8
V + jets, diboson	Sherpa
Single top	Powheg + Pythia8

tt+bb can have different signature in the detector,
with different impact of modeling.
⇒ three categories of tt+b considered in legacy analysis:

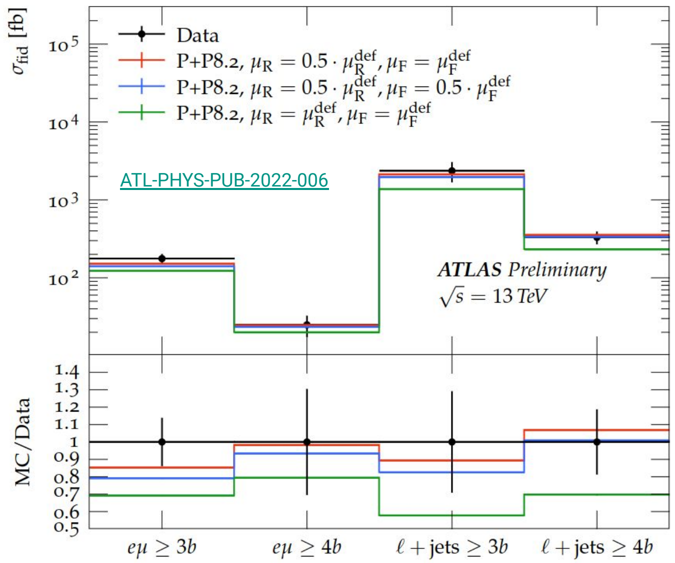
- tt+≥2b
- tt+1B (e.g. bb in single jet)
- tt+1b (e.g. 1 jet out-of-phase-space)

Modeling of tt+bb

- In-between the analyses, ttbar modeling was scrutinized [ATL-PHYS-PUB-2022-006]
  - updated renormalization scale, nominal model closer to data
  - updated systematic model

- For the first time: **consistent 4FS Monte Carlo used for tt+bb** (2021 analysis partially relied on 5FS)

Uncertainty	Legacy analysis
ISR	Var3c (PS)
FSR	μ_R FSR (PS)
ME scale	Independent μ_R/μ_F
NLO matching	PP8 4FS pthard = 1 4FS
PS & Hadronisation	PH7 4FS
Parton shower	PP8 4FS dipole recoil



Also new dedicated tt+bb measurement CERN-EP-2024-191, presented yesterday by Ricardo Goncalo

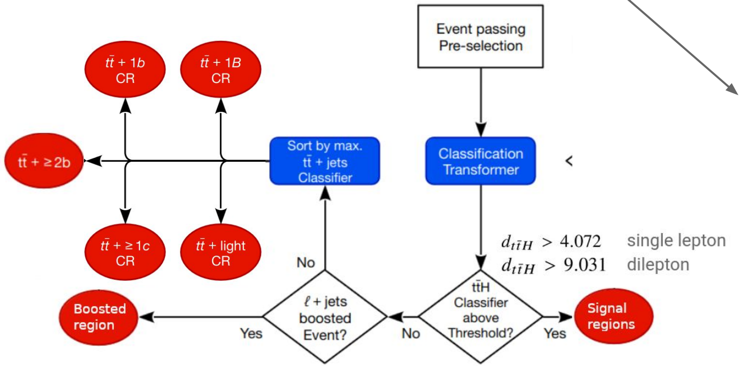
Multivariate analysis

- State-of-the art NN classifier using *transformers with attention mechanism* [arXiv:1706.03762]

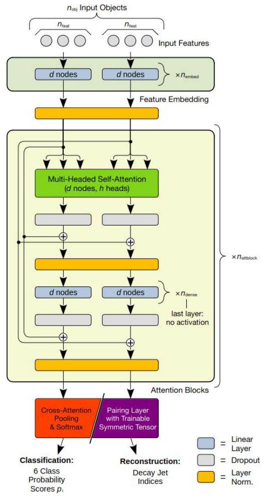
- Low-level objects used as inputs (e.g. 4-vectors, b-tagging discriminant)

- Similar structure used to classify processes and to reconstruct properties, difference in the last step and the training

- **Classification transformer:**
    - Class for ttH + all tt+jets categories, define analysis regions based on score
    - Select boosted events in single lepton



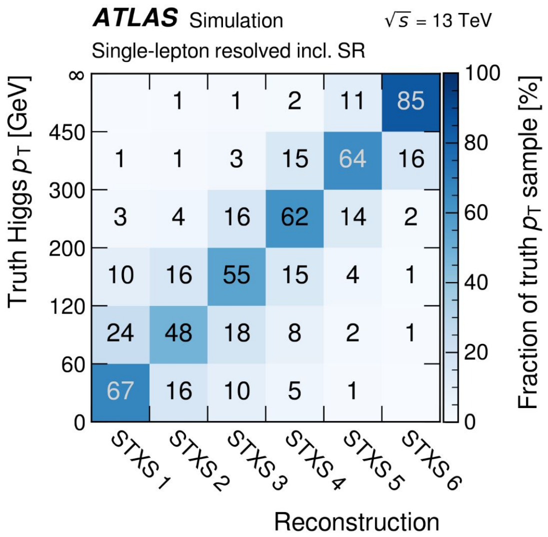
Networks are getting more complex:

