FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

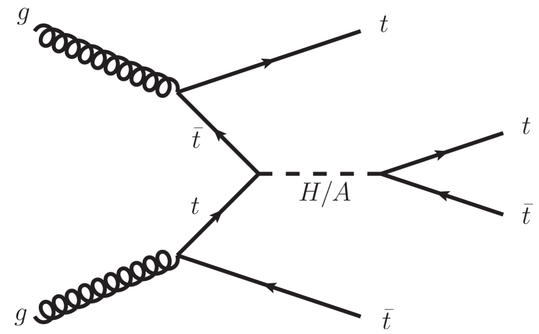
Search for a new scalar or pseudoscalar heavy Higgs boson using production of four top quarks at the LHC

Chainika Chauhan
On behalf of the ATLAS Collaboration
chainika.chauhan@cern.ch

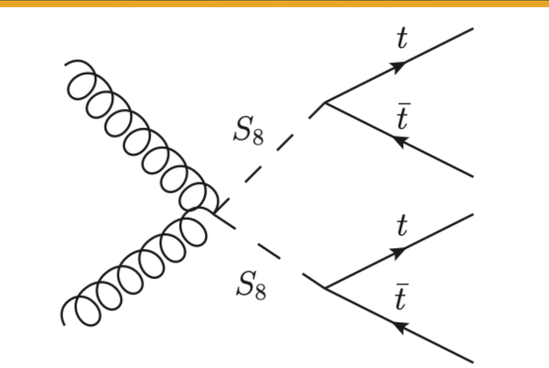
CORFU WORKSHOP 2024
Workshop on the Standard Model and Beyond
27 Aug 2024

Search targets …

- Search for Two-Higgs-Doublet-Model[1] (2HDM) type-II $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ signal in the alignment limit using data collected in pp colisions within the ATLAS detector .
  - This search is detailed in [ATLAS-CONF-2024-002].
  - Calculation of exclusion upper limits on the $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ production cross section.
  - Calculation of exclusion lower limits on $tan\beta$ as a function of mass of heavy scalar/pseudoscalar Higgs.
- Reinterpretation of results in sgluon Model[2] for signal $S_8S_8 \rightarrow t\bar{t}t\bar{t}$



[ATLAS-CONF-2024-002]



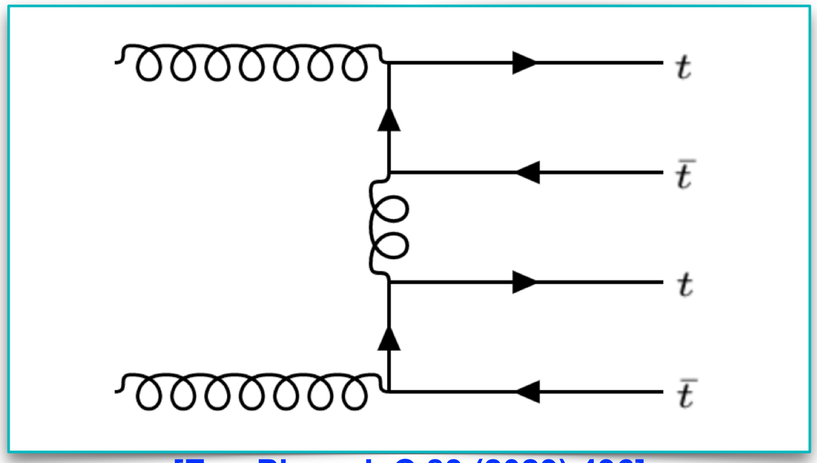
[ATLAS-CONF-2024-002]

1. Phys. Rept. 516 (2012) 1
2. JHEP 09 (2021) 143

Motivation

☑ Inconsistencies between theoretical and experimental Standard Model (SM) $t\bar{t}t\bar{t}$ process:

- Prediction[3] of $t\bar{t}t\bar{t}$ production from the SM: $13.37^{+1.04}_{-1.78}$ fb
- Measurement[4] of $t\bar{t}t\bar{t}$ production within 1.8 standard deviations of SM prediction: $22.5^{+4.7}_{-4.3}(stat)\ ^{+4.6}_{-3.4}(syst)$ fb *(QCD NLO+EW)*

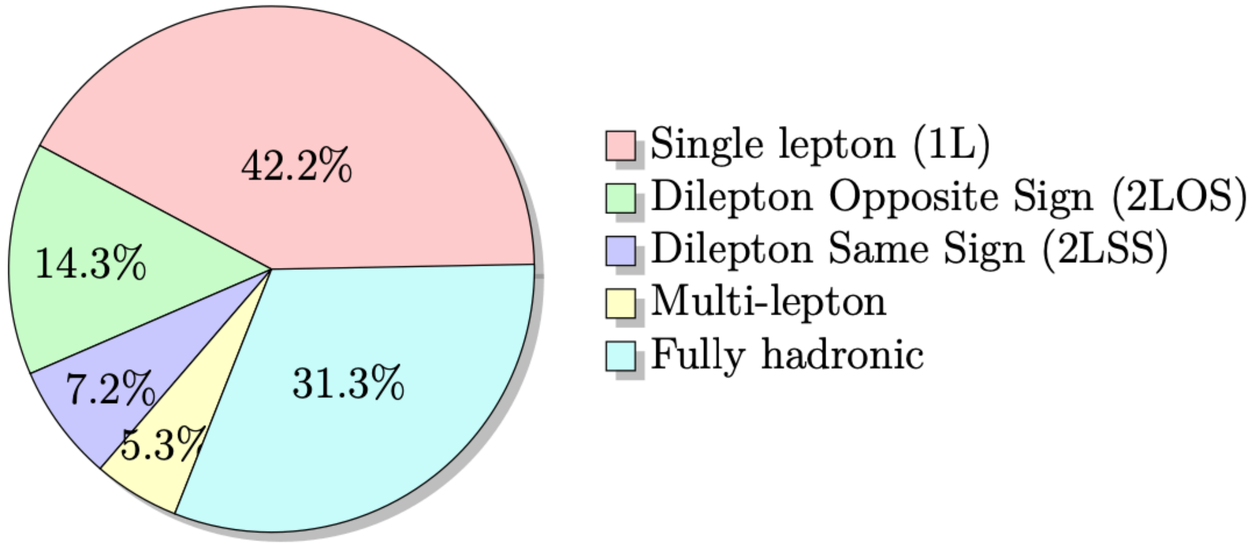


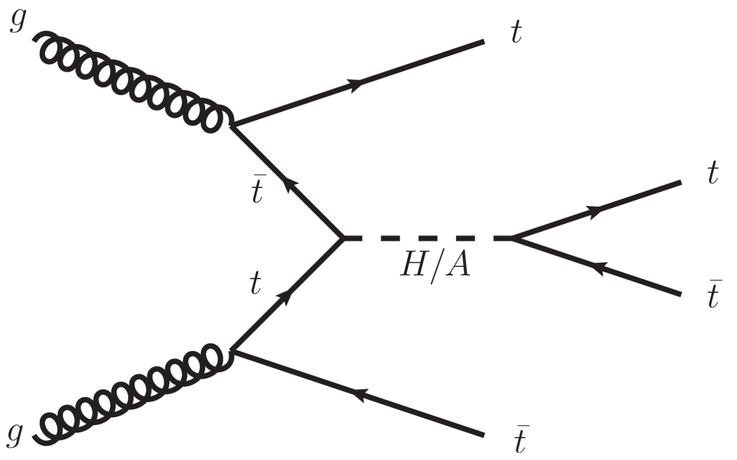
[Eur. Phys. J. C 83 (2023) 496]

3. arXiv:2212.03259
4. Eur. Phys. J. C 83 (2023) 496

Analysis Overview

- ☑ Signal: $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$
- ☑ Higgs mass ($m_{H/A}$): 400-1000 GeV
- ☑ Decay channel considered: One lepton and Two opposite-sign lepton
- ☑ Dominant background: $t\bar{t} + jets$
- ☑ $t\bar{t} + jets$ estimated using two types of data-driven correction factors:
  - Heavy flavor (HF) normalization rescaling
  - Kinematic reweighting using neural network
- ☑ Trained Graph Neutral Network (GNN) to separate signals from background.
- ☑ Profile likelihood fit in all signal-depleted and signal-rich regions.
- ☑ Similar search published in multi-lepton channel [JHEP 07 (2023) 203] .



