

# Search for a charged Higgs ( $H^\pm$ ) in the $W h$ ( $h \rightarrow b\bar{b}$ ) boosted channel using multivariate methods with the ATLAS experiment

Master's Thesis Project

Satyajit Chakraborty

Supervisors: Priv. Doz. Dr. Tatjana Lenz,  
Prof. Dr. Klaus Desch

With funding from:

57th Herbstschule HEP, 2026

10th September, 2026



**FSP ATLAS**

Erforschung von  
Universum und Materie

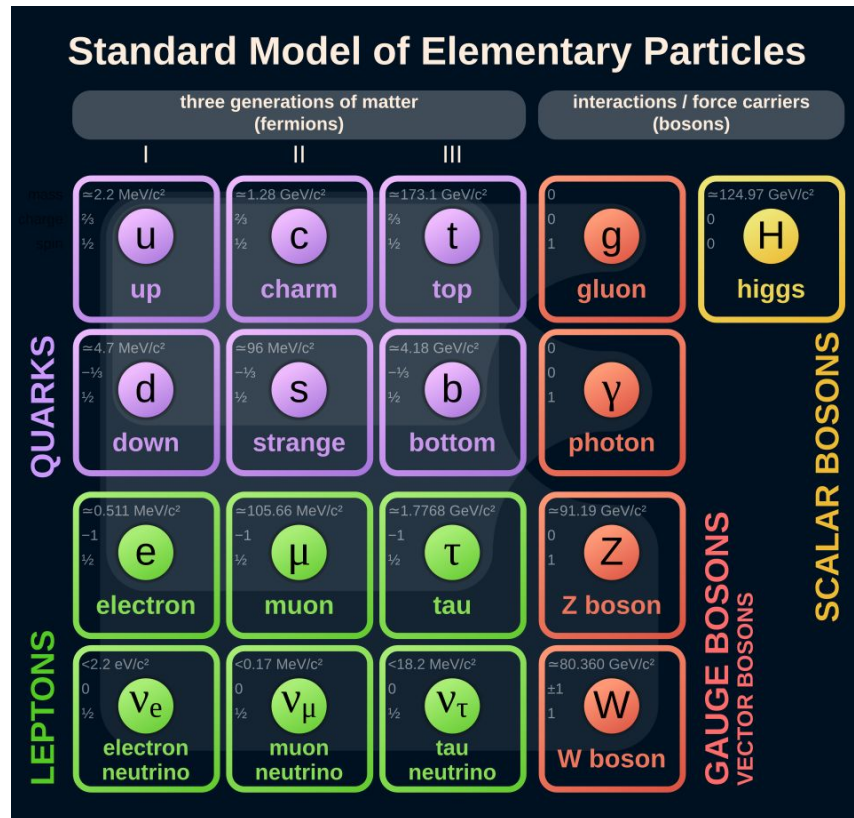
- ❑ Introduction
  - ❑ Event selection and categorisation
  - ❑ Baseline analysis: deep neural network (DNN) and  $m_{Wh}$  fit
  - ❑ Parameterized neural network (PNN) and fit on NN output score
  - ❑ Comparison of exclusion limits on signal cross section obtained using PNN and DNN (statistics-only)
- } Following previous ATLAS analysis: [arXiv:2411.03969](https://arxiv.org/abs/2411.03969)

# The Standard Model

- ❖ Well tested model explaining the fundamental interactions of matter.
- ❖ Incorporates special relativity and quantum mechanics.
- ❖ Valid over a huge energy range (eV to TeV).
- ❖ Last missing particle discovered in 2012 (Higgs) 

Some limitations:

- ❖ Matter-Antimatter asymmetry
- ❖ Dark Matter
- ❖ Mass-Hierarchy problem
- ❖ Gravity

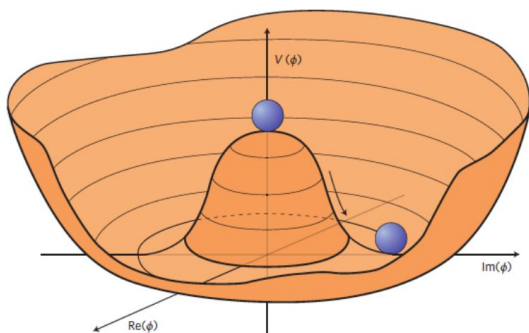


Wikimedia, Standard Model, (2016), url: [https://en.wikipedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles.svg](https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg)

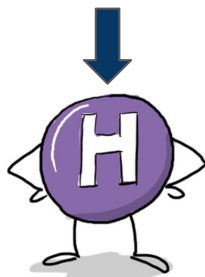
# Higgs and extended scalars

## Standard Model

One scalar  $SU(2)_L$  doublet  $\phi$



$\mu < 0, \langle \phi \rangle \neq 0$   
 $\rightarrow$  Spontaneous  
 Symmetry  
 breaking



$$V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4$$

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

$$\nu = 246 \text{ GeV}$$

Why extend the  
 scalar sector?

Incorporates  
 CP-violation

Addresses  
 Hierarchy

Connects to BSM

A simple and well  
 motivated  
 extension of the SM

## Two doublet Higgs Model

Addition of a scalar doublet:  $\phi_1$  and  $\phi_2$

$$V(\Phi_1, \Phi_2) = m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \text{h.c.}) \\
+ \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 + \lambda_3 (\Phi_1^\dagger \Phi_1) (\Phi_2^\dagger \Phi_2) \\
+ \lambda_4 (\Phi_1^\dagger \Phi_2) (\Phi_2^\dagger \Phi_1) + \frac{\lambda_5}{2} [(\Phi_1^\dagger \Phi_2)^2 + \text{h.c.}]$$

$$\langle \Phi_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix}$$

$$\langle \Phi_2 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}$$

$$\tan \beta = \frac{v_2}{v_1}$$

$$\nu = \sqrt{\nu_1^2 + \nu_2^2} = 246 \text{ GeV}$$

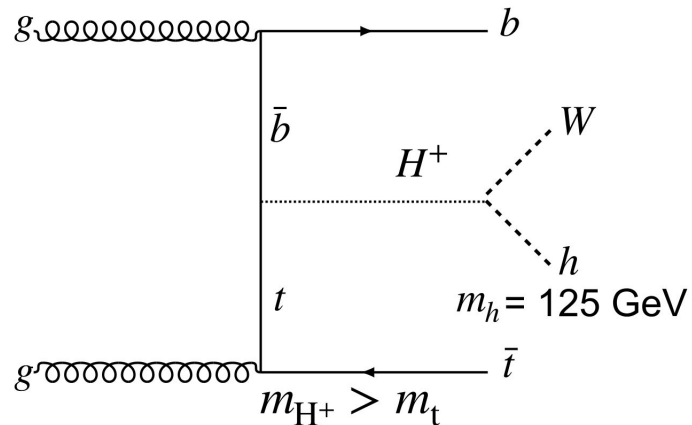
Type II considered for this analysis

# 2HDM and charged Higgs

❖ 2HDM leading to 5 physical states after EWSB:

- $h(\text{light}),$
  - $H(\text{heavy}),$
  - $A$  (CP-Odd)
  - **$H^\pm$  (charged Higgs Bosons).**
- } (CP-Even)

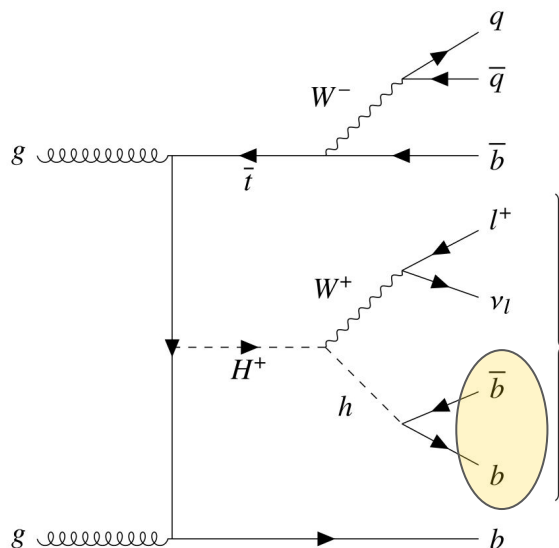
In Type II:  
 $tbH^\pm$  coupling  $\propto m_b \tan\beta + m_t \cot\beta$   
 $\rightarrow$  sizeable production