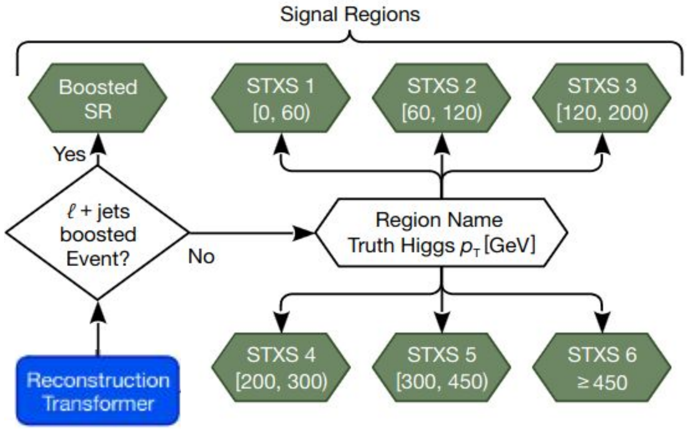


More details in the publication or [Yusong Tian's poster](#)

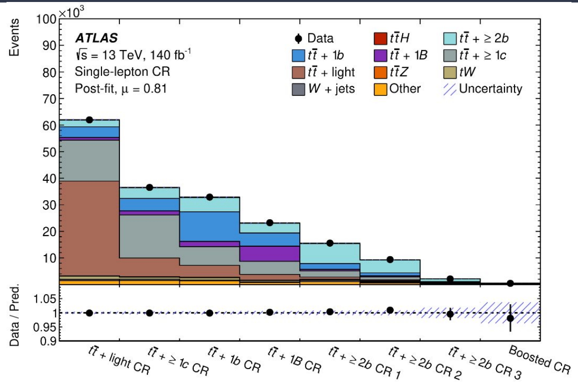
Reconstruction transformer

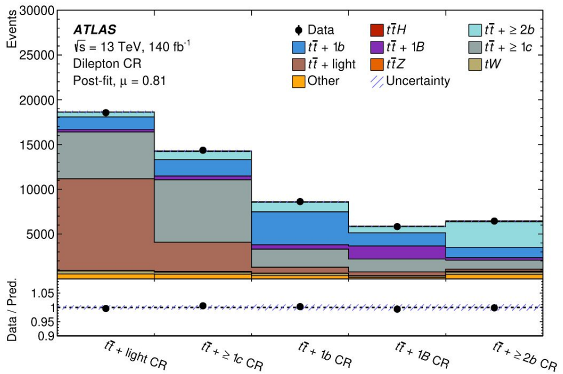
- Find b-jets from decay of the Higgs
- Structure similar to SPANet [Phys. Rev. D 105 (2022) 112008]
- Used to define regions for different bins of the $p_T(H)$ for STXS
- Also employed for tt+≥2b for better control over the background

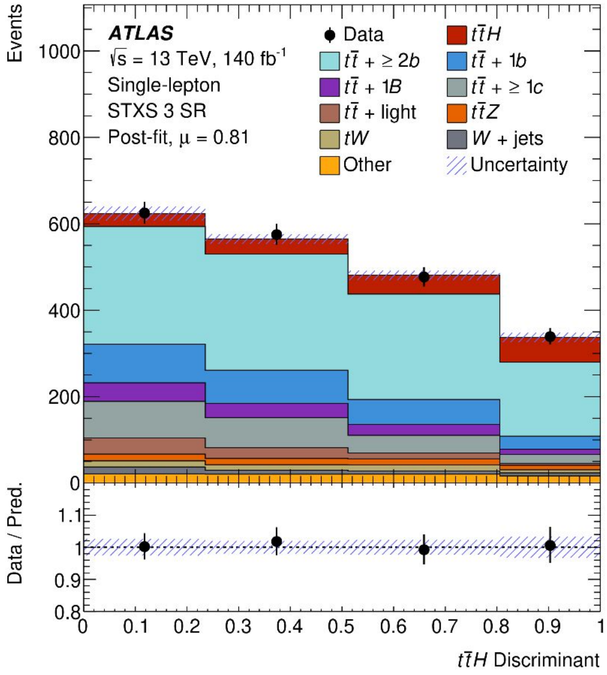




Statistical analysis







- Profile likelihood fit of all signal and control regions
- Fit variables:
  - Boosted region in single lepton: reconstructed Higgs pT
  - Otherwise discriminant variables (different for each region!)
- 8 free-floating normalization factors for major ttbar background components
- 1 (6) signal parameters for inclusive (STXS) result
- **5.4σ expected sensitivity!**

Results

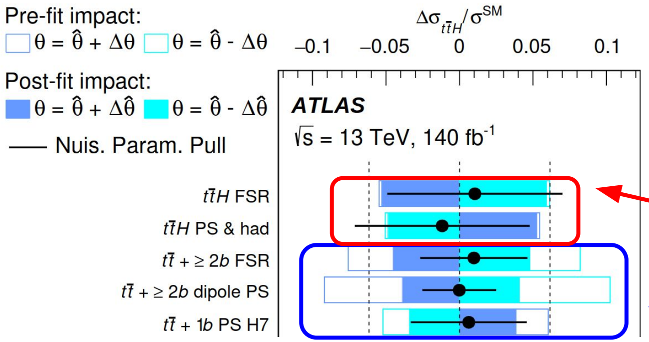
All these improvements come together to deliver significantly improved measurement of ttH(bb)!