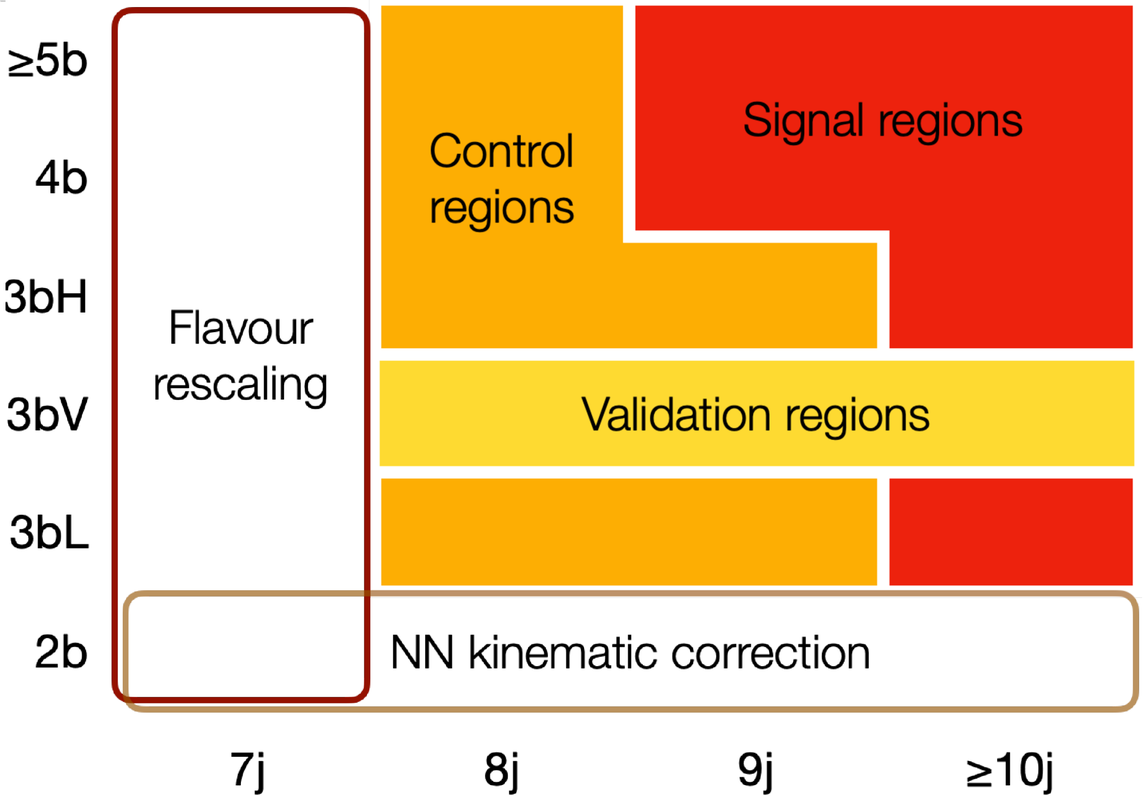


[ATLAS-CONF-2024-002]

Analysis Regions

- ☑ Events categorised based on the jets, b-tagged jets and lepton multiplicity.
- ☑ Different operating points (OPs) are calibrated having different efficiency, purity for tagging b-jets and rejection for non b-jets.
- ☑ In the analysis, a combination of these OPs is used to define regions for improved sensitivity.

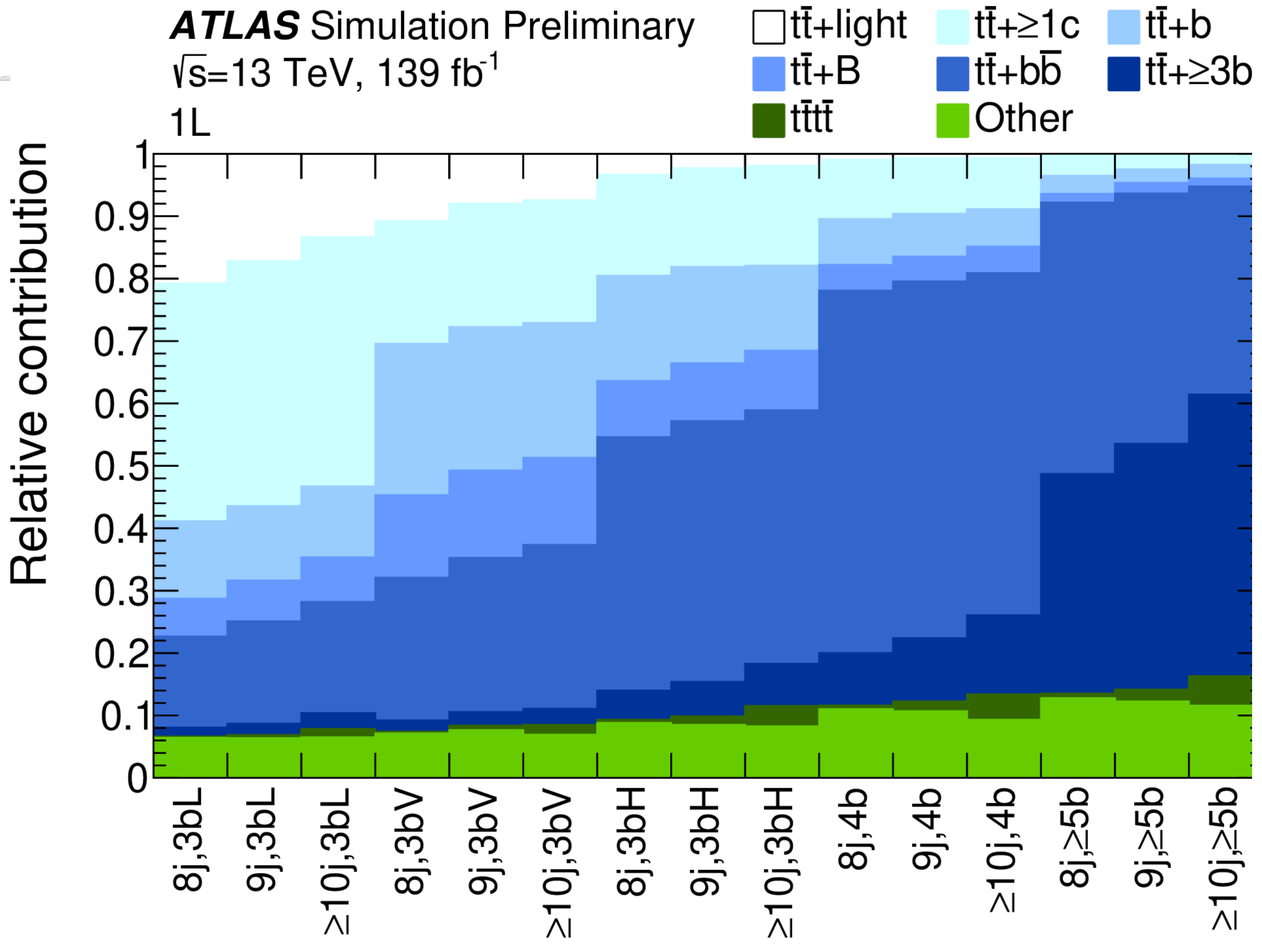
Name	$N_b^{60\%}$	$N_b^{70\%}$	$N_b^{85\%}$
2b	-	= 2	-
3bL	≤ 2	= 3	-
3bH	= 3	= 3	> 3
3bV	= 3	= 3	= 3
≥4b (2LOS)	-	≥ 4	-
4b (1L)	-	= 4	-
≥5b (1L)	-	≥ 5	-



1L [ATLAS-CONF-2024-002]

Background Estimation

- Major background: $t\bar{t} + jets$
- Minor background: $t\bar{t}H, t\bar{t}W, t\bar{t}Z$, single top quark, V(=W,Z)+jets (less than 1% - $t\bar{t}t, t\bar{t}WW, t\bar{t}WZ, tZ$)
- $t\bar{t} + jets$ sub-categorised into $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t} + light$.
- MC prediction of nominal $t\bar{t} + jets$ is under-estimated.
- Data-driven corrections applied in two steps.
  - Heavy flavor (HF) Scaling
  - Neural Network (NN) reweighting