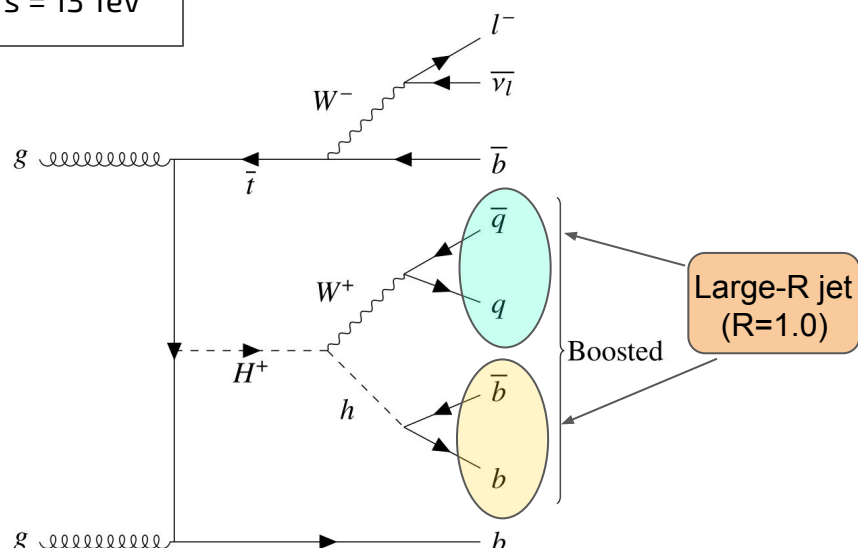
This search

# Signal Topology

$$\mathcal{L}_{\text{int}} = 140.1 \text{ fb}^{-1}, \text{ Run-2, } \sqrt{s} = 13 \text{ TeV}$$



**$\ell\nu b\bar{b}$  channel**

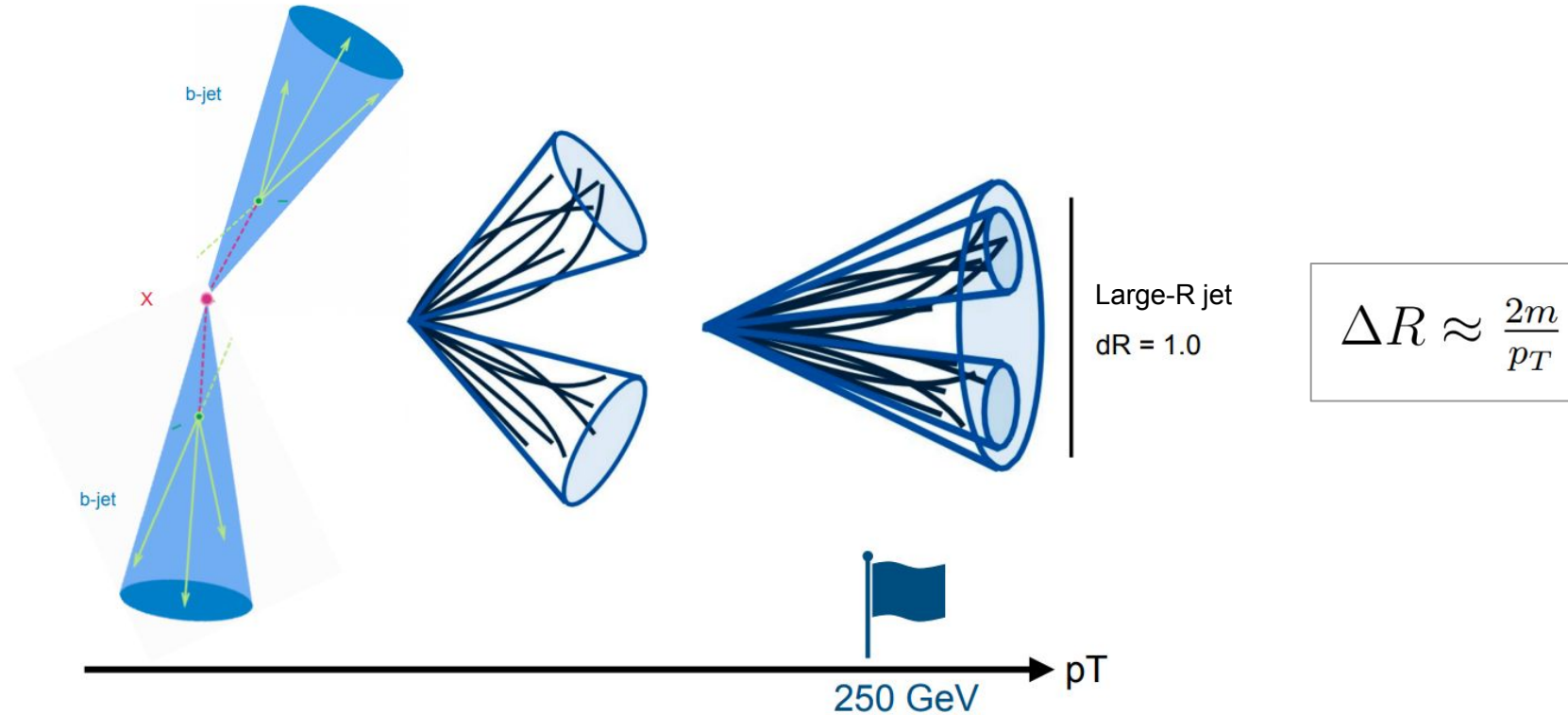


**$q\bar{q} b\bar{b}$  channel**

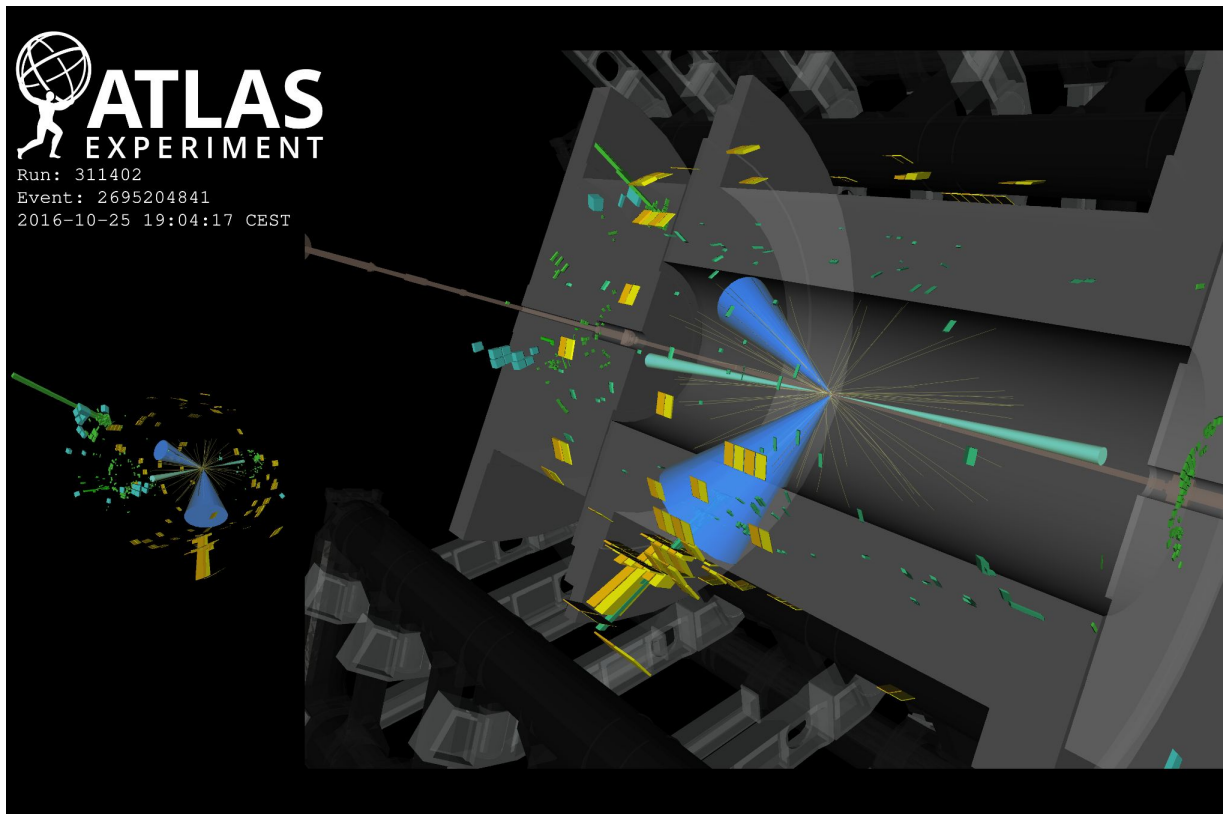
**Signal mass hypothesis studied : [1.2, 1.4, 1.6, 1.8, 2.0, 2.5, 3.0] TeV**

