Search for $t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}$ production in the final state with one or two opposite-sign leptons using the full Run 2 *pp* collisions data at $\sqrt{s} = 13TeV$

Lining Mao

Shanghai Jiao Tong University

November 18th, 2023

Motivation

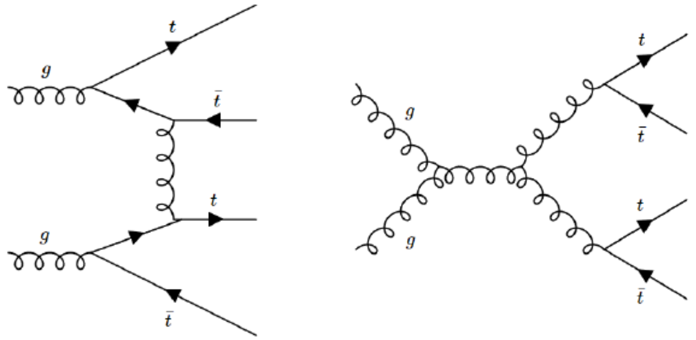
Theoretical

- Problems with the Standard Model (SM):
  - Naturalness problem, hierarchy problem, neutrinos masses…
- Two-Higgs-Doublets-Models (2HDM) as a typical BSM scenario:
  - Simplest model to extend the number of Higgs boson

$$\mathcal{L}_{\text{2HDM}}^{\text{scalar}} = \sum_{i=1,2} (D_\mu \Phi_i)^\dagger (D^\mu \Phi_i) - \boxed{V(\Phi_1, \Phi_2)}$$

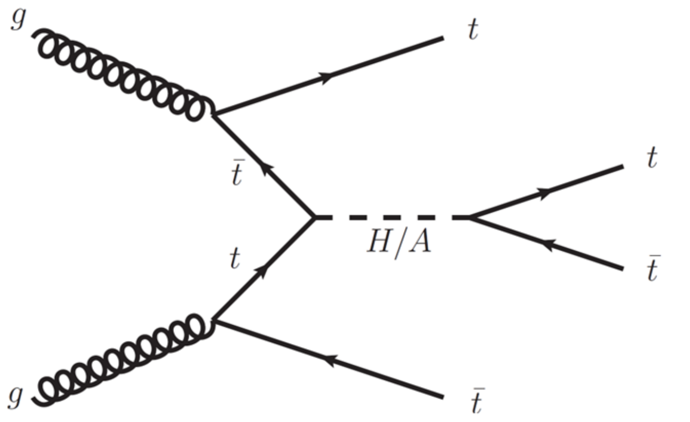
Experimental

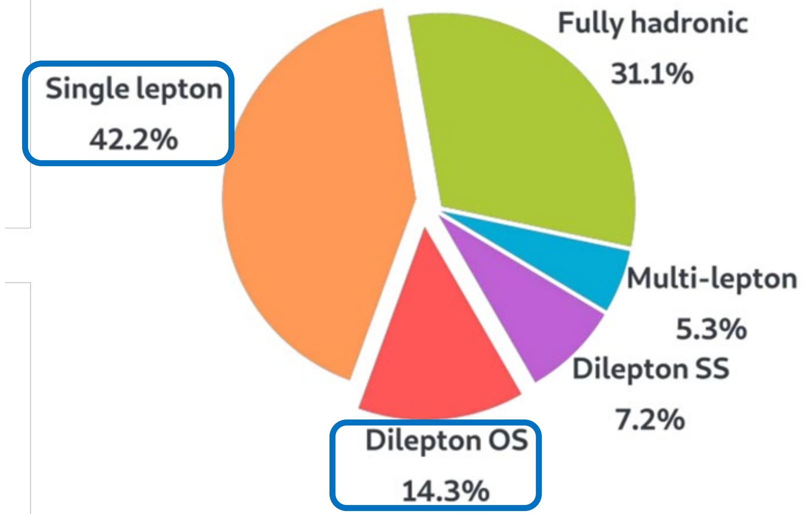
- Standard model 4-tops process
  - SM cross section expectation: $\sigma_{t\bar{t}t\bar{t}} = 13.37^{+1.04}_{-1.78} fb$
  - Latest measurement in ATLAS: $\sigma_{t\bar{t}t\bar{t}} = 22.5^{+6.6}_{-5.5} fb$
  - Inconsistence arises between the theory and the experiment



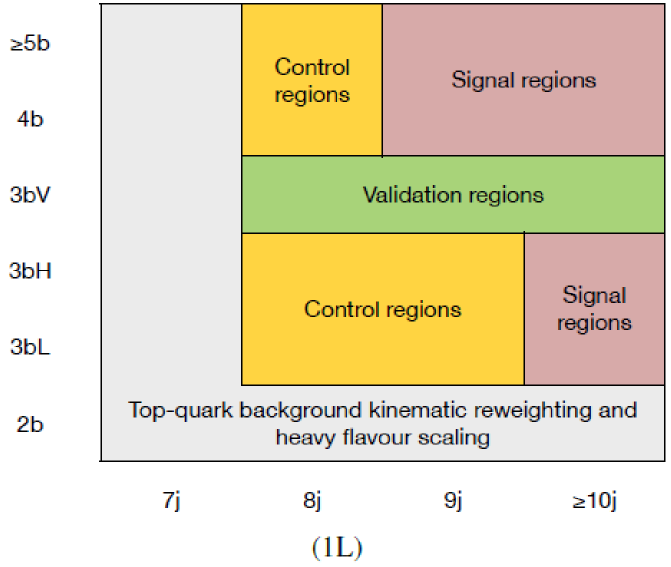
Introduction

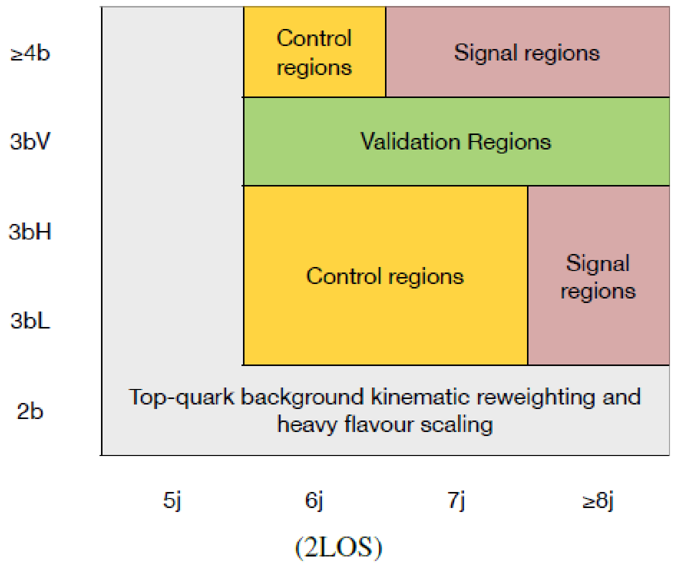
- **Aim**: Search for Heavy Higgs by 4tops production with the one-lepton/di-lepton opposite-sign final states
  - **Higgs mass**: 400GeV to 1000GeV with 100GeV granularity
  - **Main background**: $t\bar{t} + jets$
  - **Final state signature:** High jet & b-jet multiplicity
- Published for the same search in the di-lepton same-sign/multi-lepton channel (SSML channel) (link)





Analysis Strategy





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	-

- Two types of data-driven correction factors:
  - Heavy flavor (HF) normalization factors
  - Neural network (NN)-based kinematic reweighting
- Trained H/A-mass-parameterized GNN to separate signals from background
- Profile likelihood fit in all CRs and SRs simultaneously, using $\mathbf{H_T}$ in CR and **GNN-score** in SR

Data & MC Samples and Event Selection

- Data: Full Run2 (2015-2018) ~ 139fb^{-1}, using single lepton triggers

- MC samples:
  - $t\bar{t}$ + jets, SM4t, ttW, ttZ, Single top, V + jets, ttt, Other top, VV, Signal