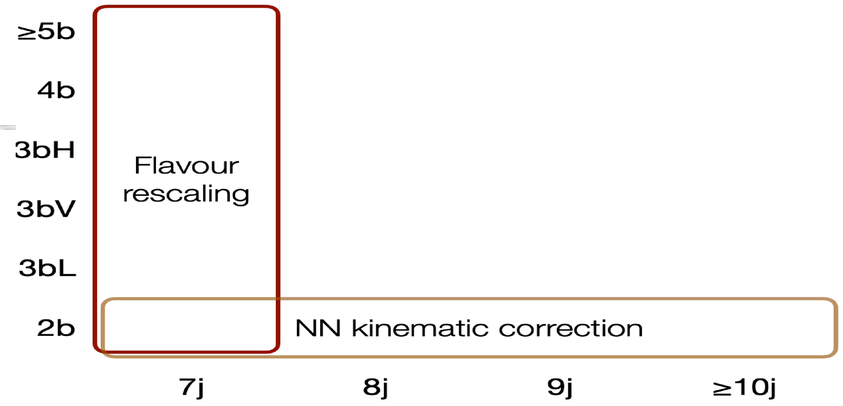
1L [ATLAS-CONF-2024-002]

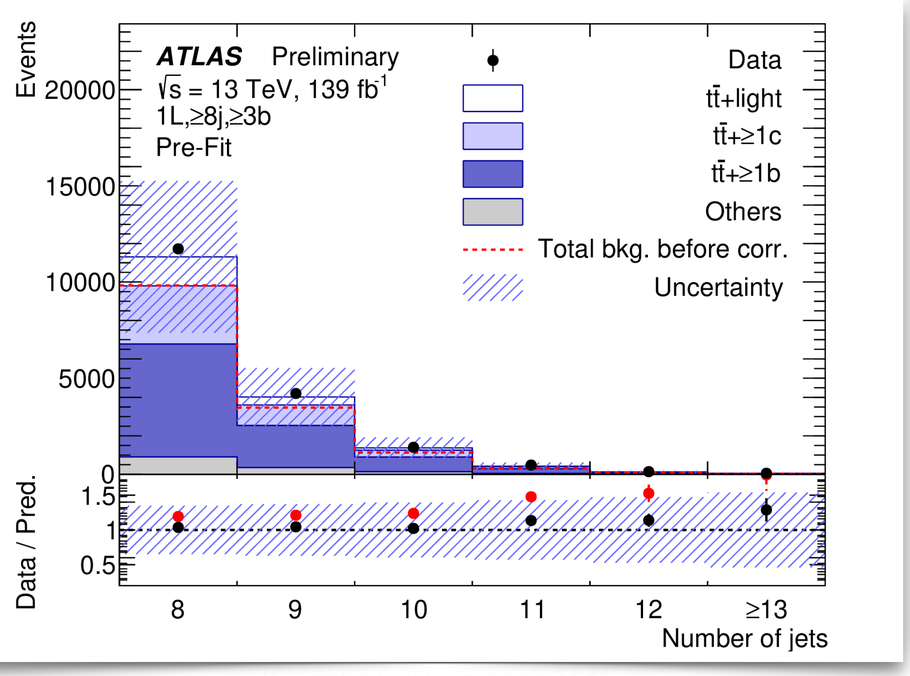
Modelling of $t\bar{t} + jets$

HF Scaling

- A combined profile likelihood fit performed to get the correct normalisation
- The fitted observable is sum of pseudo-continuous b-tagging (pcb) score of the third and foruth jets.
- Scaling factors derived as a ratio of postfit to prefit yields of $t\bar{t}$ samples: $\frac{t\bar{t}^{nominal}_{postfit}}{t\bar{t}^{nominal}_{prefit}}$ and $\frac{t\bar{t}^{nominal}_{postfit}}{t\bar{t}^{alternative}_{prefit}}$, where alternative-prefit is yield obtained before fitting the model parameters to the experimental data for $t\bar{t}$ samples generated using an alternative Monte Carlo generator.

Nominal Samples	Scaling factors
$t\bar{t} + light$ **(1L)**	**0.84 ± 0.04**
$t\bar{t} + light$ **(2LOS)**	**0.87 ± 0.03**
$t\bar{t} + \geq 1c$	**1.61 ± 0.13**
$t\bar{t} + \geq 1b$	**1.18 ± 0.03**





1L [ATLAS-CONF-2024-002]

NN Reweighting

- Neural Network (NN)-based kinematic reweighting [C.P.C 115 (1998)]
- NN output: a-posterior Bayesian probability.

$$o(\mathrm{x}) = P(data\,|\,\mathrm{x}) = \frac{\alpha_{data} P_{data}(\mathrm{x})}{\alpha_{data} P_{data}(\mathrm{x}) + \alpha_{MC} P_{MC}(\mathrm{x})}$$

- Training with $t\bar{t} + jets$ (nominal and alternative) only.
- Input variables: Jets multiplicity (N_j), large-R jets (($N_{LR\text{-}jets}$), all jets & lepton (p_T) , missing E_T
- Reweighting factor: $w(x) = e^{o(x)}$

Multivariate Analysis

- ☑ Indifference between signal and background, needs a Machine Learning algorithm.
- ☑ A mass parameterized message passing Graph Neural Network (GNN) used for signal-background discrimination.
- ☑ Assigns objects (leptons, jets, MET) to graph nodes and their relationships via edges.
- ☑ Training performed separately using events:
  - Jets multiplicity (N_j) >= 9(7) for 1L(2LOS) and b-tagged jets multiplicity ($N_b^{70\%}$) >= 3
  - The output of the GNN is used as the discriminant in the signal regions.

Systematics

- ☑ The table shows grouped impact of all nuisance parameters.
- ☑ $t\bar{t} + jets$ modelling uncertainties have highest impact for all mass points.
- ☑ Among experimental uncertainties, JES and JER dominates.
- ☑ Signal modelling uncertainties have minor impacts.

Uncertainty source	$\Delta\sigma_{t\bar{t}H/A\to t\bar{t}t\bar{t}}$ [fb] $m_{H/A}$=400 GeV		$m_{H/A}$=700 GeV		$m_{H/A}$=1000 GeV	
Signal Modelling						
BSM $t\bar{t}t\bar{t}$ modelling	< 1		+0.1	< 0.1	< 0.1	
Background Modelling						
$t\bar{t}$+$\geq 1b$ modelling	+11	-10	+3.7	-3.4	+1.9	-1.7
SM $t\bar{t}t\bar{t}$ modelling	+3	-3	+2.1	-2.1	+0.9	-0.9
$t\bar{t}$+jets reweighting	+3	-3	+1.0	-1.0	+0.5	-0.5
$t\bar{t}$+$\geq 1c$ modelling	+2	-2	+0.9	-0.8	+0.4	-0.4
$t\bar{t}$+light modelling	+1	-1	+0.2	-0.2	< 0.1	
Other background modelling	< 1		+0.4	-0.4	+0.2	-0.2
Experimental						
Jet energy scale and resolution	+4	-2	+1.3	-0.8	+0.5	-0.3
MC statistical uncertainties	+2	-3	+0.6	-0.7	+0.4	-0.4
b-tagging efficiency and high-p_{T} extrapolation	+2	-1	+0.7	-0.4	+0.4	-0.4
Other uncertainties	< 1		+0.3	-0.5	+0.1	-0.2
Luminosity	< 1		+0.3	-0.1	< 0.1	
Total systematic uncertainty	+13	-12	+4.8	-4.6	+2.5	-2.4
Statistical uncertainty	+6	-6	+3.3	-3.2	+2.3	-2.2
Total uncertainty	+14	-13	+5.6	-5.4	+3.2	-3.0