Inclusive cross-section:

$$\sigma_{t\bar{t}H} = 411\ ^{+101}_{-92}\ \text{fb} = 411 \pm 54(\text{stat.})\ ^{+85}_{-75}(\text{syst.})\ \text{fb},$$

→ dominated systematically

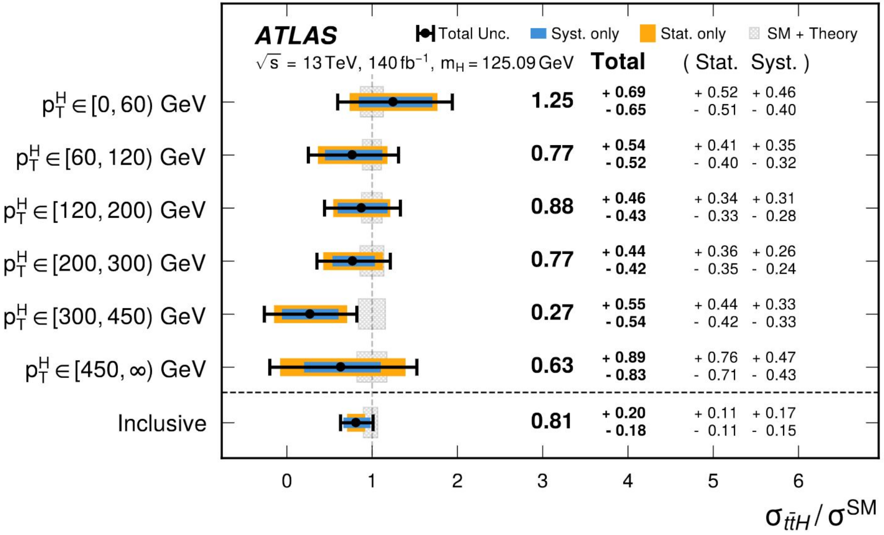
Best single channel measurement of ttH!



Dominant uncertainties:

- from ttH modeling
- tt+bb modeling

(constraints from dedicated tt+≥2b CRs and SRs!)



$p_T(H)$ measurement limited statistically

In conclusion

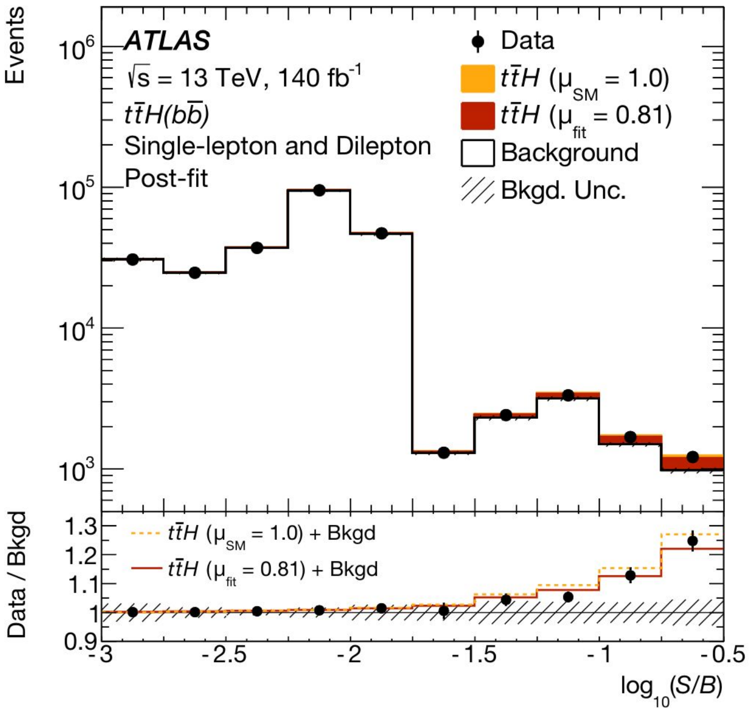
Still a lot to gain from studying Run-2 data!

Improvements coming from all aspects of analysis

- Object reconstruction/performance
- Analysis strategy, state-of-the-art methods
- Theoretical prediction and uncertainties

Improvement in measured (expected) significance from 1.0σ (2.7σ)
to 4.6σ (5.4σ) with the same Run-2 dataset!