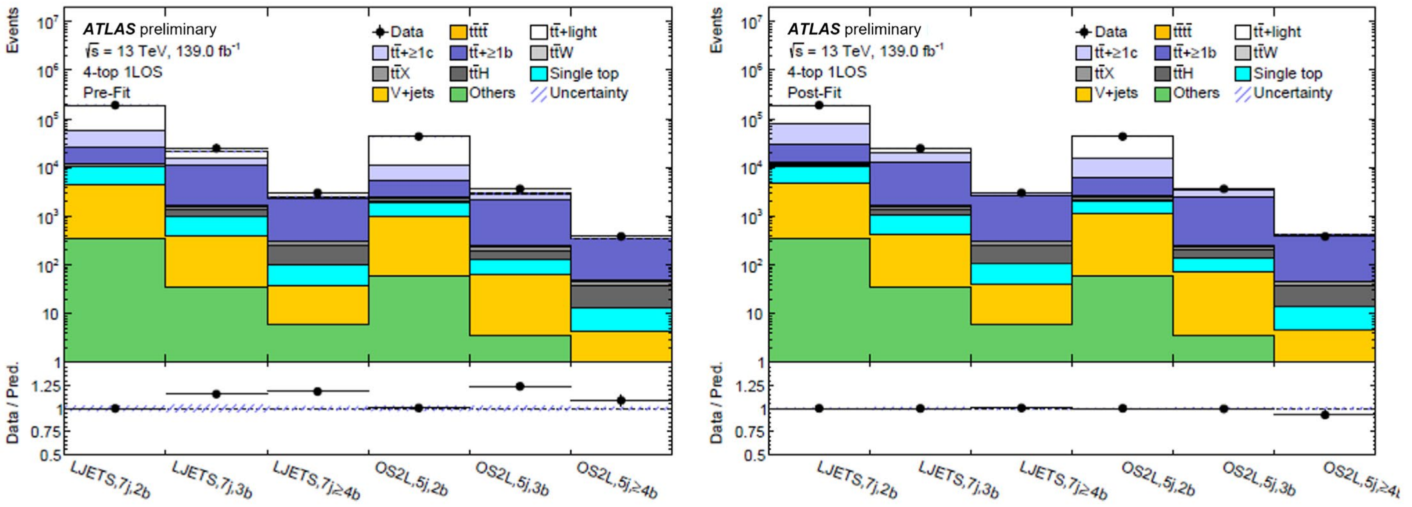
- Object definition & event selection:

Object	Baseline selection
Lepton	p_T>28GeV, $\lvert\eta\rvert$<1.37or1.52-2.47(e), $\lvert\eta\rvert$<2.5(μ) Identification: TightLH(e)/Medium(μ), Isolation: FCTight(e)/FCTightTrackOnly(μ) Impact parameter: z_0<0.5mm, σ_{d0}<5(3) for e(μ)
Jet	p_T>25GeV, $\lvert\eta\rvert$<2.5, JVT>0.5 for p_T < 60 GeV, $\lvert\eta\rvert$<2.4 Algorithm: Anti-k_T
b-jet	p_T>25GeV, $\lvert\eta\rvert$<2.5, JVT>0.5 for p_T < 60 GeV, $\lvert\eta\rvert$<2.4 Algorithm: DL1r
Event	Exactly one lepton (1L) / two opposite-charge leptons (2L) with ≥ 2 jets with b-tagging passing 70% OP

Background Modelling (HF normalization)

- To correct the under-estimation of the $t\bar{t}$ + HF production rate in MC prediction
- $t\bar{t}$ + HF includes:
  - $t\bar{t}$ + ≥1b (TTB): $t\bar{t}$ + at least one jet matched with b-hadron(s)
  - $t\bar{t}$ + ≥1c (TTC): $t\bar{t}$ + at least one jet matched with c-hadron(s)
  - $t\bar{t}$ + light (TTL): $t\bar{t}$ + jets do not match with b or c-hadrons
- Data-driven fit to sum of pseudo-continuous b-tagging scores of 3rd and 4th jets
- Regions:7j(5j) = 2b ,=3b, ≥4b for 1L(2LOS) channel
- Scale factor $\frac{t\bar{t}^{nominal}_{postfit}}{t\bar{t}^{nominal}_{prefit}}$ & $\frac{t\bar{t}^{nominal}_{postfit}}{t\bar{t}^{alternative}_{prefit}}$:

	TTL (1L)	TTL (OS)	TTC	TTB
Nominal	0.84 ± 0.04	0.87 ± 0.03	1.61 ± 0.13	1.18 ± 0.03
ttbb 4FS	0.83 ± 0.04	0.87 ± 0.04	1.60 ± 0.10	1.17 ± 0.02
aMcAtNloPy8	0.94 ± 0.04	0.96 ± 0.04	1.78 ± 0.11	1.27 ± 0.01
PhHerwig	0.66 ± 0.03	0.73 ± 0.03	2.21 ± 0.14	1.56 ± 0.02



Pre-fit & post-fit yields for HF normalization

Background Modelling (NN-reweighting)

- Kinematic reweighting based on Neural Network (NN)
  - NN output: a-posterior Bayesian probability

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

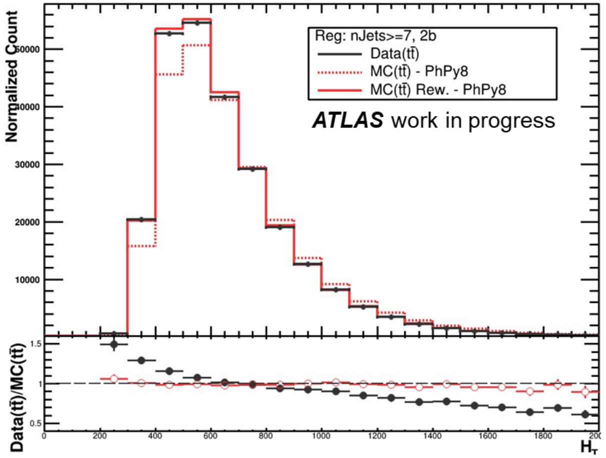
  - Input list: Njets, NRCjets, each jets & lep pT, missing ET
  - Training regions: ≥7j(5j), = 2b for 1L(2LOS)
- Using an exponential lose function:

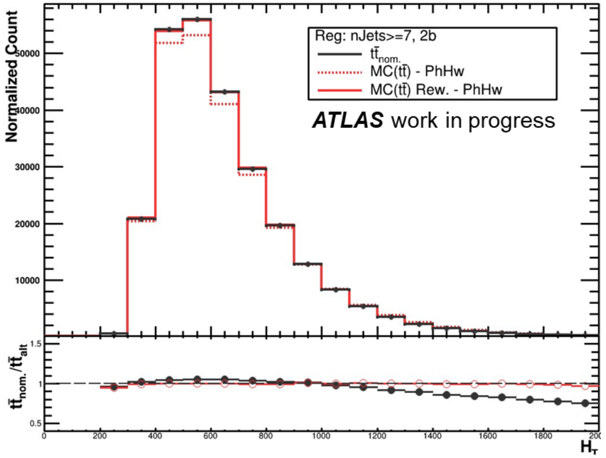
$$\mathcal{L} = P_{data} e^{-o(\boldsymbol{x})/2} + P_{MC} e^{o(\boldsymbol{x})/2} \quad \frac{d\mathcal{L}}{do(\boldsymbol{x})}$$

$$\mathcal{L} = 0 \Rightarrow -\frac{P_{data}}{2} e^{-\frac{o(\boldsymbol{x})}{2}} + \frac{P_{MC}}{2} e^{\frac{o(\boldsymbol{x})}{2}} = 0$$