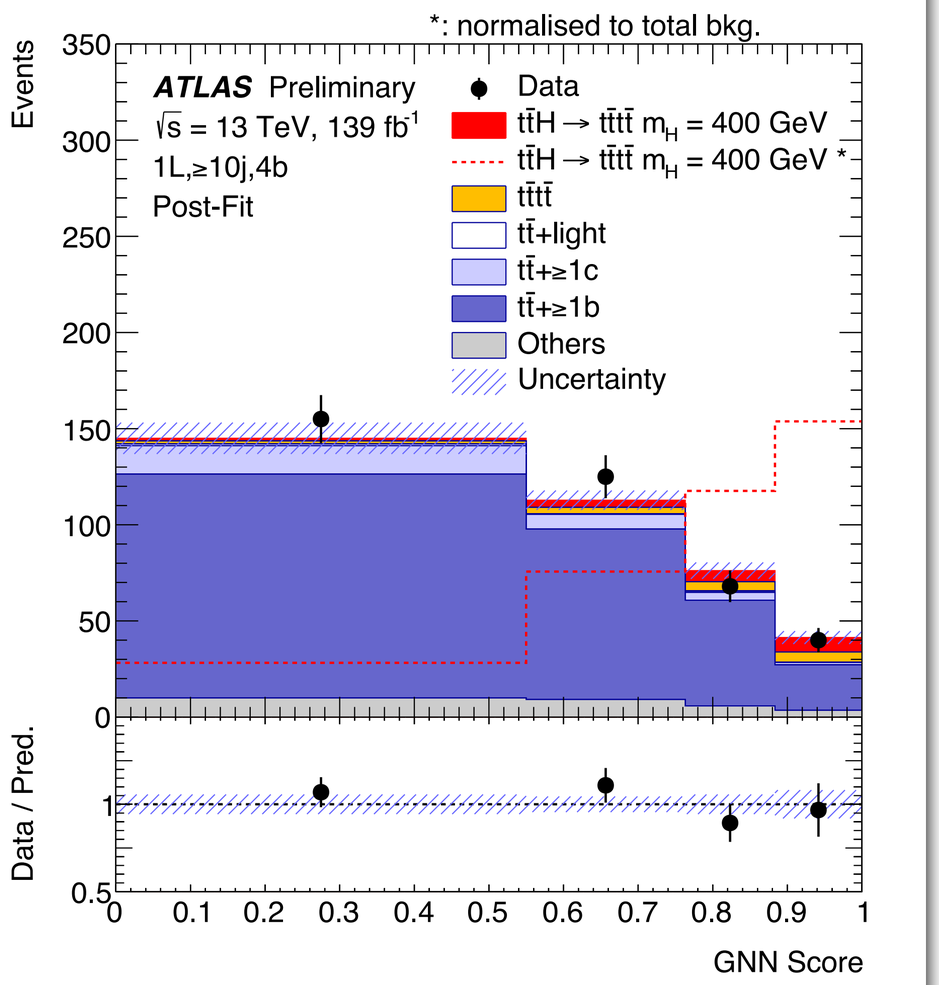
[ATLAS-CONF-2024-002]

Statistical Analysis

- ☑ Binned Profile Likelihood Fit performed simultaneously in control and signal regions for both decay channels.
- ☑ Stacked signal and background Monte Carlo samples as a function of GNN score.
- ☑ $t\bar{t} + jets$ modelling uncertainties dominate the results.
- ☑ Reasonable agreement between data and background.

$$L\left(\vec{n}|\mu,\vec{\theta}\right) = \prod_{r\in\text{region}}\prod_{i\in\text{bin}} \text{Pois}\left(n_{i,r}|\mu S_{i,r}(\vec{\theta}) + B_{i,r}(\vec{\theta})\right) \times \prod_{j\in\text{NP}} G\left(\theta_j\right),$$

Data vector · POI · Poisson distribution · Background yield · Gaussian Distribution · NPs affecting signal & Background · Signal yield · Binned likelihood Function

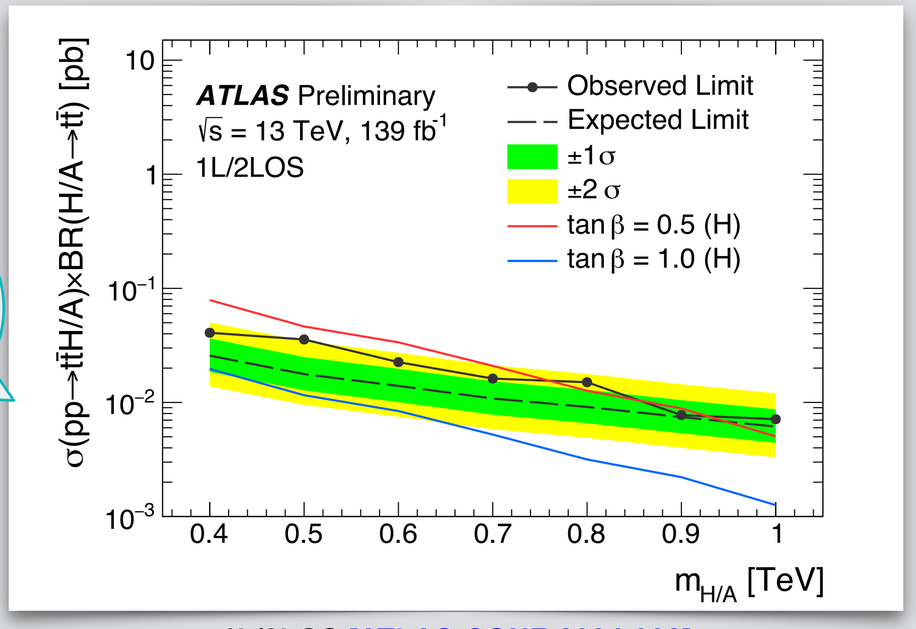


1L [ATLAS-CONF-2024-002]

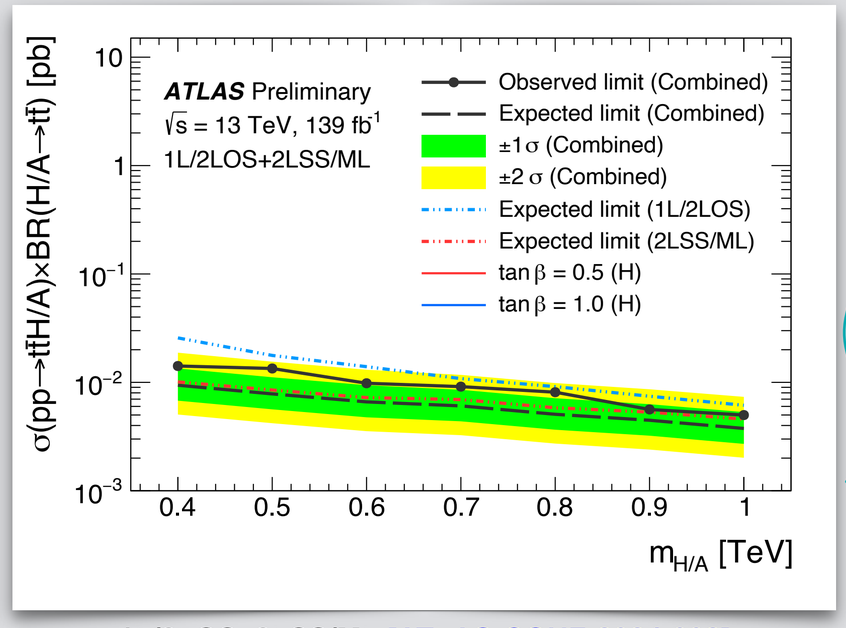
2HDM type-II: Exclusion limits on the cross-section of $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ process

☑95% CLs upper limits on the cross section of the production estimated for 1L/2LOS.

1L/2LOS:
14.2(5.0)fb for heavy Higgs mass of 400(1000) GeV



1L/2LOS [ATLAS-CONF-2024-002]



1L/2LOS+2LSS/ML [ATLAS-CONF-2024-002]

Combined channel:
15.3(5.09)fb for heavy Higgs mass of 400(1000) GeV

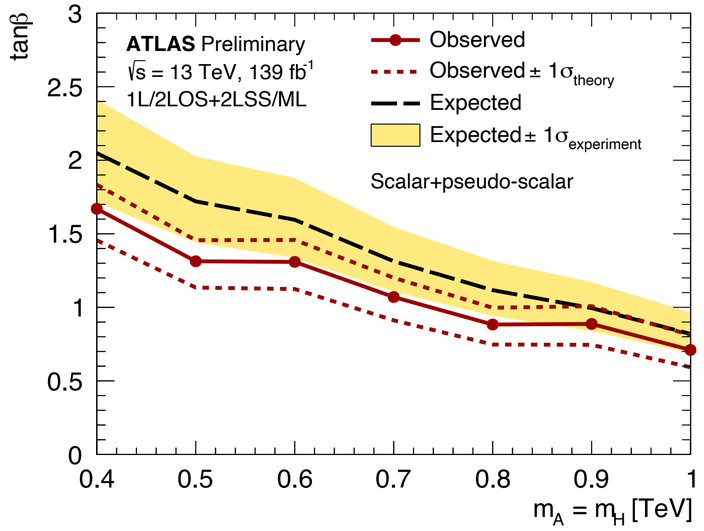
☑Search in 1L/2LOS channel combined with previous search in multileptonic channel (2LSS/ML) [JHEP 07 (2023) 203].

☑Combination performed via a simultaneous profile likelihood fit including all signal and control regions of both the channels.

☑No significant evidence for the heavy Higgs production.