Ultimate precision will be reached in combination with the other ttH channels, ttH(bb) to play significant role



Backup

Classification transformer

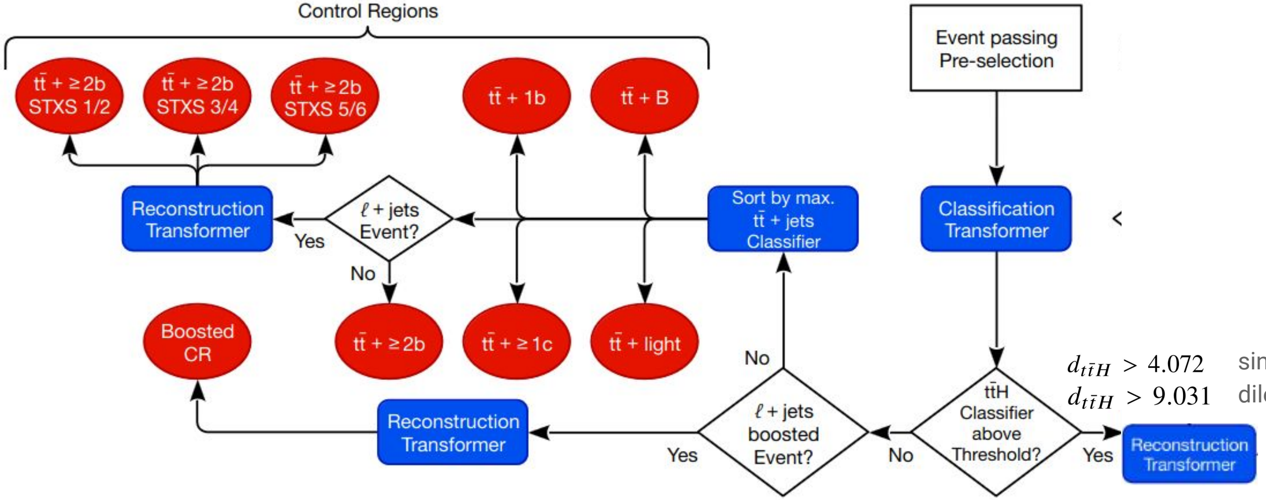
Define signal and control regions for all tt+jets background categories
+ separate classifier for boosted events (using large-R jets) in single lepton channel

Discriminant:

$$d_i = \frac{p_i}{\sum_{j \neq i} p_j \cdot \hat{N}_{ij}}.$$

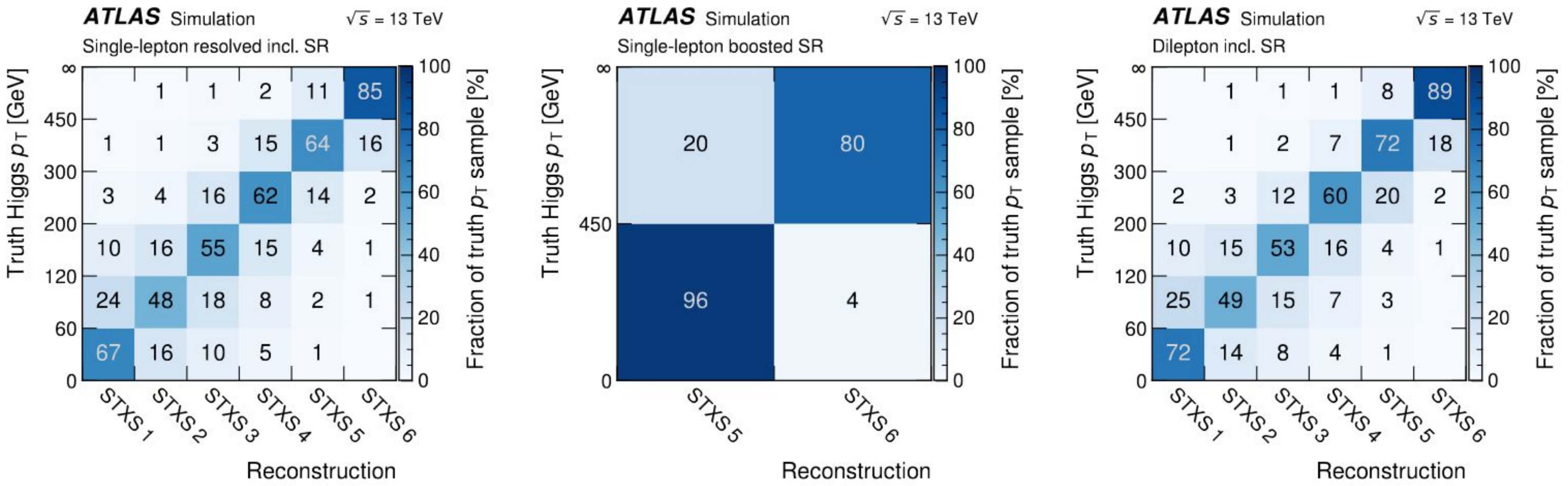
$$\hat{N}_{ij} = N_j / \sum_{k \neq i} N_k$$