**Leading backgrounds:  $t\bar{t}$  + LF jets,  $t\bar{t}$  + HF jets, W/Z + jets**

# How $h \rightarrow b\bar{b}$ actually looks like?



# How $h \rightarrow b\bar{b}$ actually looks like?



<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HDBS-2022-02/index.php>

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)



Phys. Lett. B 858 (2024) 139007  
DOI: [DOI:10.1016/j.physletb.2024.139007](https://doi.org/10.1016/j.physletb.2024.139007)



CERN-EP-2024-092  
14th October 2024

**Search for pair production of boosted Higgs bosons  
via vector-boson fusion in the  $b\bar{b}b\bar{b}$  final state using  
 $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS  
detector**

The ATLAS Collaboration

- ❑ Introduction
  - ❑ Event selection and categorisation
  - ❑ Baseline analysis: deep neural network (DNN) and  $m_{Wh}$  fit
  - ❑ Parameterized neural network (PNN) and fit on NN output score
  - ❑ Comparison of exclusion limits on signal cross section obtained using PNN and DNN (statistics-only)
- } Following previous ATLAS analysis: [arXiv:2411.03969](https://arxiv.org/abs/2411.03969)

# Event selection and categorization

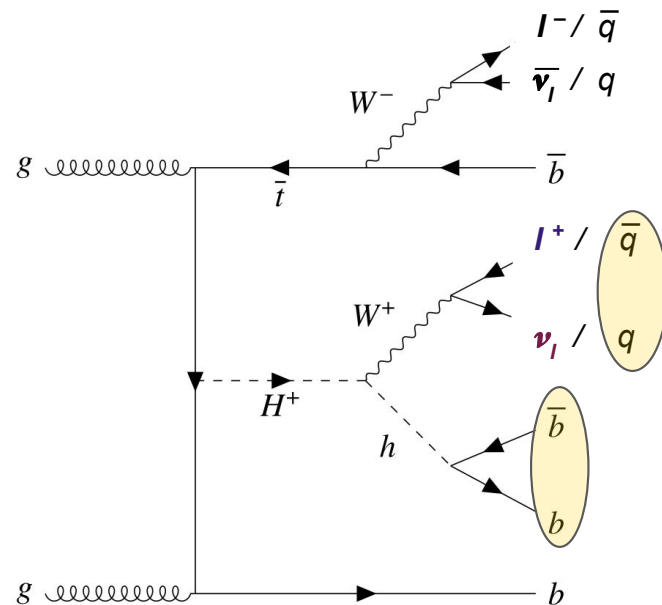
## Event Selection:

- ❖ **Trigger:** pass at least one of the single  $e/\mu$  triggers.
- ❖ **Missing transverse momentum:**  $E_T^{miss} > 30$  GeV
- ❖ **Leptons:** Exactly one  $e$  or  $\mu$  with  $p_T > 27$  GeV
- ❖ **Large-R jets**  $\geq 1$  ( $p_T \geq 250$  GeV)
- ❖ Exactly one **Large-R jet** passing the **60% working point** of the **Higgs-Tagger\***.
- ❖ **Overlap removal:** **Small-R jets** considered to be **b-tagged** if they satisfy  $\Delta R > 1.4$  from the **Large-R jet**

## Categorization:

| Channel          | Large-R jets | W boson candidate                                                                        | Higgs candidate          |
|------------------|--------------|------------------------------------------------------------------------------------------|--------------------------|
| $qqbb$           | $\geq 2$     | Fails $X \rightarrow b\bar{b}$ tagging<br>$50 \text{ GeV} \leq M_j \leq 110 \text{ GeV}$ | $X \rightarrow b\bar{b}$ |
| $\ell vb\bar{b}$ | $\geq 1$     | $W_{lep}$                                                                                | $X \rightarrow b\bar{b}$ |

B-tagging working point: 77%



\*  $M_j$ : Jet Mass

\* [ATL-PHYS-PUB-2020-019](#)

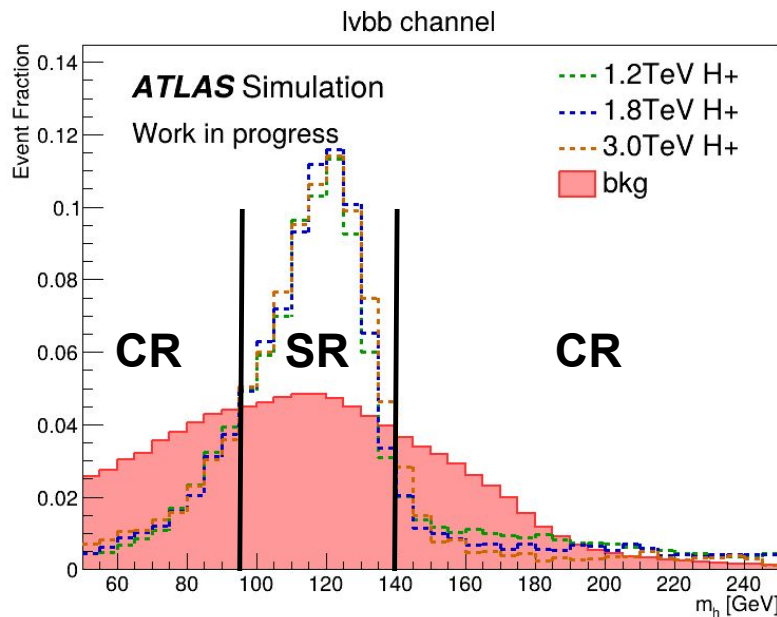
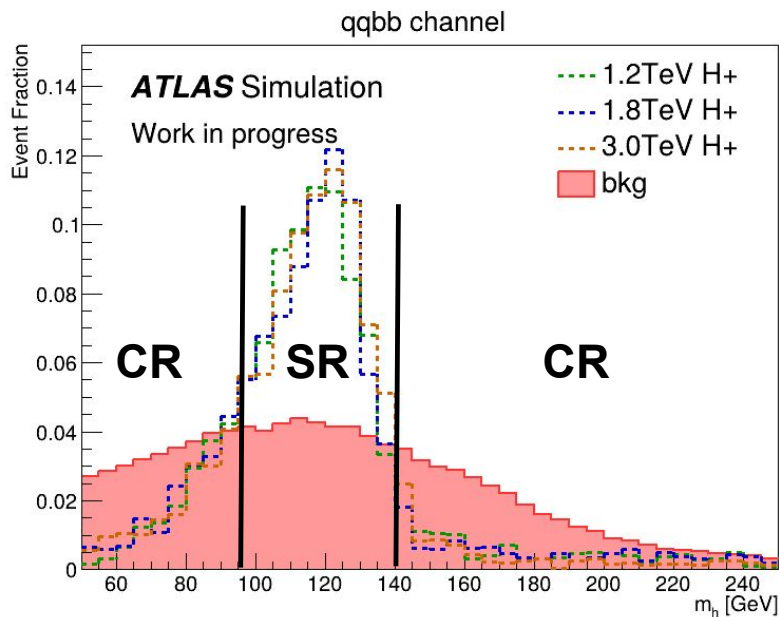
# Signal and Control Regions