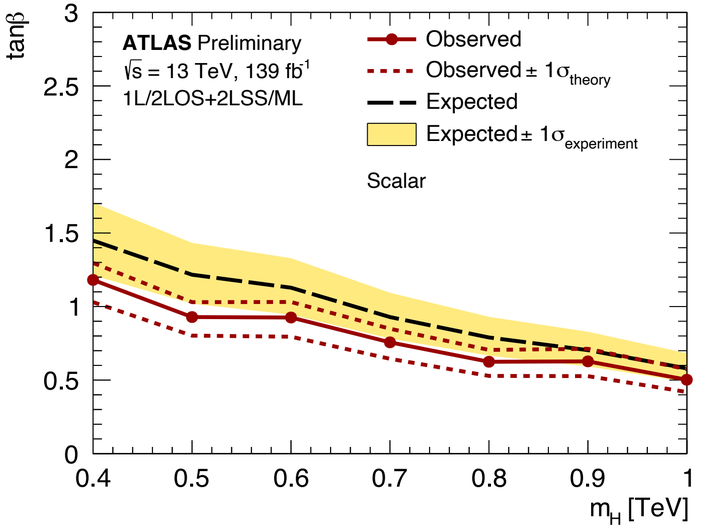
2HDM type-II: Exclusion limits on $tan\beta$

- Interpretation for low $tan\beta$ region in the alignment limit $sin(\beta - \alpha) \rightarrow 1$, where h couplings are similar to the SM Higgs boson and $tan\beta$ is the ratio of the vacuum-expectation-values of the two Higgs doublets.
- 95% CLs lower limits for $tan\beta$ as a function of mass of heavy Higgs for different scenarios in combined channel:
  - *When both scalar and Pseudoscalar contribute to the final state*
  - *When either scalar or Pseudoscalar contribute to the final state*



[ATLAS-CONF-2024-002]

scalar (H) and pseudoscalar (A) contribute:
$tan\beta$ **values below 1.7(0.7) excluded at heavy Higgs mass of 0.4(1.0) TeV**



[ATLAS-CONF-2024-002]

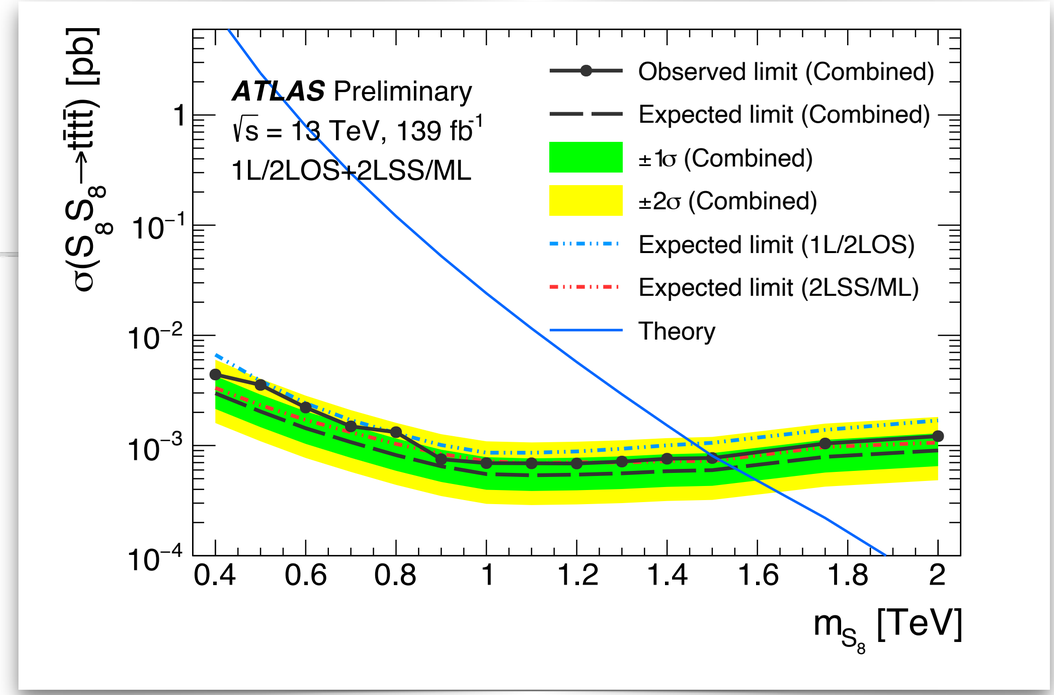
scalar (H) or pseudoscalar (A) contribute:
$tan\beta$ **values below 1.2(0.5) excluded at heavy Higgs mass of 0.4(1.0) TeV**

s-gluon model

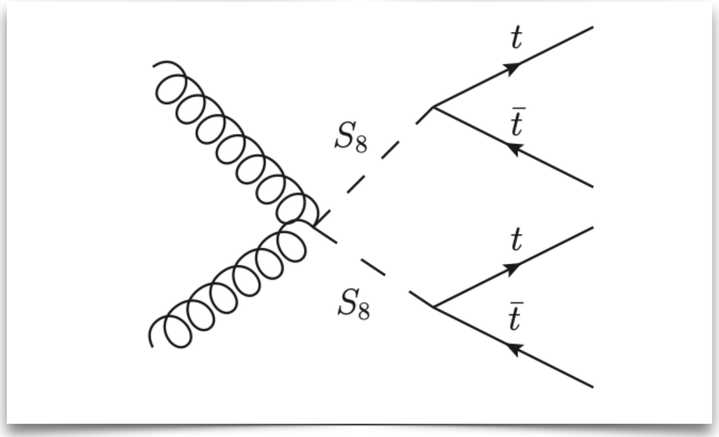
- Reinterpretation of results obtained in the $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ 1LOS channel analysis in the context of s-gluon model [JHEP 09 (2021) 143].
- Signal: $S_8S_8 \rightarrow t\bar{t}t\bar{t}$ with mass of s-gluon $(m_{S_8}) \in [0.4,2.0]$ TeV
- GNN training used: (masses are in TeV)

m_{S_8}	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.75	2.0
GNN	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**

- Studies combined with multilepton channel using same binning, MC background and systematics as in 2HDM.
- 95% CLs upper limits on the cross section of the production estimated. 4.4(0.7) fb excluded at m_{S_8} of 0.4(1.0) TeV
- S-gluon masses m_{S_8} are excluded below 1.5 TeV.
- No significant evidence for the sgluon production.



[ATLAS-CONF-2024-002]



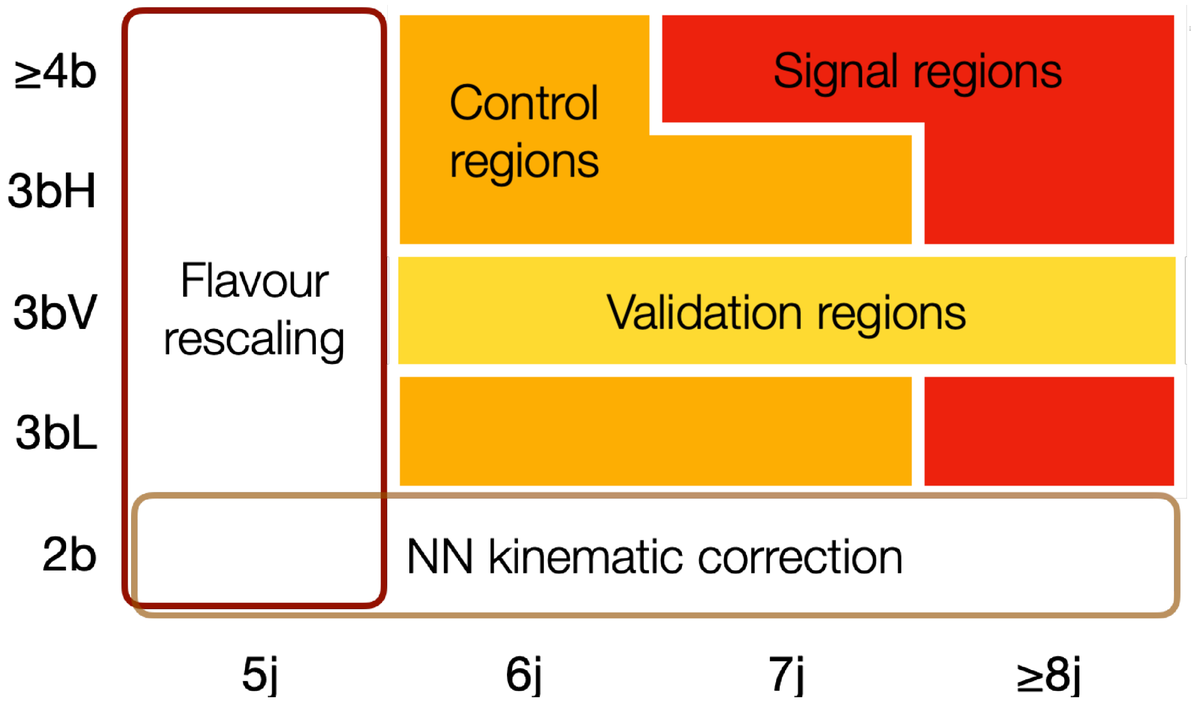
Summary

- ☑ A search for heavy scalar or pseudoscalar higgs in $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ 1LOS channel in the context of 2HDM performed.
  - Data-driven corrections applied to most dominant background: $t\bar{t} + jets$
  - A mass-parameterised GNN used, to optimise signal-background discrimination.
  - No significant excess of events above the SM prediction is observed.
  - Combination with SSML channel
    - ☐ Excluded $tan\beta$ values below 1.7(0.7) when H+A contribute and 1.2(0.5) when H or A contribute, at 0.4(1.0) TeV.
- ☑ Reinterpretation of results in the context of **sgluon model**.
  - Mass signals $m_{S_8} < 1.5$ TeV excluded.