  - Reweighting factor can be derived as:

$$w(x) = \frac{P_{data}}{P_{MC}} = e^{o(\boldsymbol{x})}$$

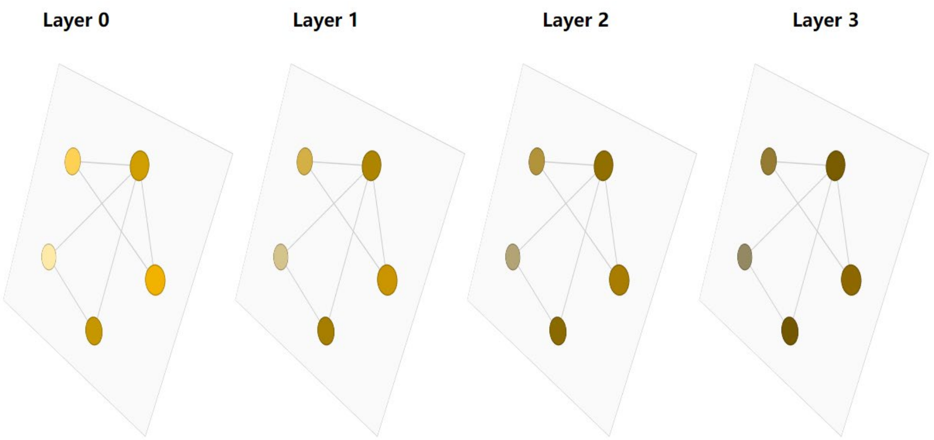




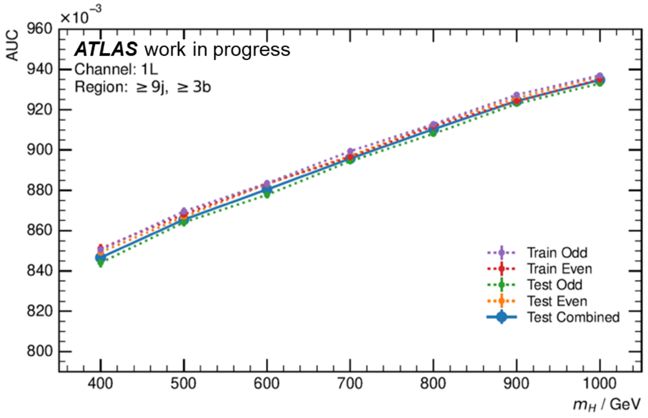
NN-reweighting effect of nominal sample (up) & alternative sample (down)

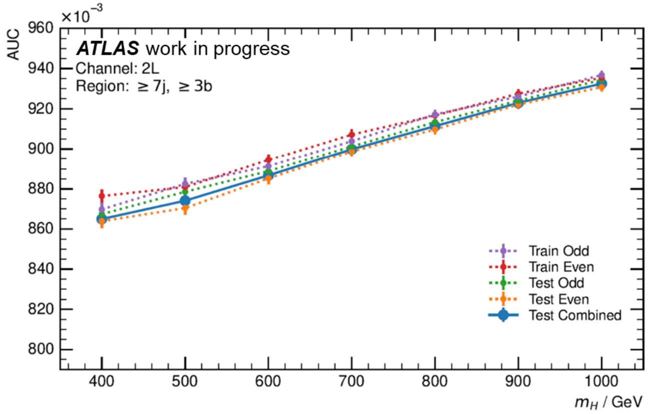
Signal Background Discrimination (GNN)

- For signal discrimination, GNN (graph neural network) has been used
  - GNN is agnostic to the number of nodes and are permutation invariant.
  - A relatively simple model can be used on events of varying multiplicity/topologies
  - Well suitable in our case with complex jet & b-jet multiplicity and structure



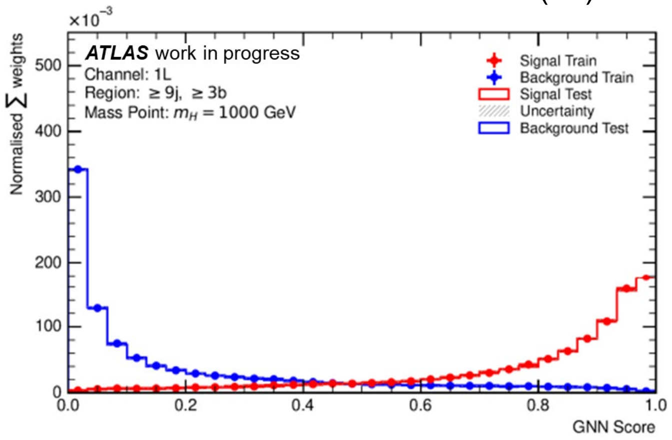
Simple sketch of node information accumulation through the network

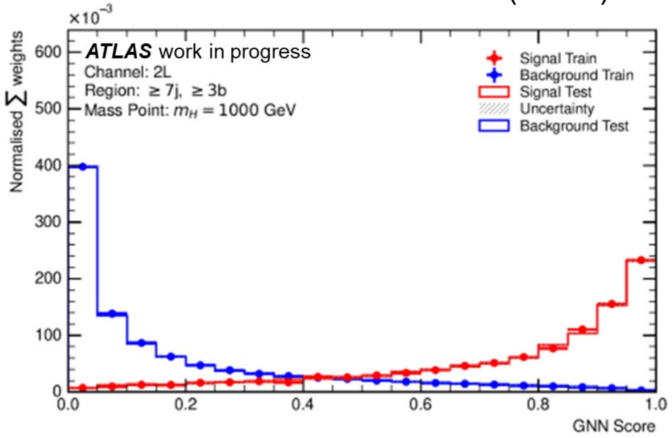




AUC curve & score of GNN (1L)

AUC curve & score of GNN (2LOS)





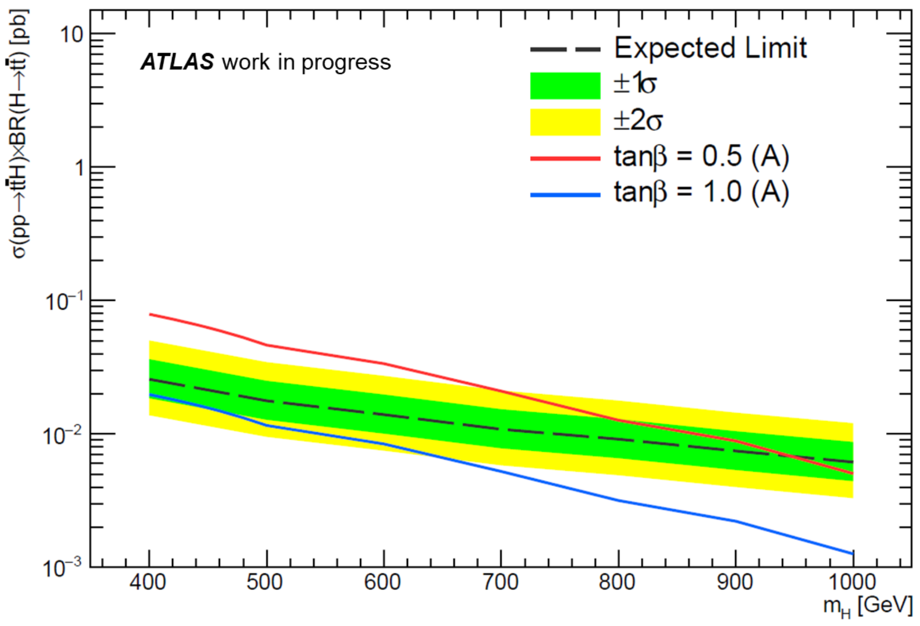
Systematics

- **$t\bar{t}$ theoretical uncertainties:**
  - $t\bar{t}$ modelling: μF/μR/ISR/FSR (decorrelated to $t\bar{t}$ + ≥ 1b, $t\bar{t}$ + ≥ 1c and $t\bar{t}$ + light based on $t\bar{t}$ event classification)
  - $t\bar{t}$ MC choice (decorrelated to $t\bar{t}$ + b, $t\bar{t}$ + bb, $t\bar{t}$ + bbb, $t\bar{t}$ + B, $t\bar{t}$ + ≥ 1c and $t\bar{t}$ + light based on $t\bar{t}$ event classification)
    - ME: (PhPy8 as nominal v.s Mg5Py8 as alternative)
    - PS: (PhPy8 as nominal v.s PhHw as alternative)
    - FS: (ttbb 5FS as nominal v.s ttbb 4FS as alternative)
  - $t\bar{t}$ + ≥ 1b and $t\bar{t}$ + ≥ 1c XS (50%), $t\bar{t}$ + light XS (5%, decorrelated to 1L & 2LOS channels)
  - NN-reweighting
    - Statistic uncertainties extrapolation
    - Non-$t\bar{t}$ subtraction
- **Non-$t\bar{t}$ theoretical uncertainties:**
  - Modelling & cross section
- **SM $t\bar{t}t\bar{t}$:**
  - Generator, PS, scale, XS (+7.8%, -13.3%)
- **Experimental uncertainties:**
  - Luminosity, pileup reweighting, object reconstruction, jet tagging, etc.
- **Signal modelling:**
  - PDF (1%)
  - Renormalization and factorization scales