❖ **Signal region:**  $95 \text{ GeV} \leq m_h \leq 140 \text{ GeV}$

❖ **Control region:**

➤ **Lower sideband:**  $m_h < 95 \text{ GeV}$

➤ **Higher sideband:**  $m_h > 140 \text{ GeV}$



# Signal and Control Regions

❖ **Signal region:**  $95 \text{ GeV} \leq m_h \leq 140 \text{ GeV}$

❖ **Control region:**

➤ **Lower sideband:**  $m_h < 95 \text{ GeV}$

➤ **Higher sideband:**  $m_h > 140 \text{ GeV}$

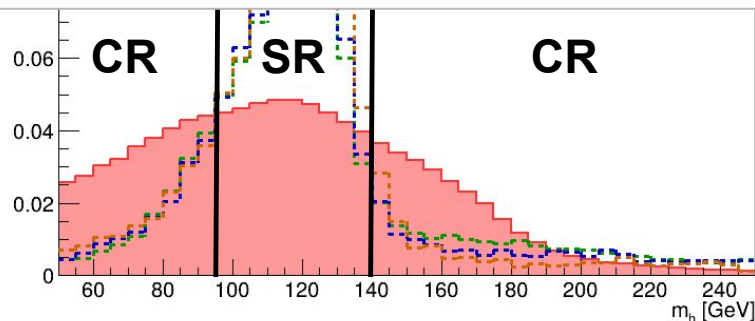
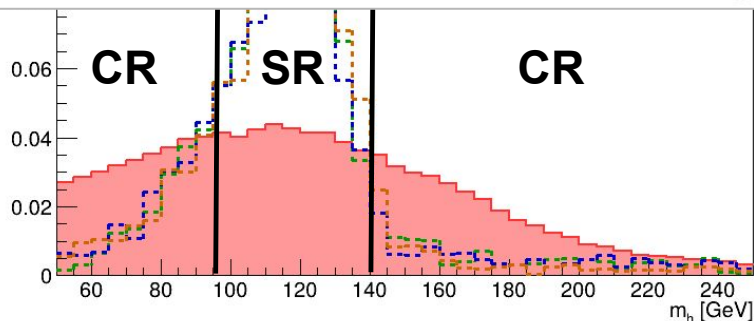
❖ Further division based on the number of small radius b-tagged jets outside Large-R jets.

❖ Additional division into two regions: **0 b-tagged,  $\geq 1$  b-tagged.**

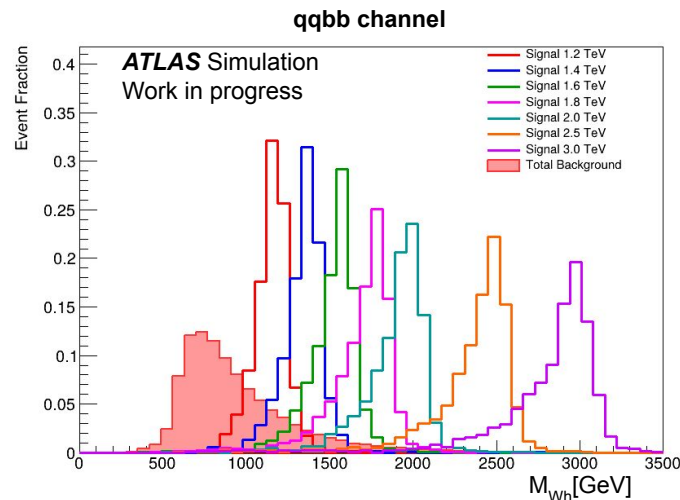
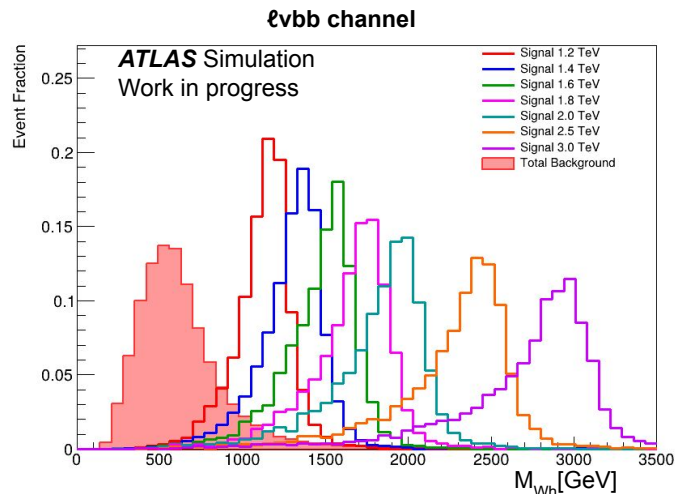
❖ **Signal regions** are further divided into



- High nn score
  - Medium nn score
  - Low nn score
- } after using a DNN for event S vs B classification