where:
p_i: average network output for process i
N_i: event yield of process i



$d_{t\bar{t}H} > 4.072$ single lepton
$d_{t\bar{t}H} > 9.031$ dilepton

Reconstruction transformer performance



CP nature of Higgs boson HIGG-2020-03

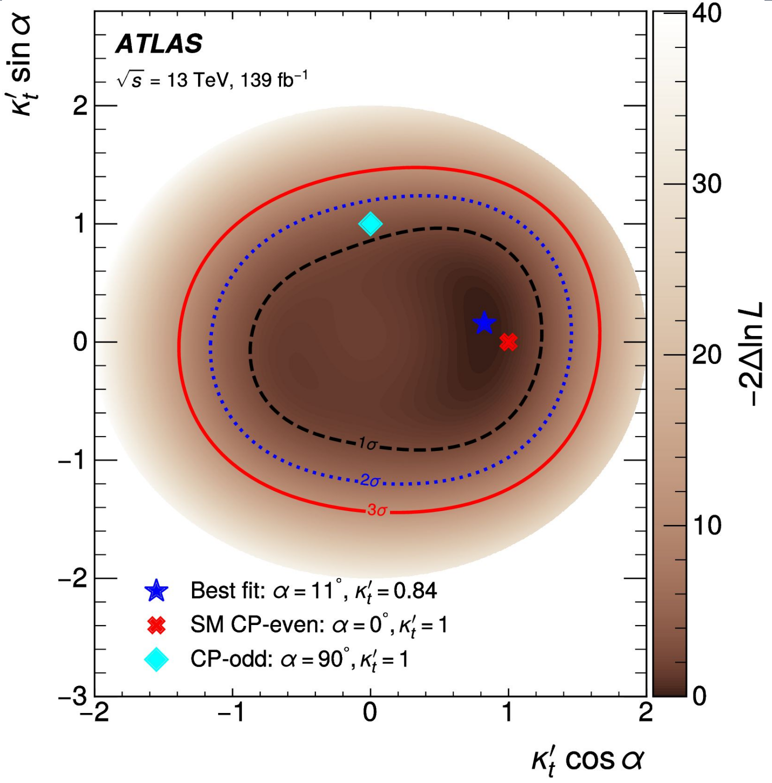
Phys. Lett. B 849 (2024) 138469

- Standard model predicts CP-even top-Higgs Yukawa

  -> is there CP-odd component?

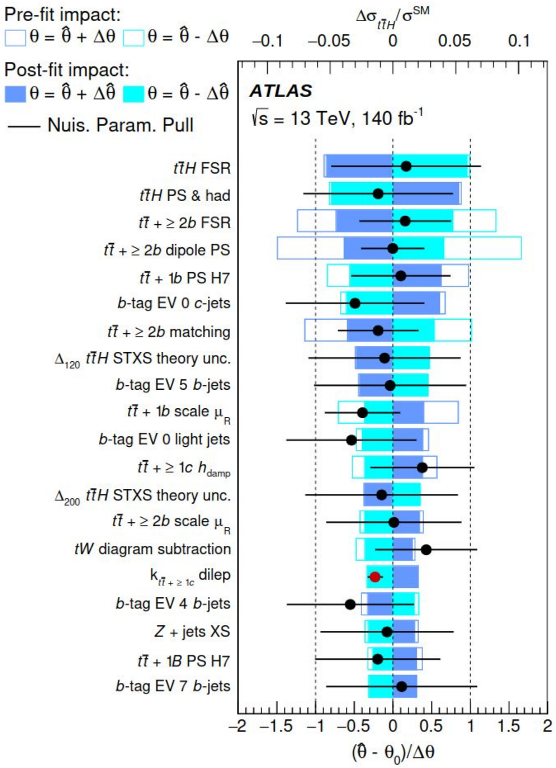
- Measured in ttH/tH with H->bb
- Measuring alpha mixing angle between CP-even and CP-odd
- Dedicated BDT sensitive to CP

- Result $\alpha = 11^{\circ}{}^{+52^{\circ}}_{-73^{\circ}}$,compatible with SM prediction