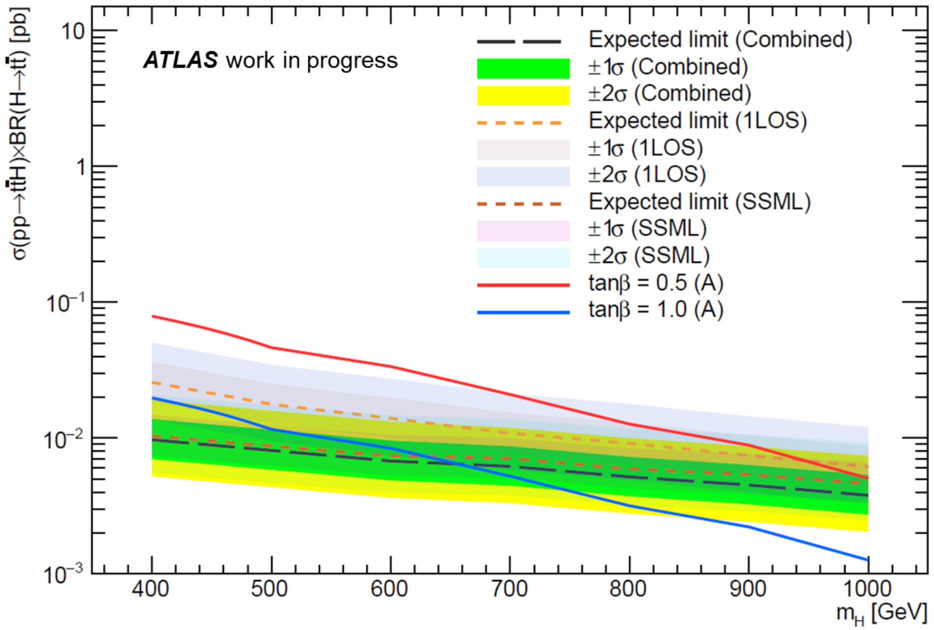
Expected Fit Results

- Expected upper limit of the cross section extracted from Asimov data fit
  - Ranged between 25.7 fb to 6.1 fb for heavy Higgs mass between 400 GeV and 1000 GeV, respectively

- Combination with SSML channel
  - Ranged between 9.7 fb to 4.6 fb for heavy Higgs mass between 400 GeV and 1000 GeV, respectively



Expected upper limit of 1LOS channel



Expected upper limit of 1LOS + SSML channel

Summary

- Heavy Higgs production in the 4-top process has been searched using ATLAS full run-2 data
- Background correction with 2 data-driven factors
  - HF normalization
  - NN-reweighting
- Signal-background discriminated by GNN
- Systematics has been robustly studied
- Expected upper limit of the cross section:

	400 GeV	**1000 GeV**
1LOS	25.7fb	6.1fb
Combined	9.7fb	4.6fb

- Under collaboration review

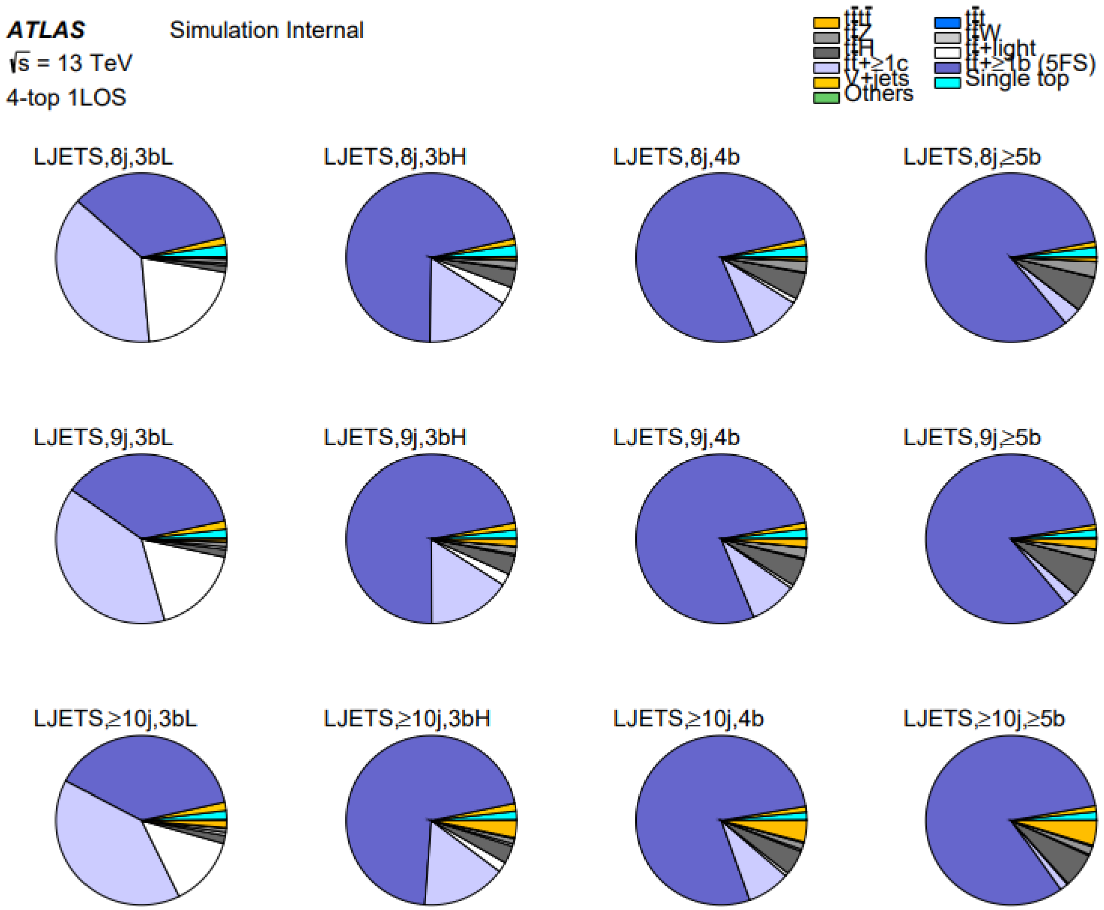
BACKUP

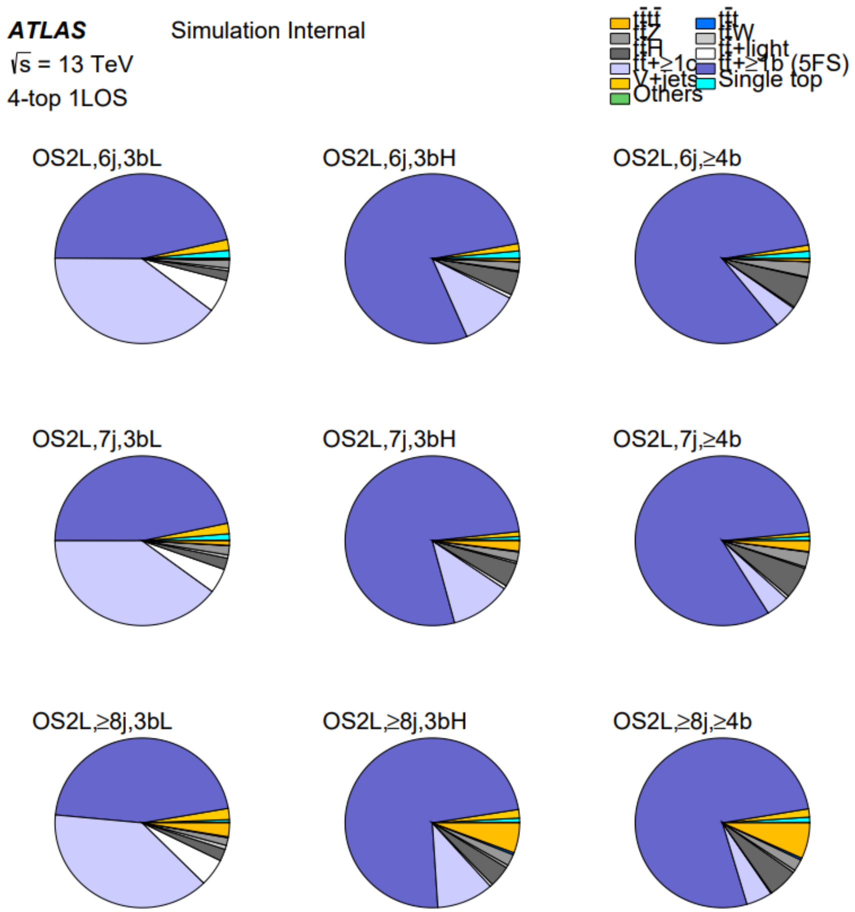
Samples

- Recent p-tag from the TOPQ1 has been used for over all MC/Data samples.
- AnalysisBase 21.2.169.
- Framework: TTHbb analysis (Link).
- Only prompt SM processes modelled with MC for background:

	Nominal	Alternative
$t\bar{t}$ + jets	PhPy8 (5FS ttbb, inclusive + HF filtered +HT sliced)	Mg5Py8, PhHw, PhPy8 (4FS ttbb)
SM4t	aMcAtNloPy8	aMcAtNloHerwig7, Sherpa
ttW	Sherpa	aMcAtNloPy8
ttZ	aMcAtNloPy8	Sherpa
Single top	PhPy8	PhH7, aMcAtNloPy8
V + jets	Sherpa	-
ttt	Mg5Py8	-
Other top, VV	MgPy8, aMcAtNloPy9	-
Signal	Mg5Py8	-

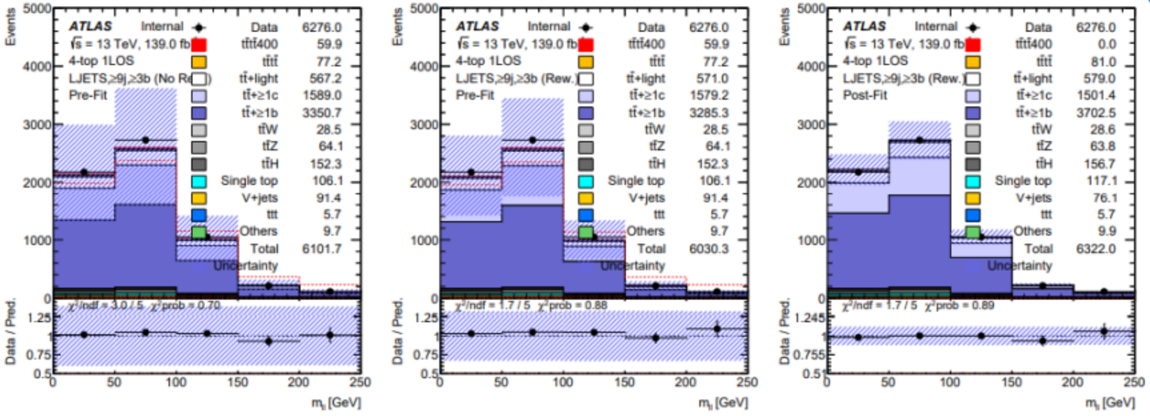
Background Composition



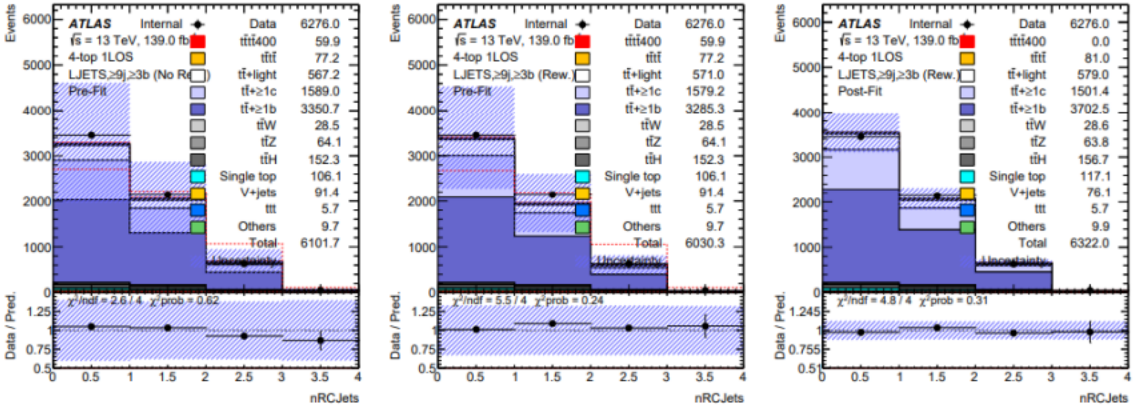


Performance of NN-reweighting

- Added several Data/MC comparison plots for the variables used in the MVA training (Appendix K.2)
  - Pre-fit plots without reweighting (left), pre-fit plots with reweighting (middle), post-fit plots (right)



(a) mtw



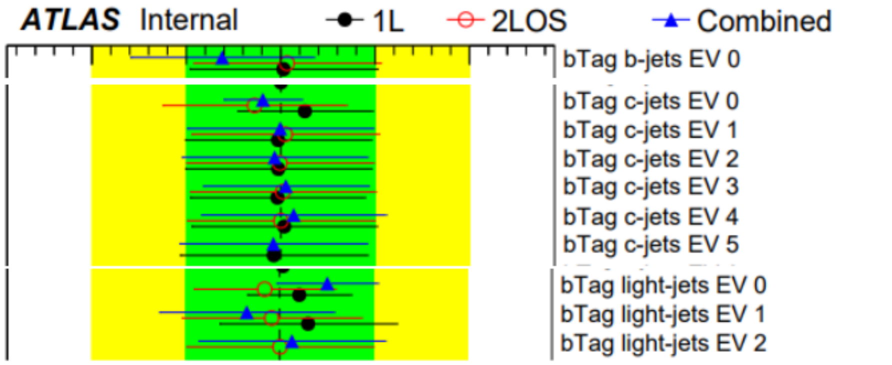
(a) nRCJets

Example plots

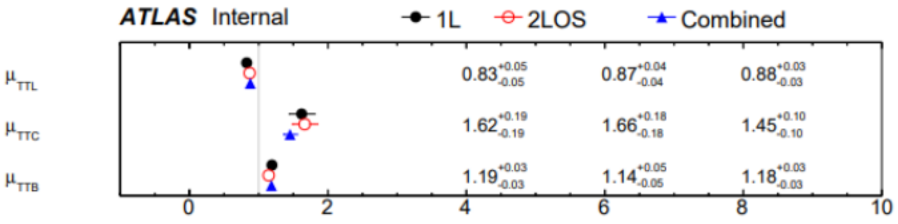
- Shows good modelling after NN-reweighting and fitting