# $M_{Wh}$ as a fit variable

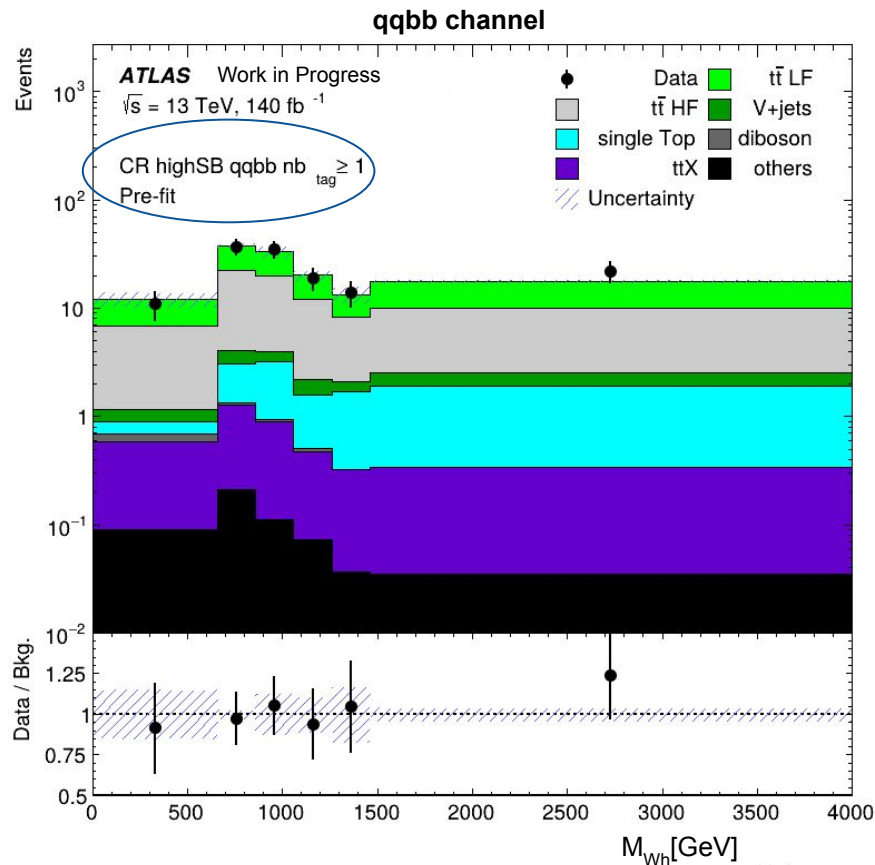
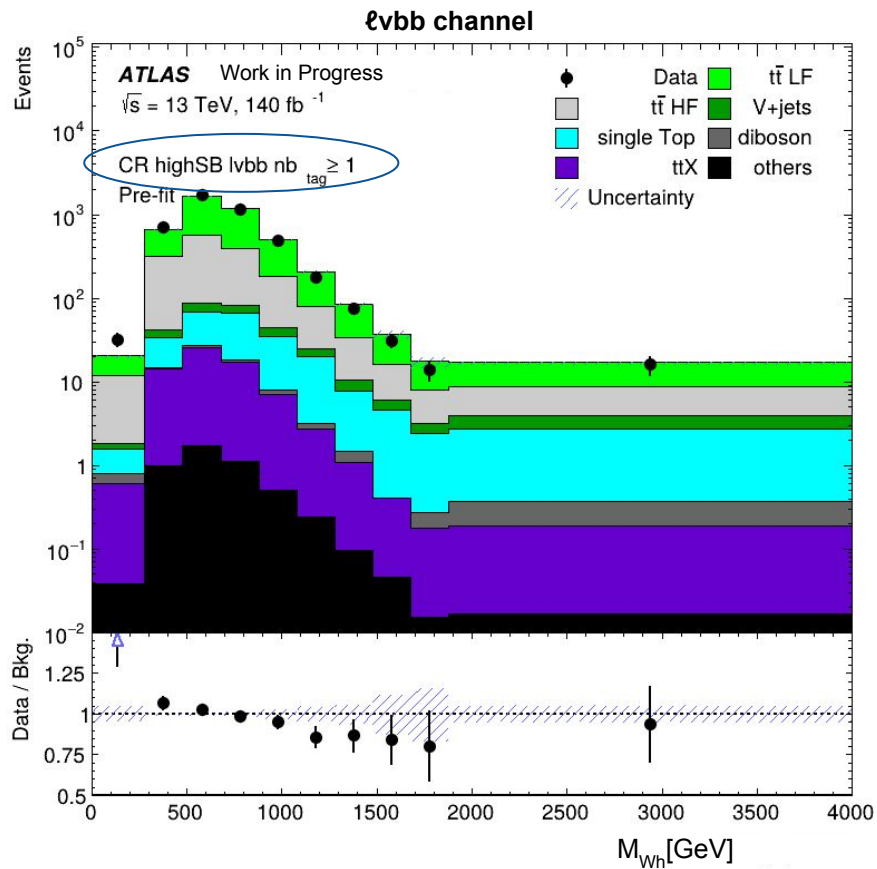


**Clear separation between the signal distributions and the total background**

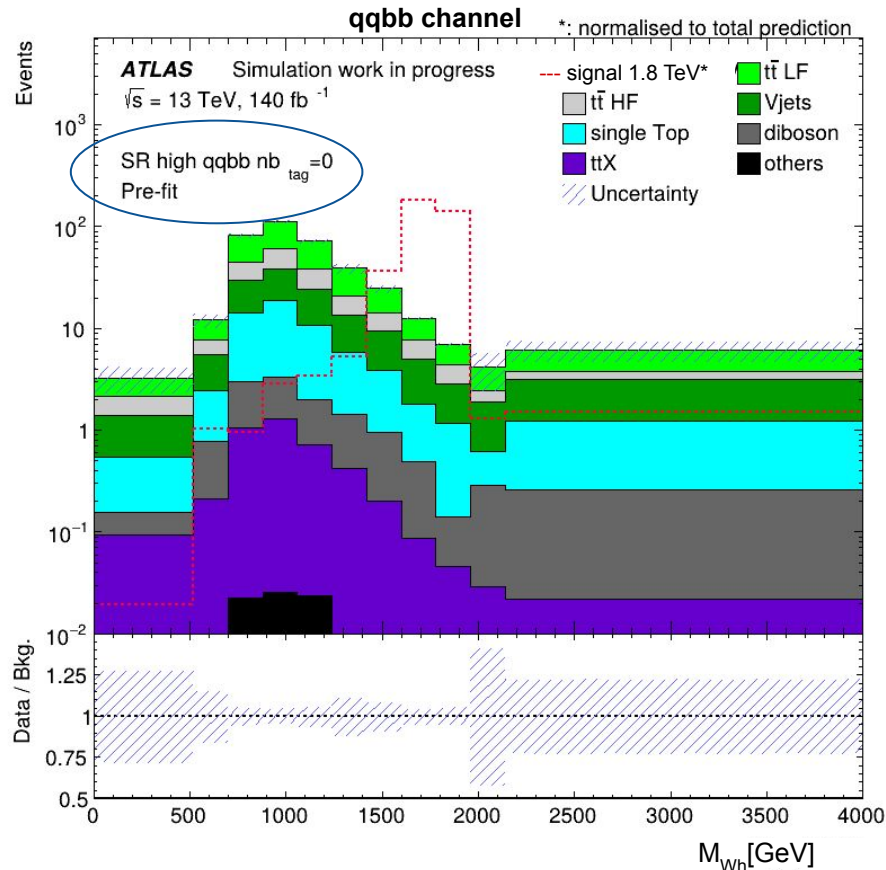
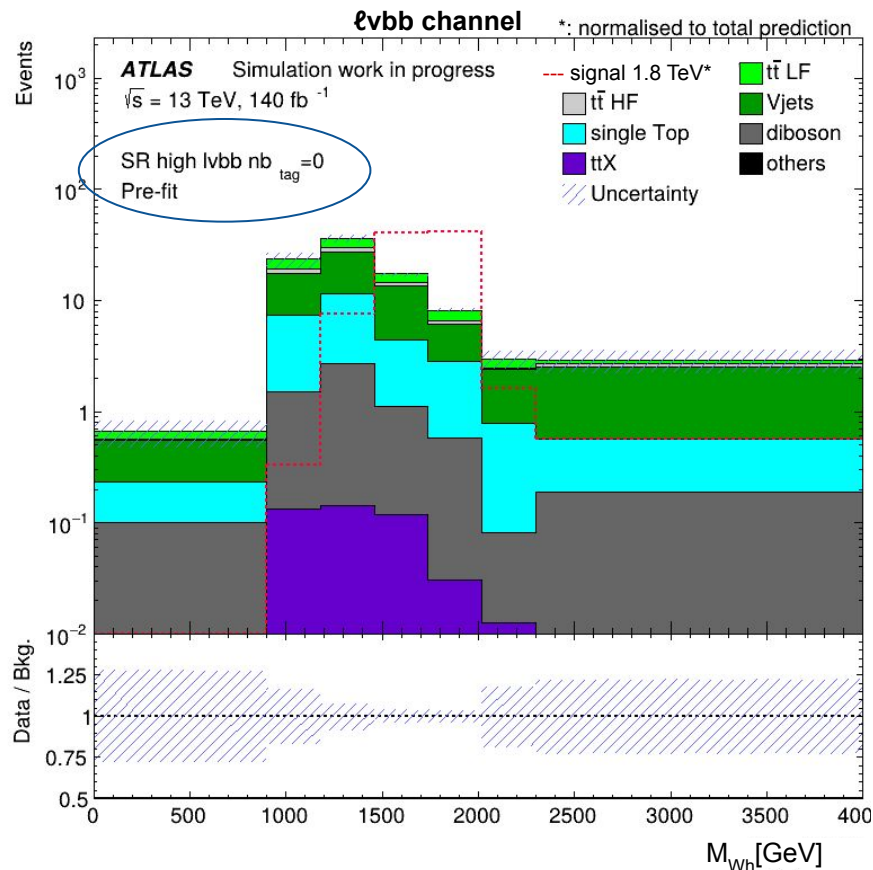
- ❖ Normalization factors for  $t\bar{t} + \text{LF}$ ,  $t\bar{t} + \text{HF}$ ,  $W/Z + \text{jets}$  -> **free parameters**
- ❖ **Background-only fit to data in control regions** -> for obtaining the normalization factors.
- ❖ Simultaneous **Signal + Background blinded fit** in all **signal** and **control** regions to obtain the **signal strength parameter ( $\mu$ )** and **upper limits on ( $\sigma \times \text{BR}$ )**.
- ❖ Both  **$q\bar{q}b\bar{b}$**  and  **$\ell\nu b\bar{b}$**  channels included in the fit setup.