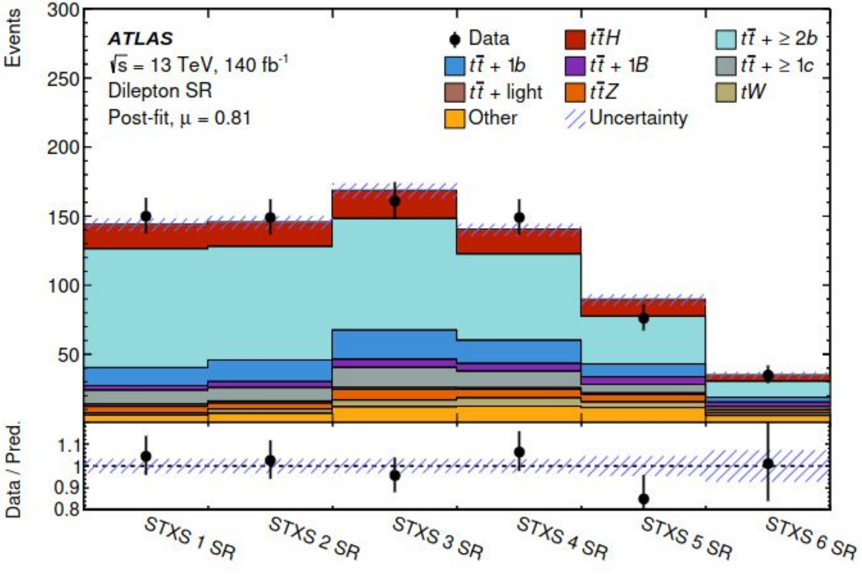


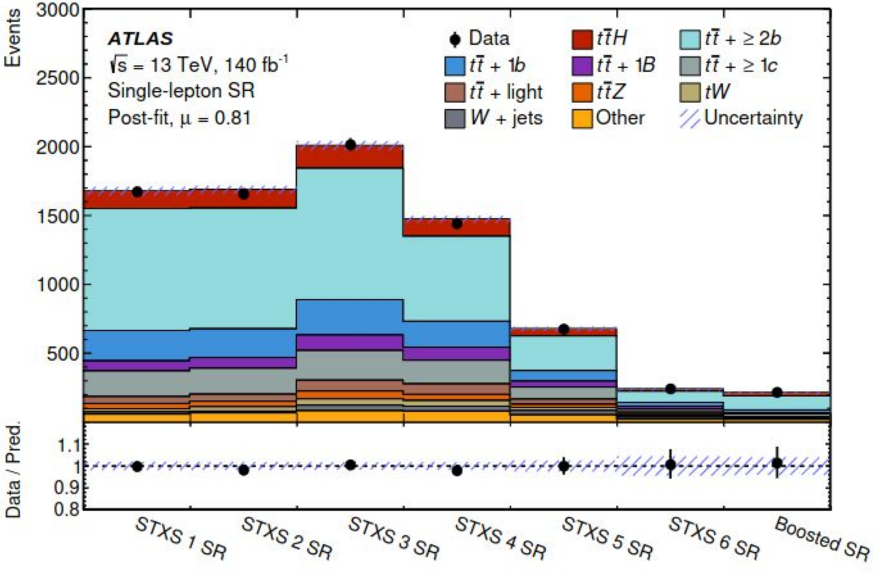
Uncertainty sources and impact

Uncertainty source	$\Delta\sigma_{t\bar{t}H}$ (fb)		$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ (%)	
Process modelling				
$t\bar{t}H$ modelling				
$t\bar{t}H$ radiation	+35	−21	+9	−5
$t\bar{t}H$ parton shower	+32	−19	+8	−5
$t\bar{t}H$ matching	<0.1	−0.3	<0.1	−0.1
$t\bar{t}H$ theory	+25	−17	+6	−4
$t\bar{t} + \geq 1b$ modelling				
$t\bar{t} + \geq 1b$ radiation	±31		±8	
$t\bar{t} + \geq 1b$ parton shower	±29		±7	
$t\bar{t} + \geq 1b$ matching	±19		±5	
$t\bar{t} + \geq 1c$ modelling	±18		±4	
$t\bar{t}$ + light modelling	±5		±1	
tW modelling	±16		±4	
Minor background modelling	±19		±5	
Flavour tagging	±36		±9	
Jet modelling	±22		±5	
Monte-Carlo statistics	±17		±4	
Other instrumental	±10		±2	
Total systematic uncertainty	+85	−75	+21	−18
Normalisation factors	±21		±5	
Total statistical uncertainty	±54		±13	
Total uncertainty	+101	−92	+25	−22