Heavy Flavor Correction Factor

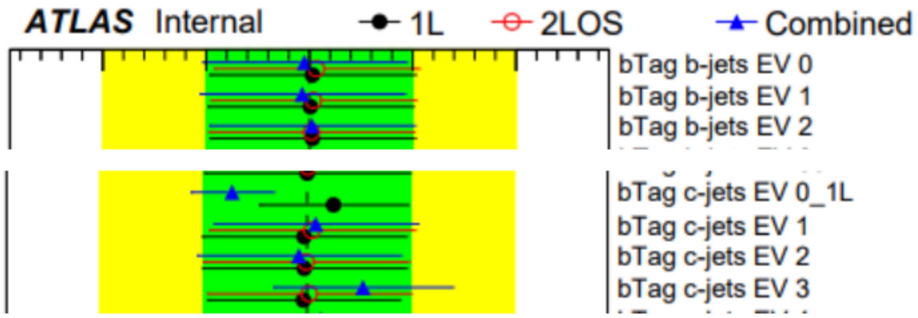
- Regions: 7j (5j) for 1L (2LOS) with 2b, 3b, ≥4b

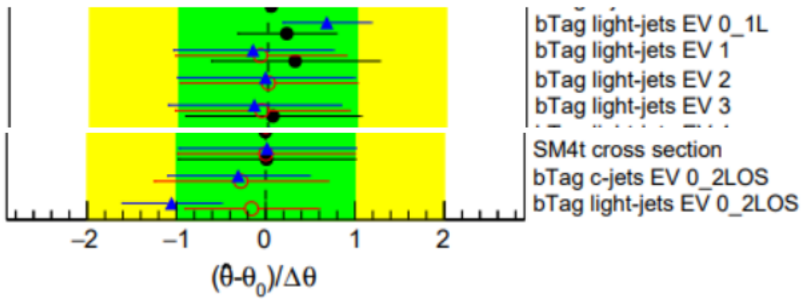


- Combined Fit with Nuisance Parameters Correlated among Channels
  - Unstable pulls is come from the performance differences among the two channels.
  - major differences can be found in NP bTag c-jets EV0 (C0) and bTag light-jets EV0 (L0)



- Combined Fit with C0 and L0 Decorrelated among Channels
  - Do not entirely resolve the issue





Heavy Flavor Correction Factor

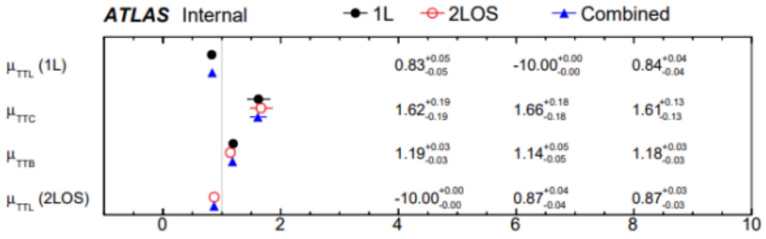
- Combined Fit with Different Correction Factors Decorrelation among Channels:

✓**Decorrelated TTL**:

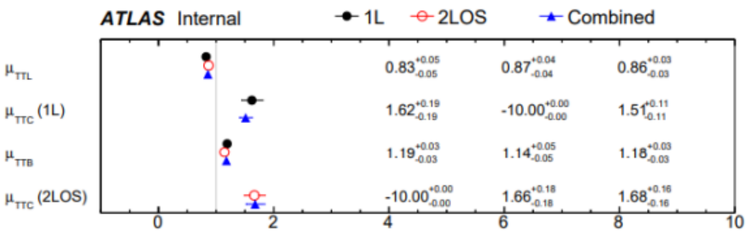
  - The correction factors derived from the combined fit are closed to and in between the results derived from 1L and 2LOS separated fits.

- **Decorrelated TTC**:
  - C0 pulls in the combined fit are opposite to the separated fits
  - μTTC (1L) derived in the combined fit is deviated from the result derived in the 1L separate fit

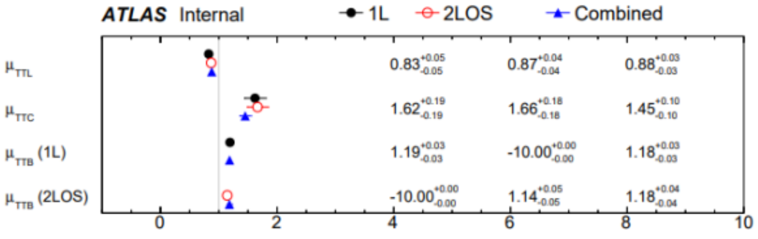
- **Decorrelated TTB**:
  - The NP pulls and the HF correction factors are nearly the same as the case of without decorrelating the correction factors.