**Fit  
Setup**

# $M_{Wh}$ MC-data Agreement in CR



# $M_{Wh}$ distribution in SR (Blinded)



- ❑ Introduction
  - ❑ Event selection and categorisation
  - ❑ Baseline analysis: deep neural network (DNN) and  $M_{wh}$  fit
  - ❑ **Parameterized neural network (PNN) and fit on NN output score**
  - ❑ Comparison of exclusion limits on signal cross section obtained using PNN and DNN (statistics-only)
- } Following previous ATLAS analysis: [arXiv:2411.03969](https://arxiv.org/abs/2411.03969)

# PNN for event classification

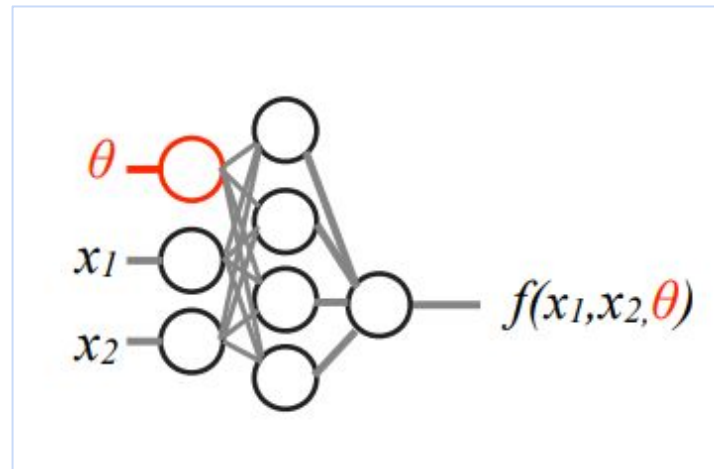
## PNN for improved sensitivity

- ❖ Identical Neural network structure as DNN
- ❖ Additional input: **true  $H^\pm$  mass** as parameter.  
Signal: True  $H^\pm$  mass  
Background: Random  $H^\pm$  mass from considered hypotheses
- ❖ Advantages: smooth interpolation

## Two Approaches chosen:

- ❖ Trained a PNN with the same DNN setup and variables: **yielded minimal improvement (PNN 1)**
- ❖ Training a PNN with a combination of new inputs and old inputs and a more optimized network (**PNN 2**) ➡ 

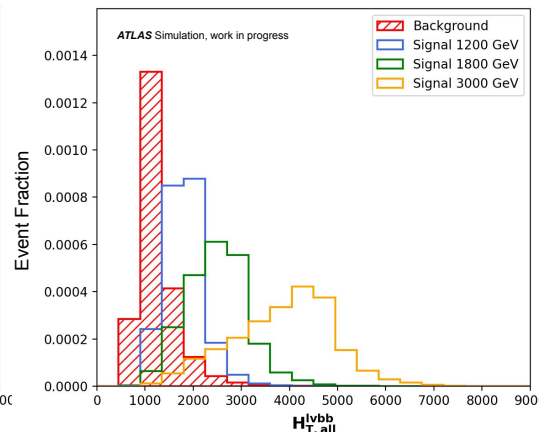
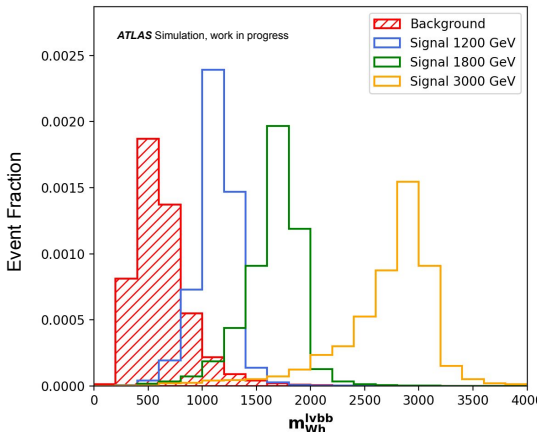
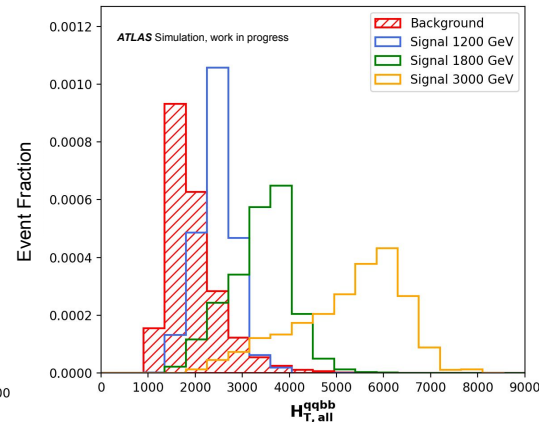
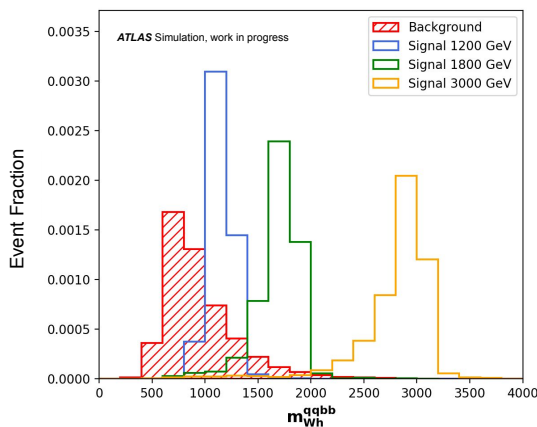
up-ahead



Baldi, Pierre, et al. [arXiv:1601.07913](https://arxiv.org/abs/1601.07913) (2016).

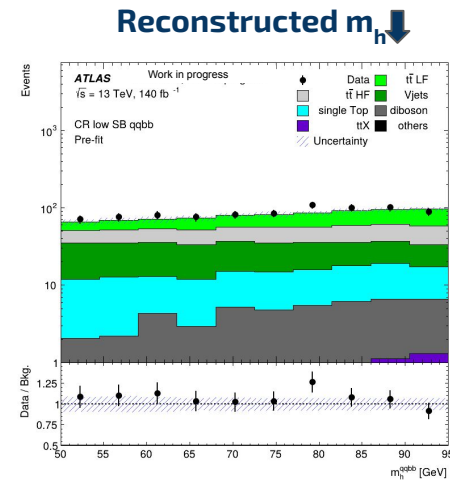
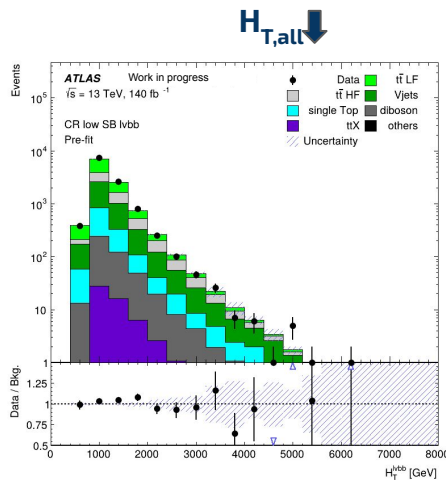
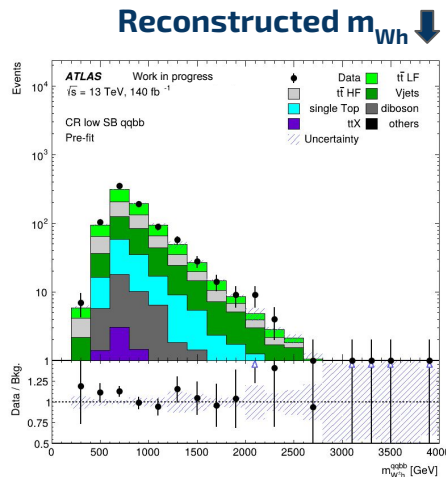
Caveat: the PNN score is mass-dependent. Therefore, both signal and background events are re-evaluated at each mass point, giving a different score distribution for each hypothesis.