Thank you!

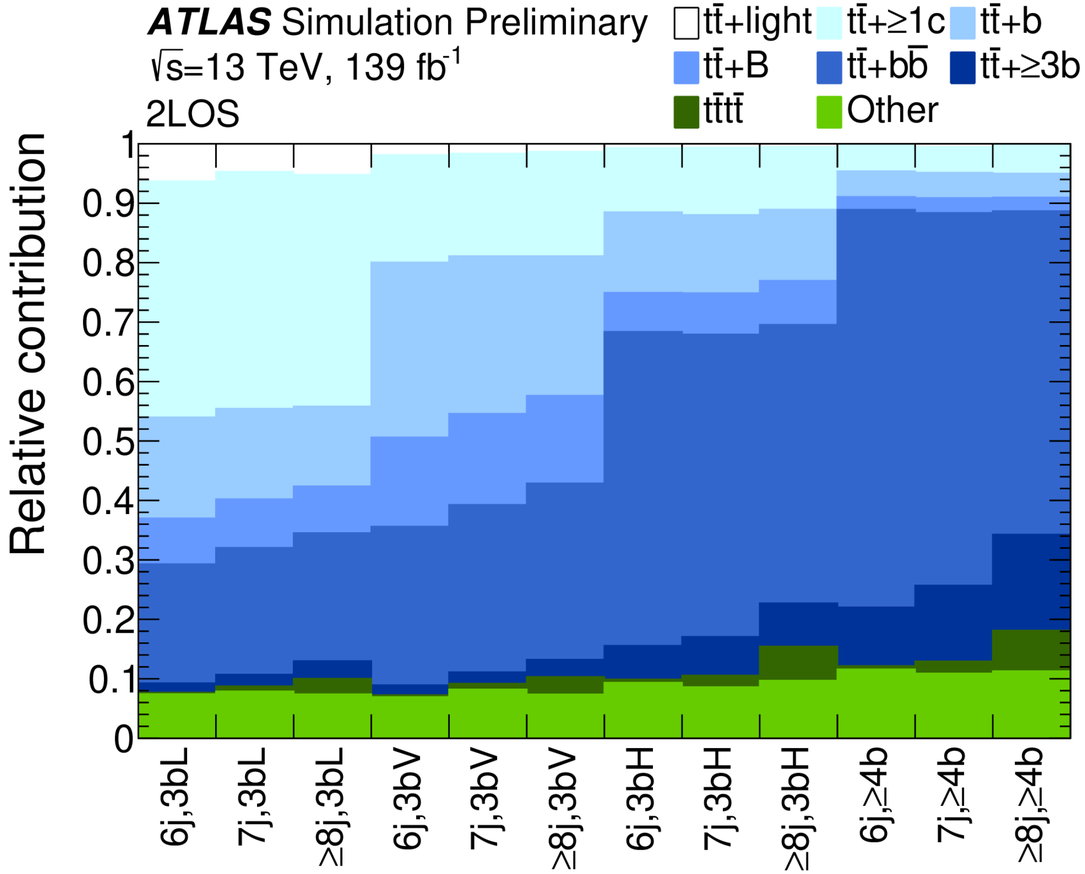
Backup slides

Analysis Regions : 2LOS



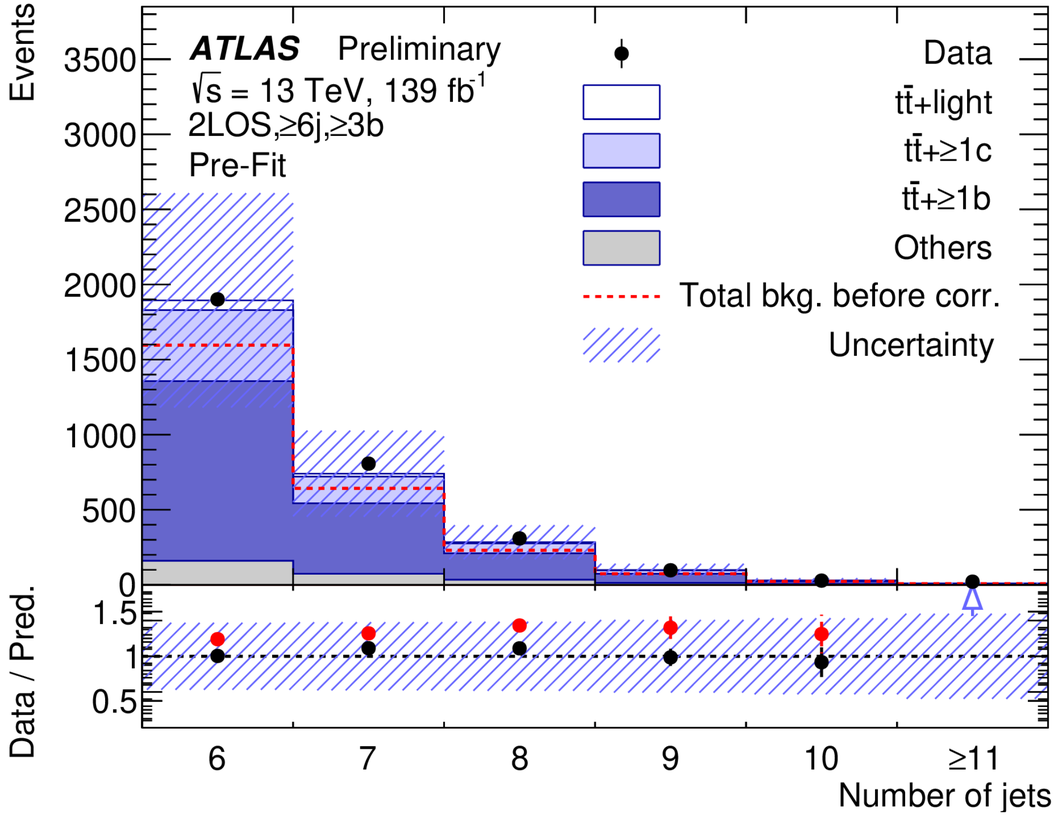
[ATLAS-CONF-2024-002]