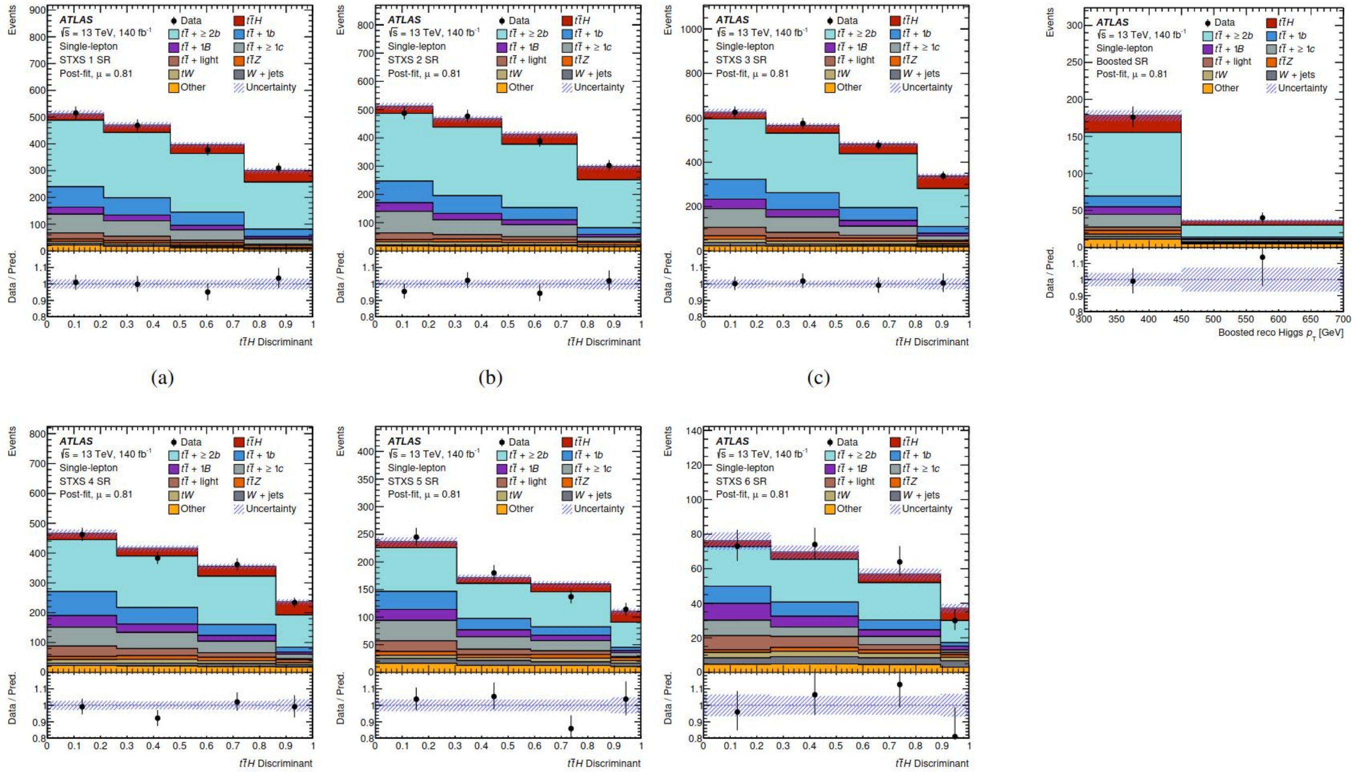


Post-fit overview of signal regions





Single lepton signal regions postfit

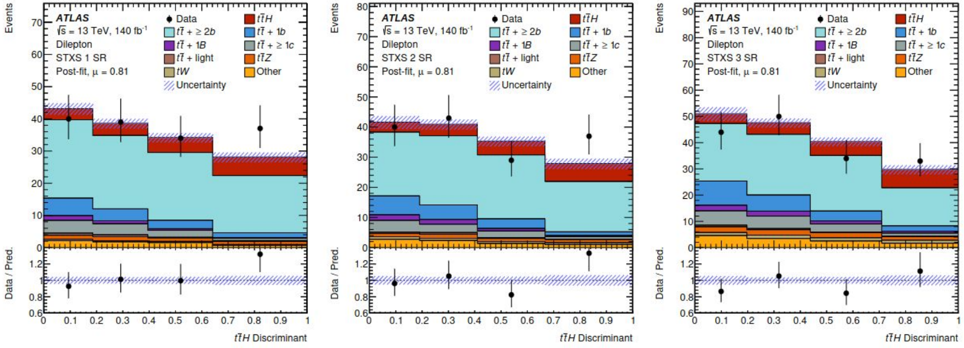


(a)

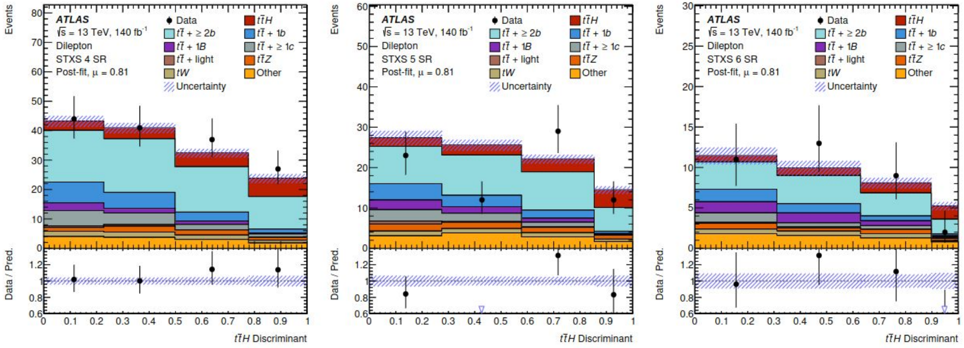
(b)

(c)

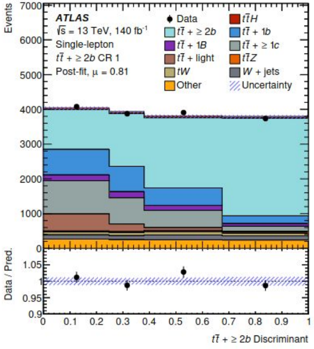
Dilepton signal regions postfit



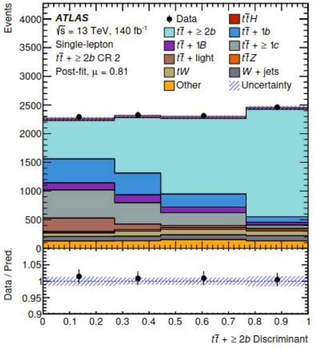
(a) (b) (c)



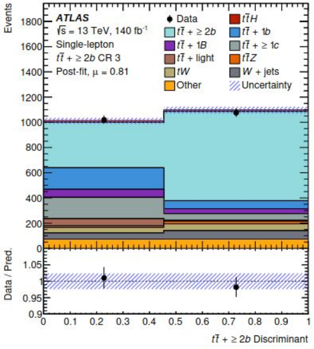
Single lepton control regions postfit



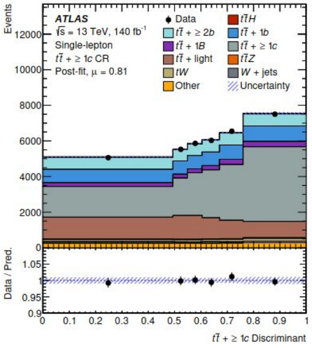
(a)