- ❖ Best input features selected using **mass-normalized permutation importance**.  
(More in back-up)
- ❖ **Most sensitive variable: reconstructed charged Higgs mass**
- ❖ Hyperparameter optimization done using **optuna (Bayesian optimization)**  
Far efficient than tedious grid and random search.

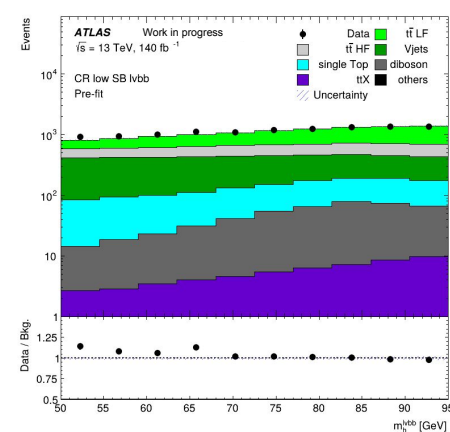
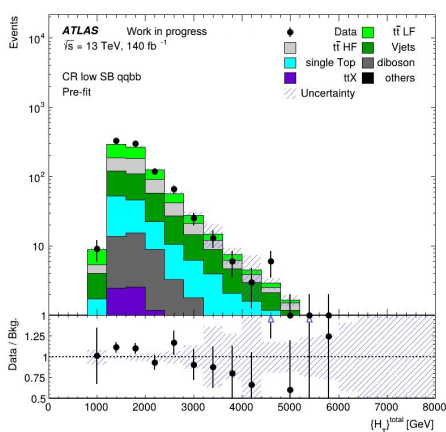
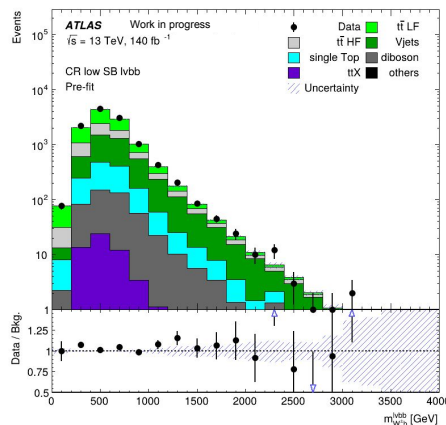


# Data-MC comparison in CR (PNN-2 inputs)

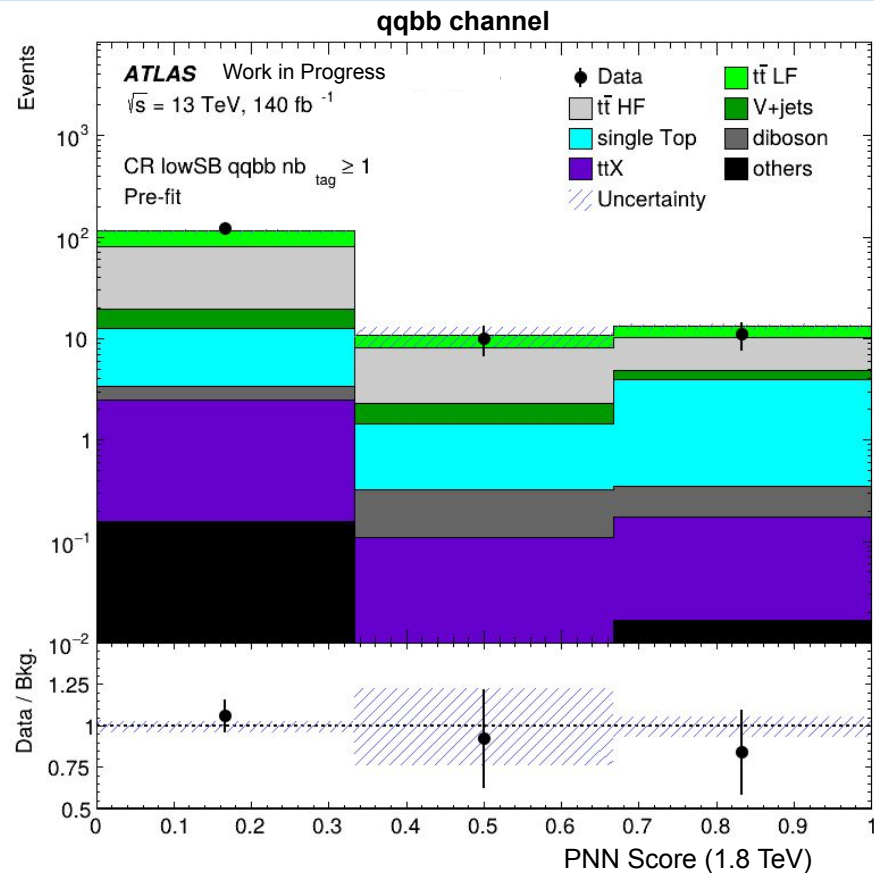
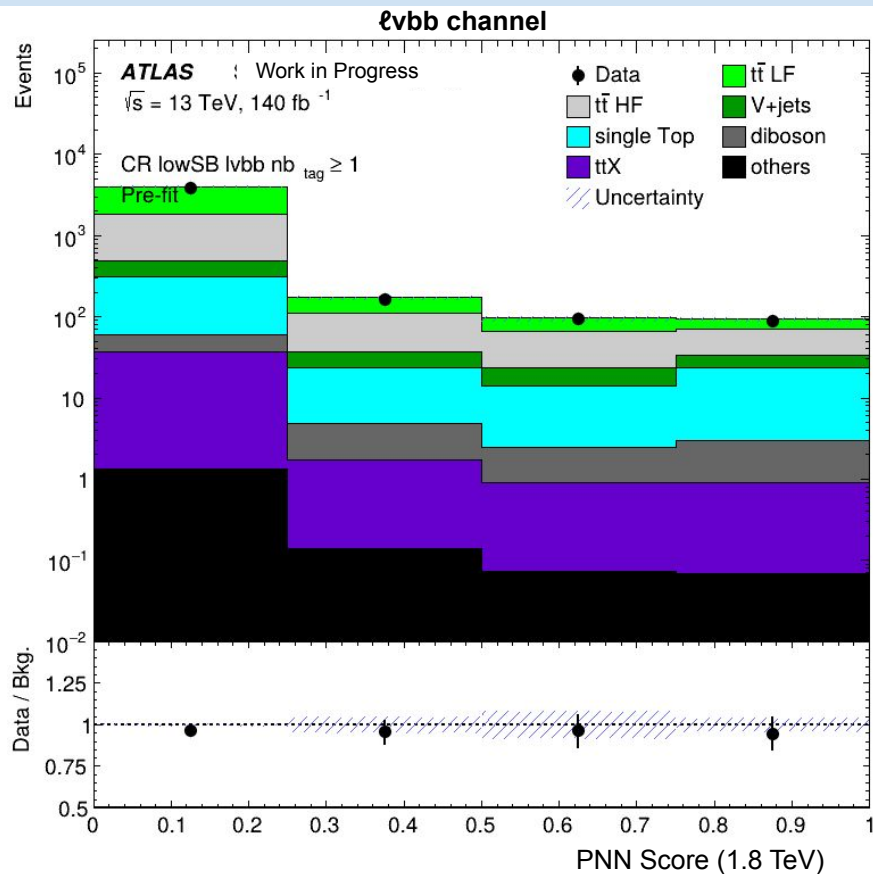
qqbb →