Decorrelated TTL



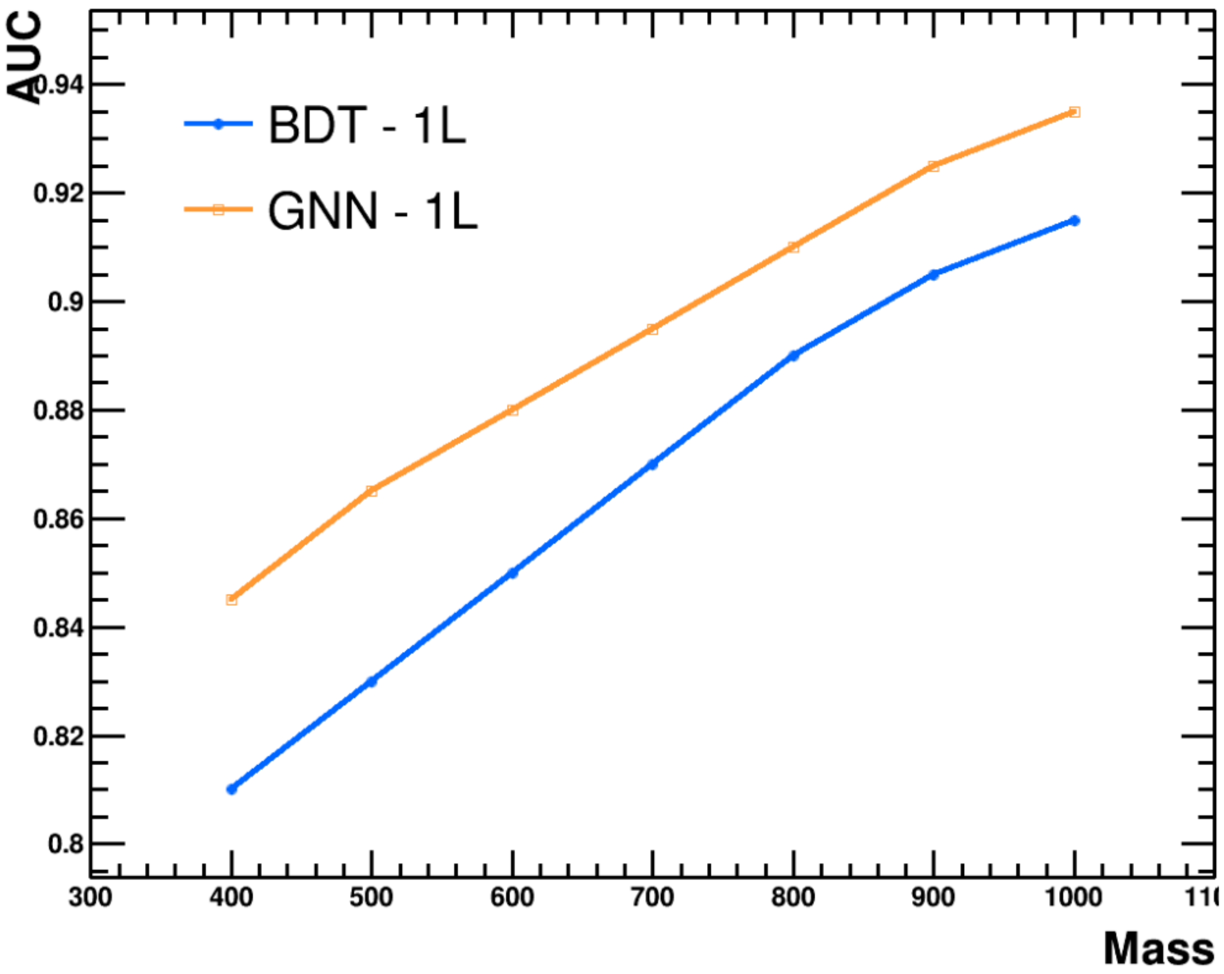
Decorrelated TTC



Decorrelated TTB

GNN VS BDT

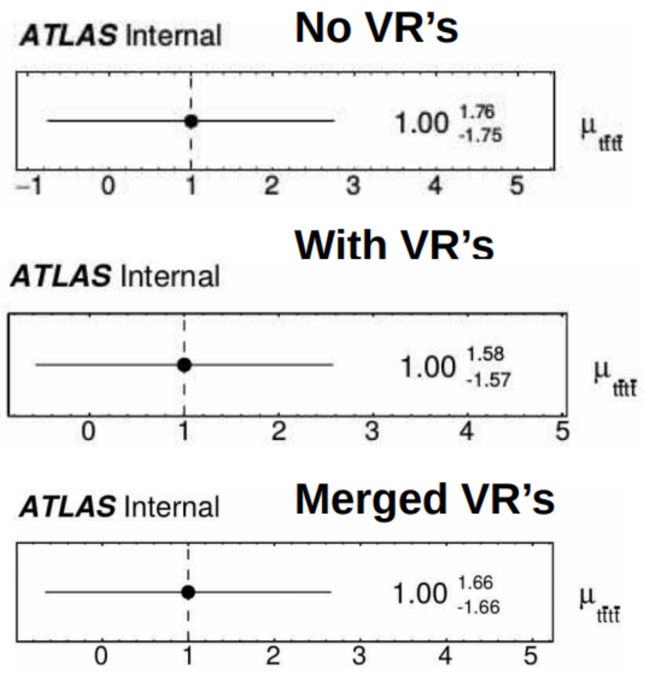
- SM4top like BDT has been also studies with the same input list used as global feature in GNN
- Purposed as a benchmark, and used only for control
- GNN shows significant improvement comparing to BDT
- GNN and BDT shows same level agreement over the data/MC in SR
- Observed same level of pulls/constrains in fits using BDT or GNN
- GNN found to be more sensitive

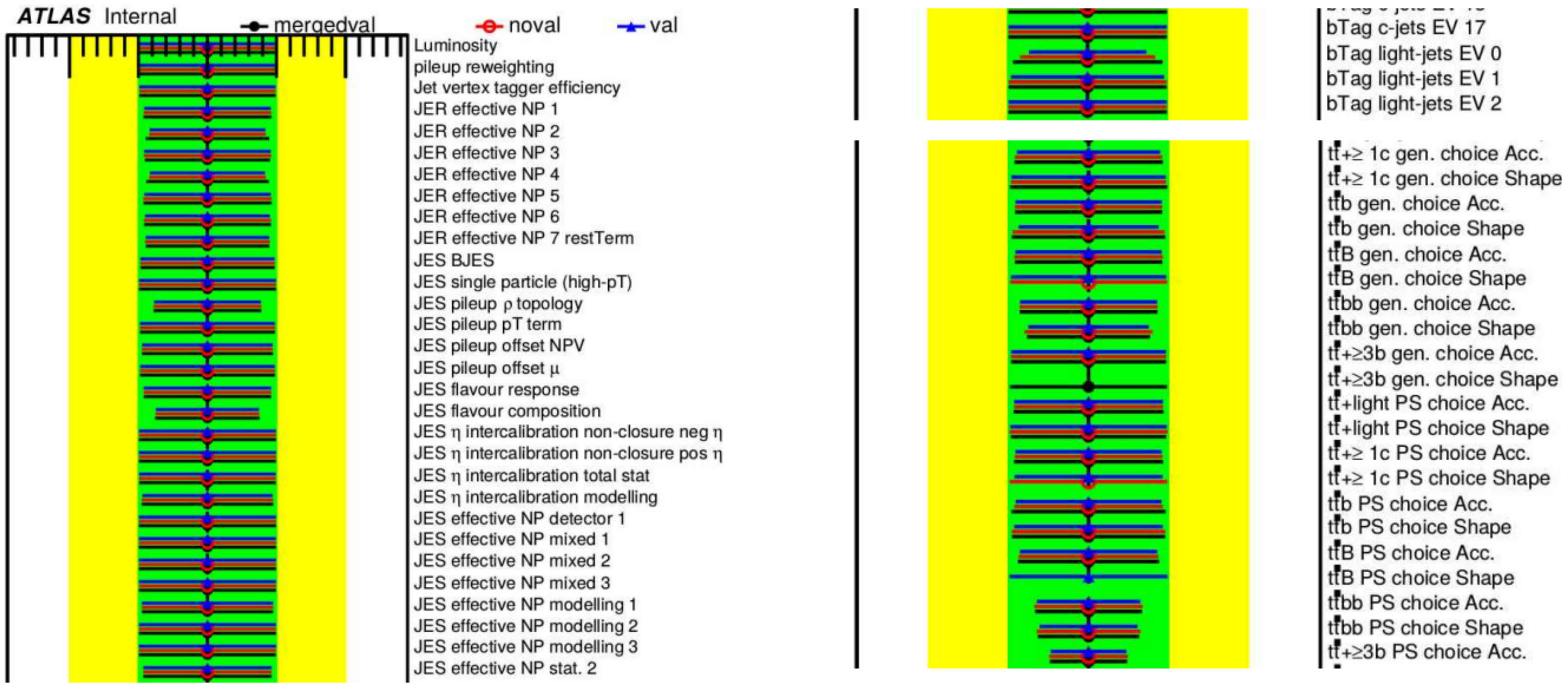


Validation Regions

- Validation region defined in SM 4-top is studied 3bV(SM 4-top)
  - In our analysis this region found to be more sensitive than 3bH
  - Simply this two region definition swapped in our analysis
    - 3bV (SM 4-top) = 3bH & 3bH(SM 4-top) = 3bV

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



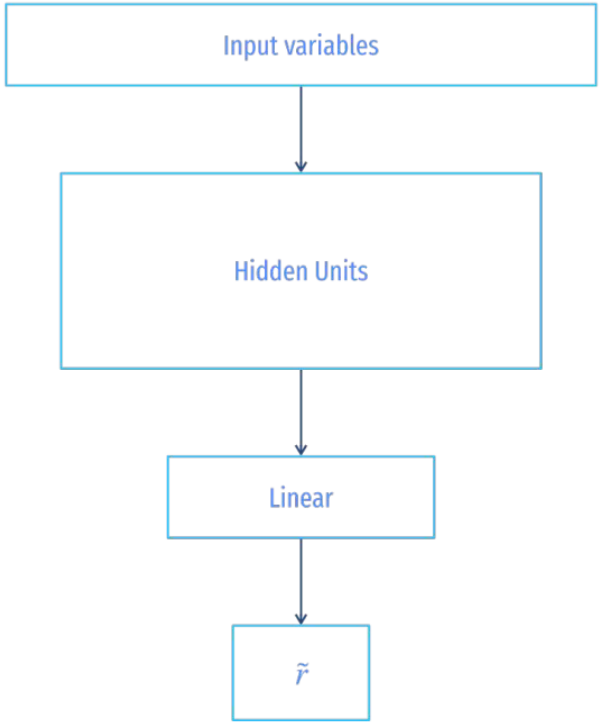


NN-reweighting (loss function)