Background Estimation: 2LOS



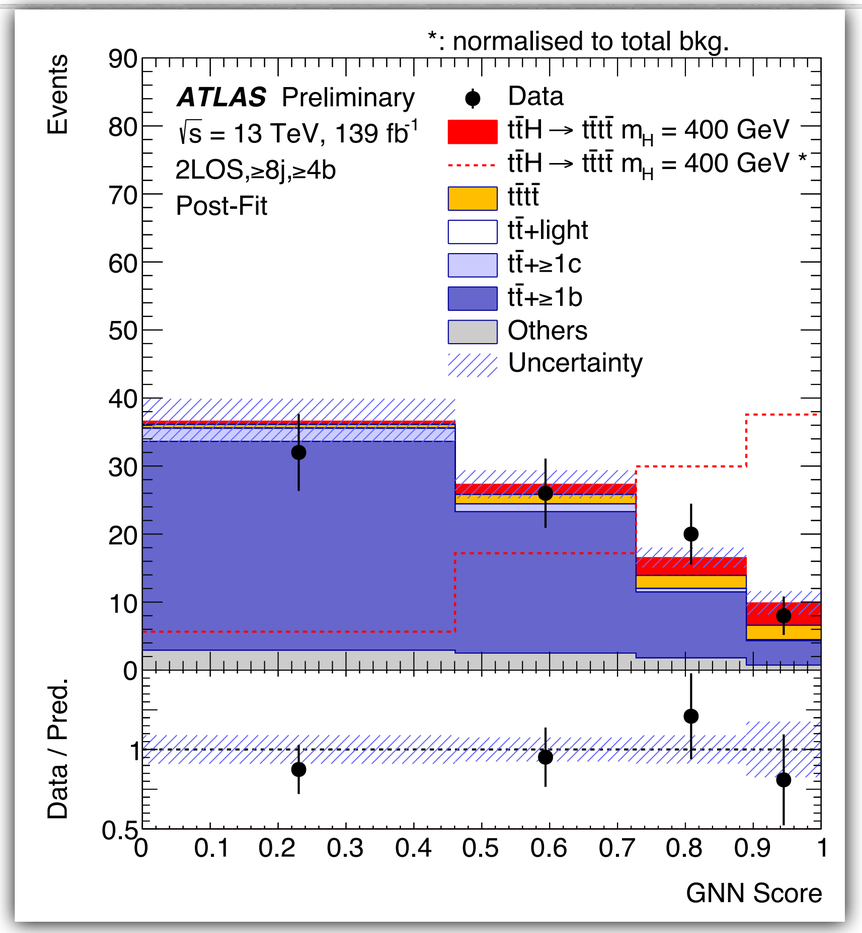
[ATLAS-CONF-2024-002]

Background estimation: NN reweighting (2LOS)



[ATLAS-CONF-2024-002]

Statistical Analysis: (2LOS)



[ATLAS-CONF-2024-002]

Background estimation: $t\bar{t} + jets$ truth classification

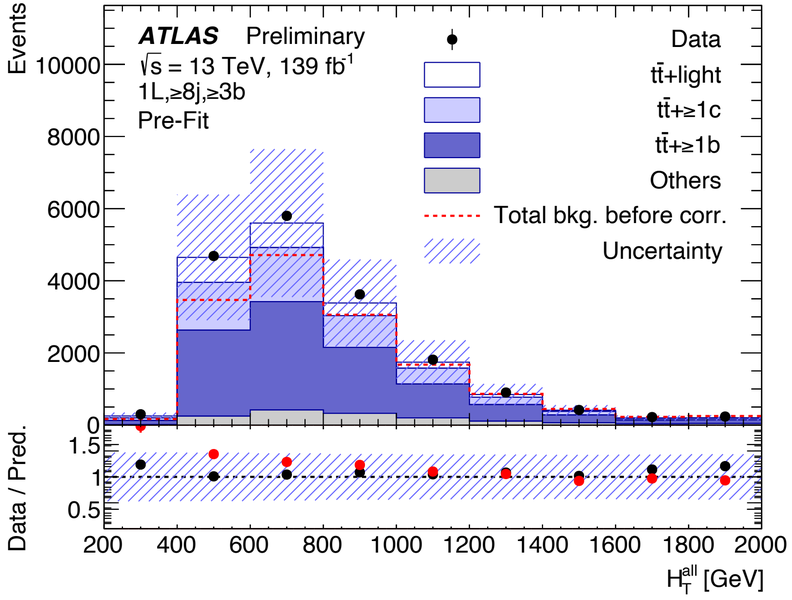
- ☑ Only applied for jets not from t-quark decay.
- ☑ Particle jets are reconstructed from stable truth particle by anti-kt.
  - ☑ R = 0.4
  - ☑ $p_T > 15$ GeV
  - ☑ $|\eta| < 2.5$
- ☑ Hadron-jet Matching:

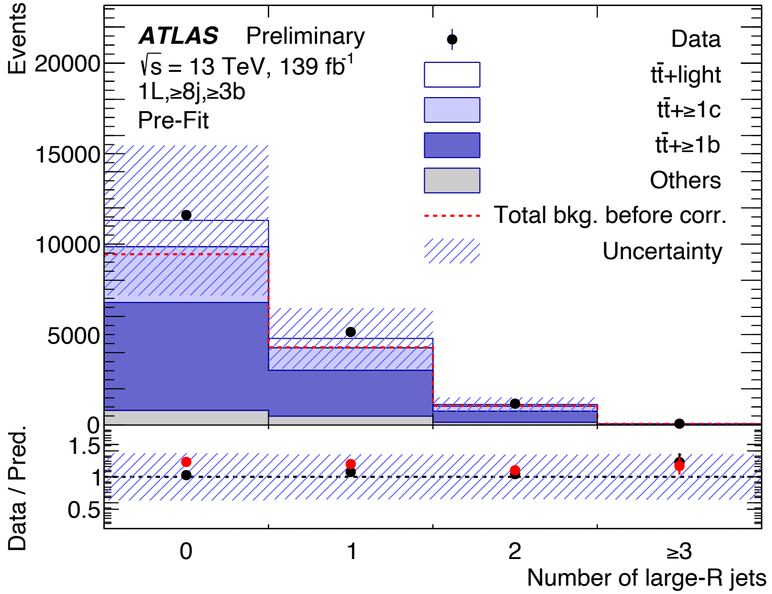
jets	contains hadrons	ΔR	$p_T(GeV)$	decay from
b-jets	b-hadron(s)	<0.3	>5	not t-quark
c-jets	c-hadron(s)	<0.3	>5	not t-quark or W-boson

Background Estimation: NN Reweighting

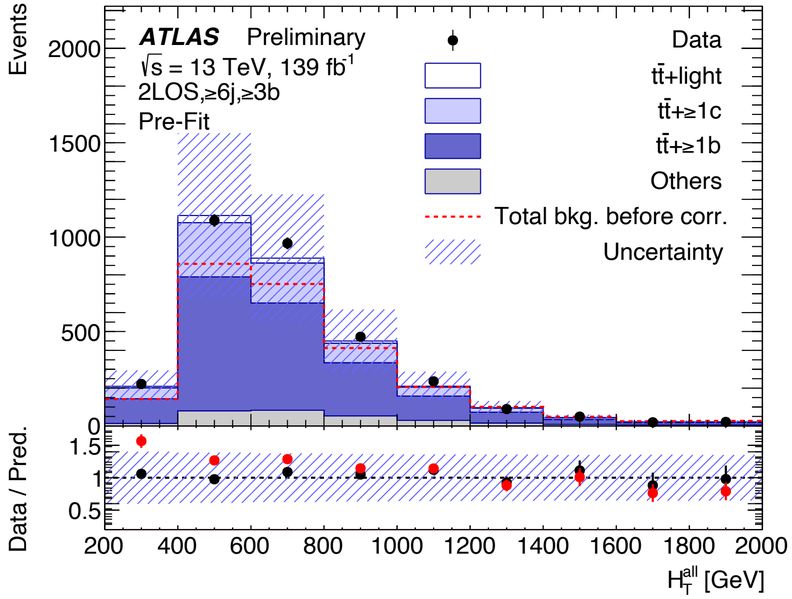
[ATLAS-CONF-2024-002]