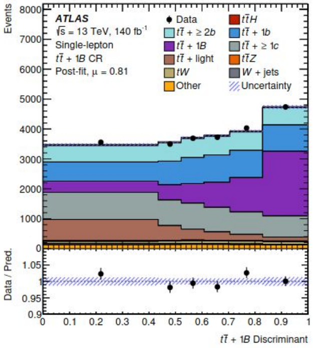


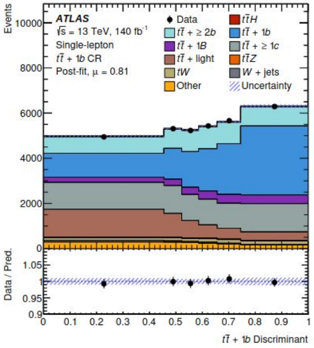
(b)

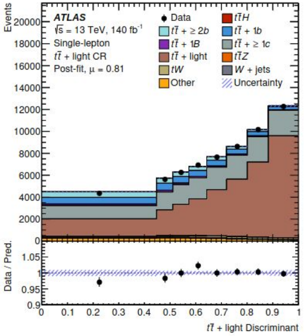


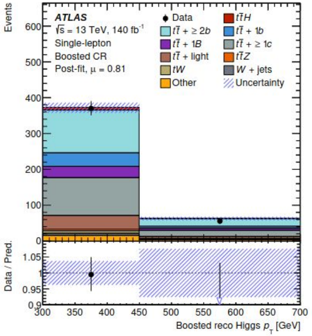
(c)



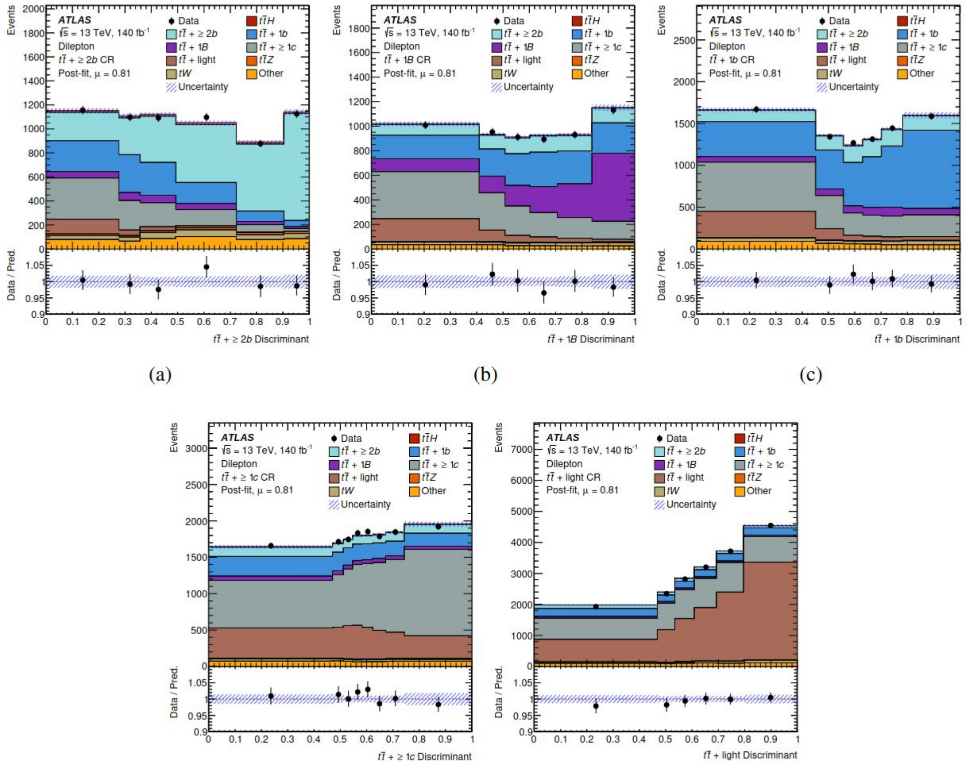








Dilepton control regions postfit



Systematic sources comparison

Uncertainty	Previous analysis	Legacy analysis	$t\bar{t}$ + jets components
ISR	Var3c (PS) and μ_R/μ_F (ME)	Var3c (PS)	All
FSR	μ_R FSR (PS)	μ_R FSR (PS)	All
ME scale	-	Independent μ_R/μ_F	All
NLO matching	aMC@NLO + Pythia8	PP8 4FS pthard = 1 4FS	$t\bar{t} + \geq 1b$
		PP8 4FS pthard = 1 5FS	$t\bar{t}$ + light, $t\bar{t} + \geq 1c$
PS & Hadronisation	Powheg + Herwig7 (PH7)	PH7 4FS	$t\bar{t} + \geq 1b$
		PH7 5FS	$t\bar{t}$ + light, $t\bar{t} + \geq 1c$
Parton shower	-	PP8 4FS dipole recoil	$t\bar{t} + \geq 1b$