Exponential Loss

$$\mathscr{L} = p_A e^{-r'/2} + p_B e^{r'/2}$$

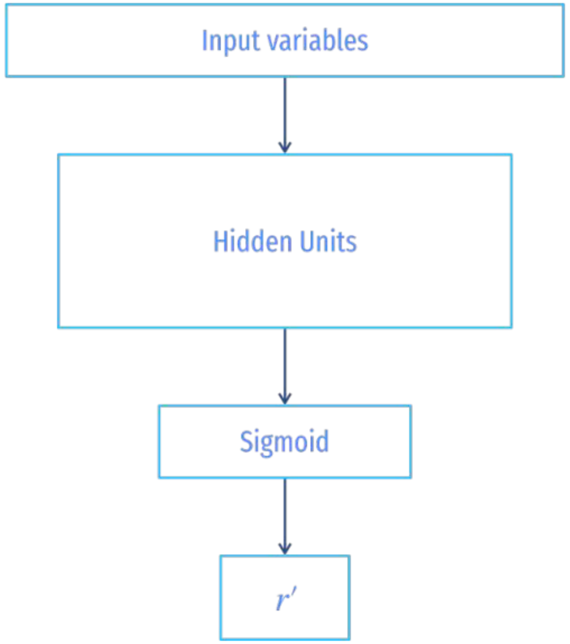
$$\frac{\partial \mathscr{L}}{\partial r'} = -\frac{p_A}{2} e^{-r'/2} + \frac{p_B}{2} e^{r'/2} = 0 \rightarrow \boxed{e^{r'} = \frac{p_A}{p_B} = r}$$



Binary Cross Entropy

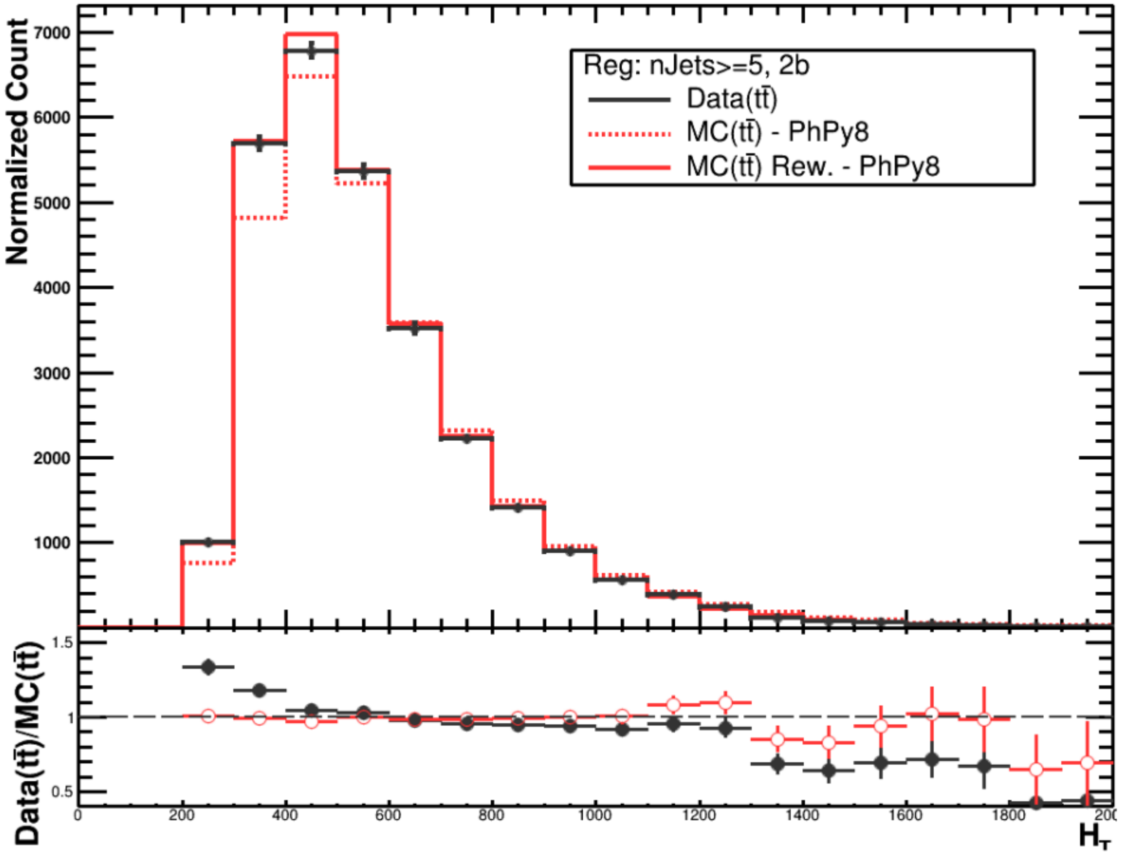
$$\mathscr{L} = -p_A \log r' - p_B \log(1 - r')$$

$$\frac{\partial \mathscr{L}}{\partial r'} = -\frac{p_A}{r'} + \frac{p_B}{1 - r'} = 0 \rightarrow \boxed{r' = \frac{p_A}{p_A + p_B} \rightarrow r = \frac{r'}{1 - r'}}$$

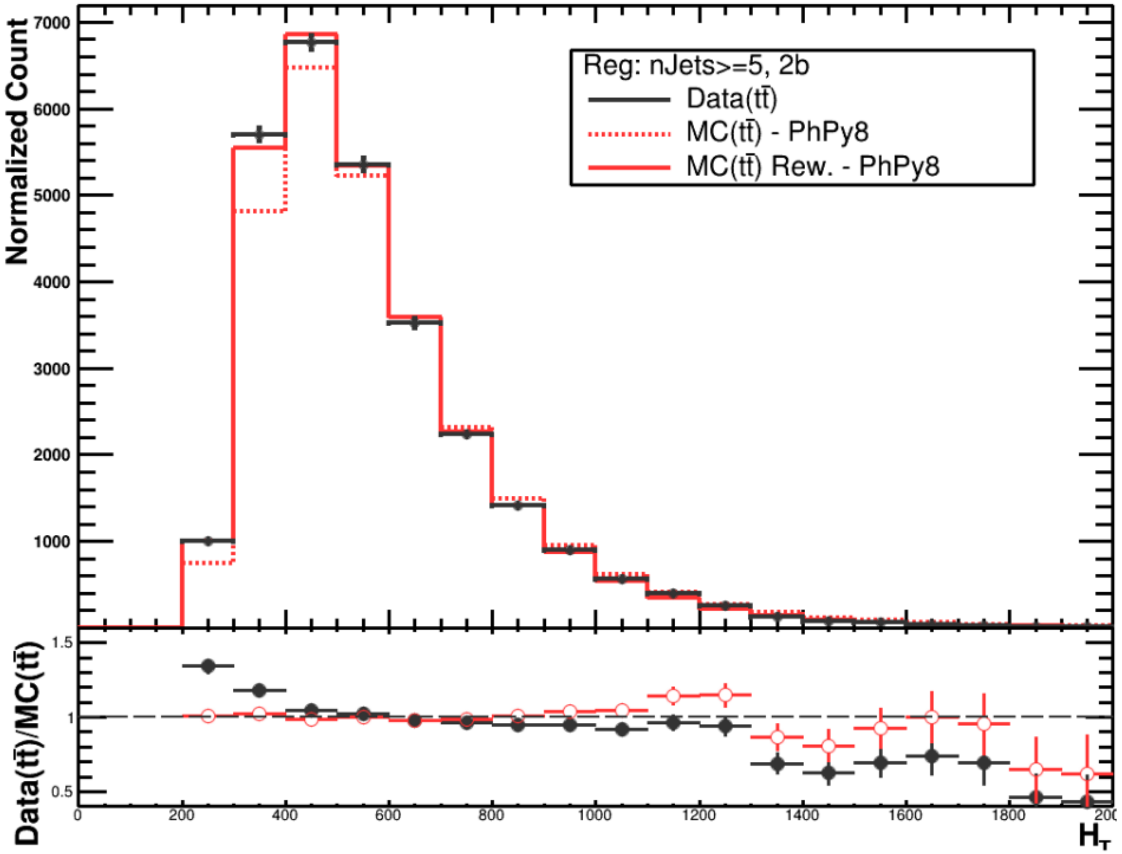


NN-reweighting (loss function)

Exponential Loss



Binary Cross Entropy



Expected significance