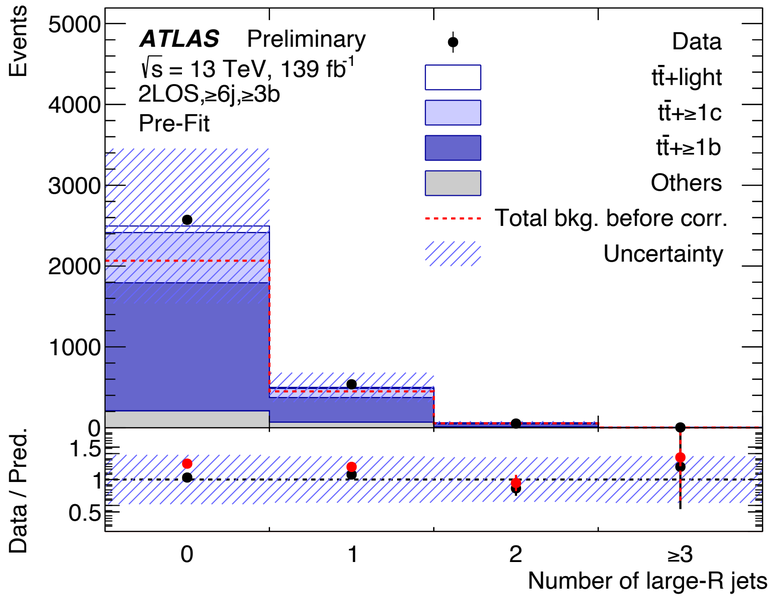
1L





2LOS





List of global features in GNN training

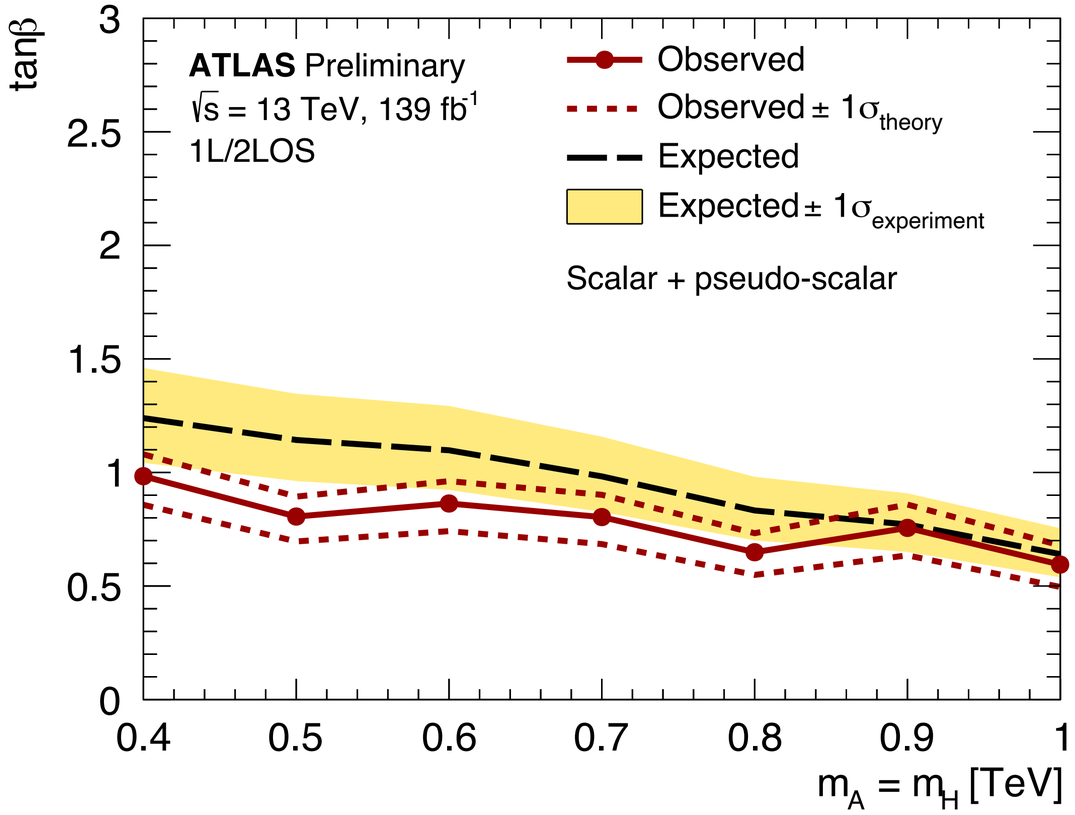
[ATLAS-CONF-2024-002]

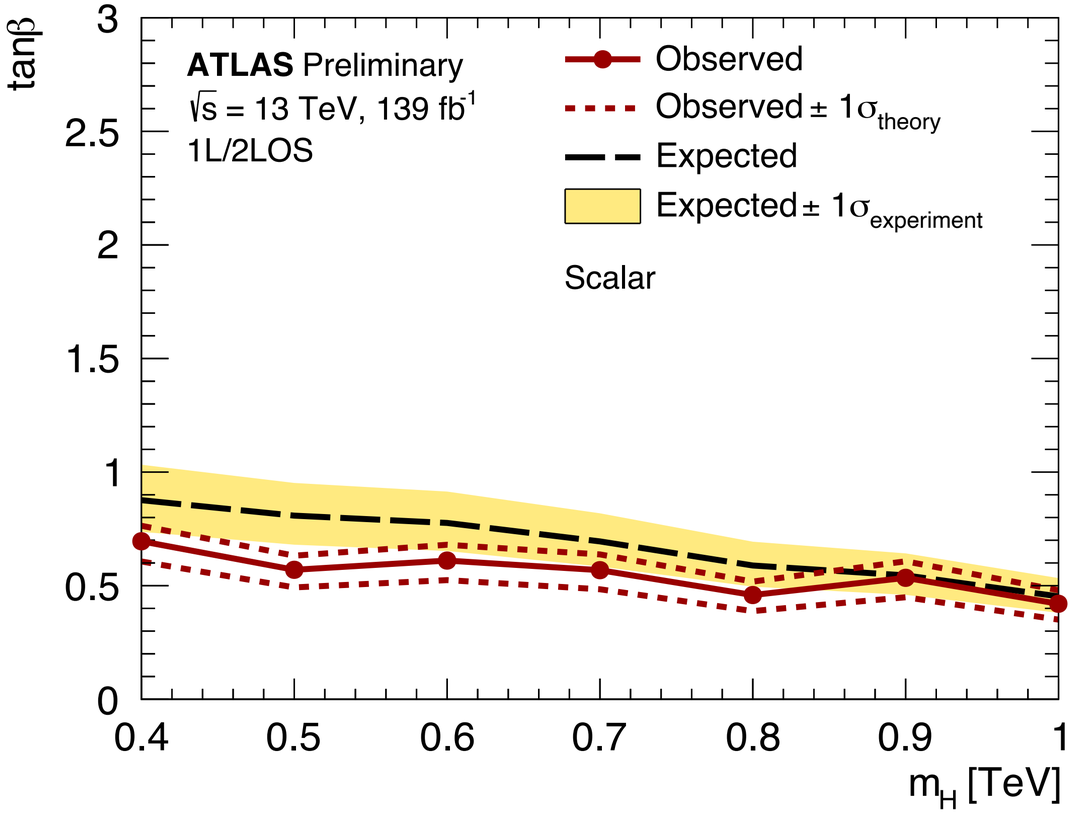
Node	Edge	Global
Object p_T, η, E, b-tagging score, Object type encoding number	$\Delta\eta$, $\Delta\phi$, ΔR between pairs of objects	H_T, m_{ll} for 2LOS and m_T for 1L,, N_{jets}, $NRCjets_{m>100}$, M_{bbb}^{avg}, ΔR_{bb}^{min}, ΔR_{bl}^{min}, $\frac{\sum_i p_{Ti}}{\sum_i E_i}$, Sum of pcb for the first 6 jets $\sum_{i<6} pcb_i$, $\sum d_{12}$, $\sum d_{23}$, $\frac{\sum_{i=0}^{3} p_{Ti}}{\sum_{i\geq 4} p_{Ti}}$

Variable	Description
$\sum_{i\in[1,6]} \text{pcb}_i$	Sum of the pcb scores of the six jets with the highest scores
H_{T}	p_{T} sum of all reconstructed leptons and jets
N_{jets}	Jet multiplicities
$H_{\text{T}}^{\text{ratio}}$	p_{T} sum of the four leading jets in p_{T} divided by the p_{T} sum of the remaining jets
$dR_{jj}^{\text{avg.}}$	Average ΔR across all jet pairs
m_{T}^{W}	W-boson transverse mass calculated using the lepton four-momenta and $E_{\text{T}}^{\text{miss}}$ (1L only)
$\Delta R_{bb}^{\text{min.}}$	Minimum ΔR between any pair of jets b-tagged at the 70% OP
$\Delta R_{\ell b}^{\text{min.}}$	Minimum ΔR between any pair of lepton and jet b-tagged at the 70% OP
$m_{bbb}^{\text{avg.}}$	Average invariant mass of all triplets of jets b-tagged at the 70% OP
$m_{jjj}^{\text{avg.}}$	Average invariant mass of all jet-triplets with an angular separation of $\Delta R < 3$
$\sum d_{12}$	Sum of the first k_t splitting scale d_{12} over all large-R jets
$\sum d_{23}$	Sum of the second k_t splitting scale d_{12} over all large-R jets
$N_{\text{LR-jets}}$	Number of large-R jets with a mass greater than 100 GeV
Centrality	$\sum_i p_{\text{T}}^i / \sum_i E_i$ where the sums are performed over all reconstructed jets and leptons
$m_{\ell\ell}$	Invariant mass of the two leptons (2LOS only)

2HDM: Exclusion limits on $tan\beta$ (1L/2LOS)

[ATLAS-CONF-2024-002]





scalar (H) and pseudoscalar (A) contribute

scalar (H) or pseudoscalar (A) contribute