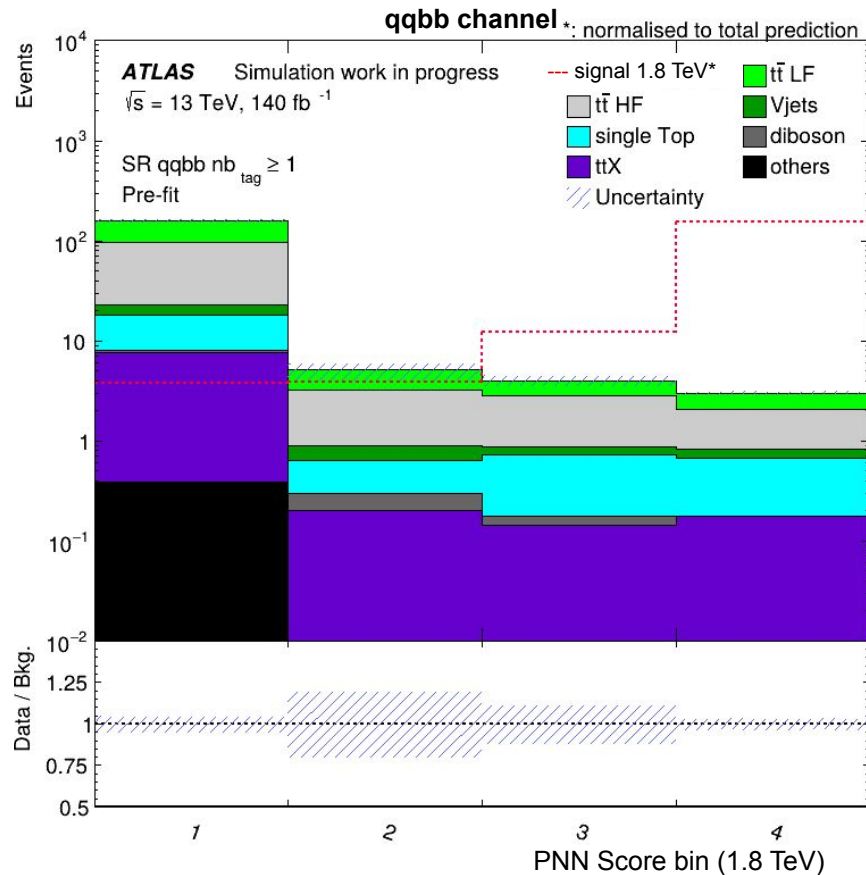
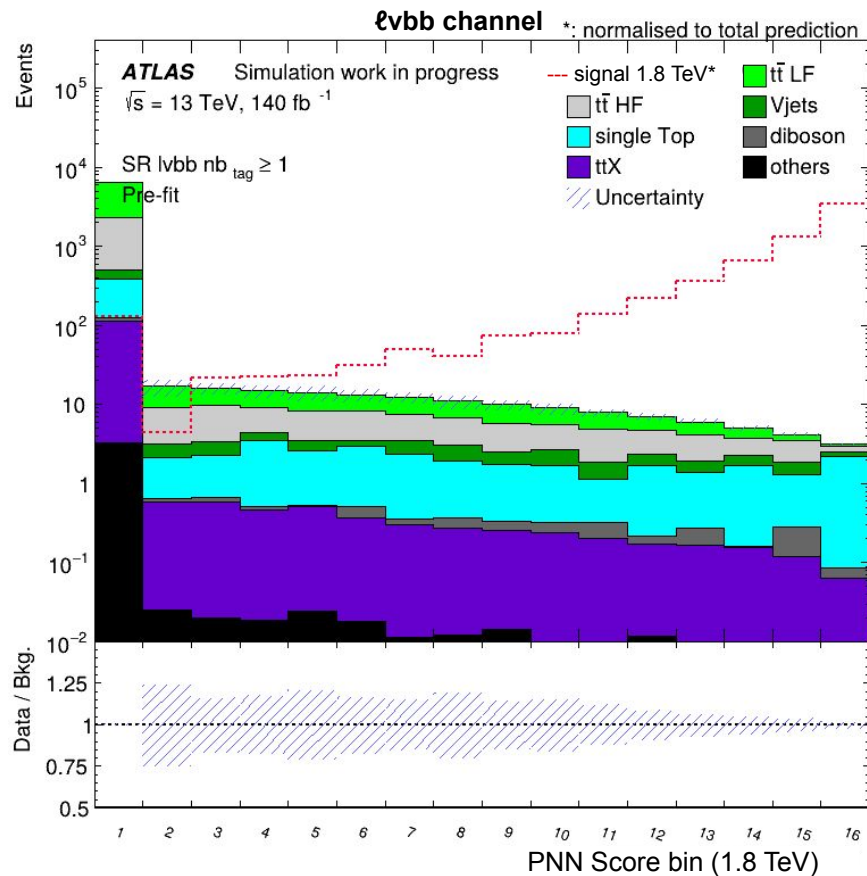
lvbb →



# PNN-2 score MC-data Agreement in CR



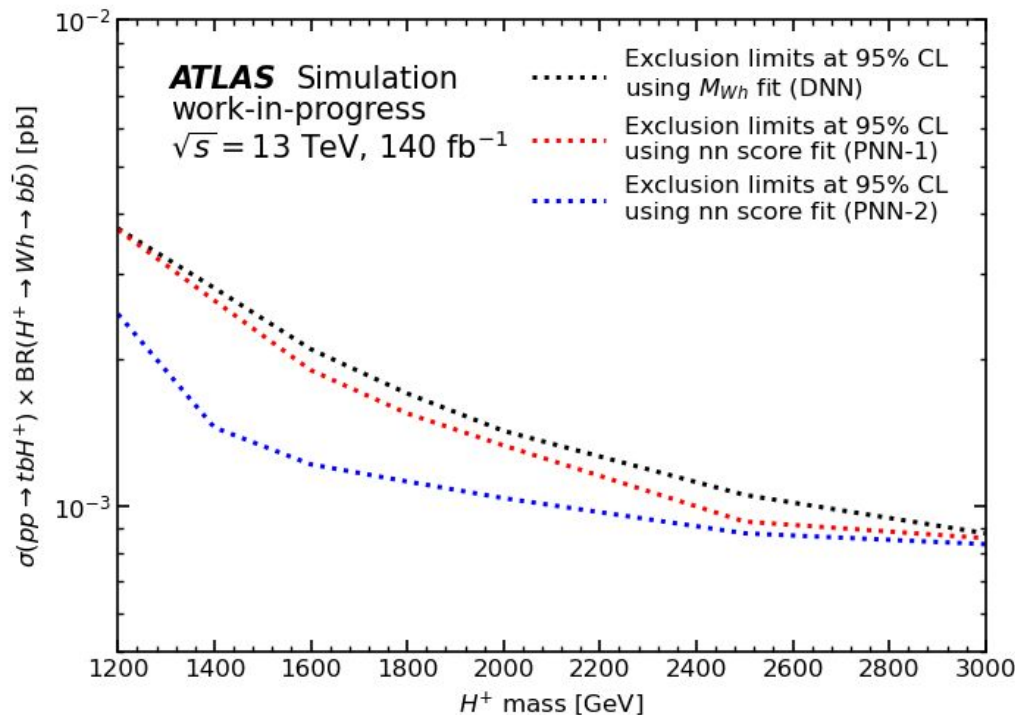
# PNN-2 score distribution in SR (Blinded)



- ❑ Introduction
  - ❑ Event selection and categorisation
  - ❑ Baseline analysis: deep neural network (DNN) and  $M_{Wh}$  fit
  - ❑ Parameterized neural network (PNN) and fit on NN output score
  - ❑ Comparison of exclusion limits on signal cross section obtained using PNN and DNN (statistics-only)
- } Following previous ATLAS analysis: [arXiv:2411.03969](https://arxiv.org/abs/2411.03969)

# Comparison of exclusion limits

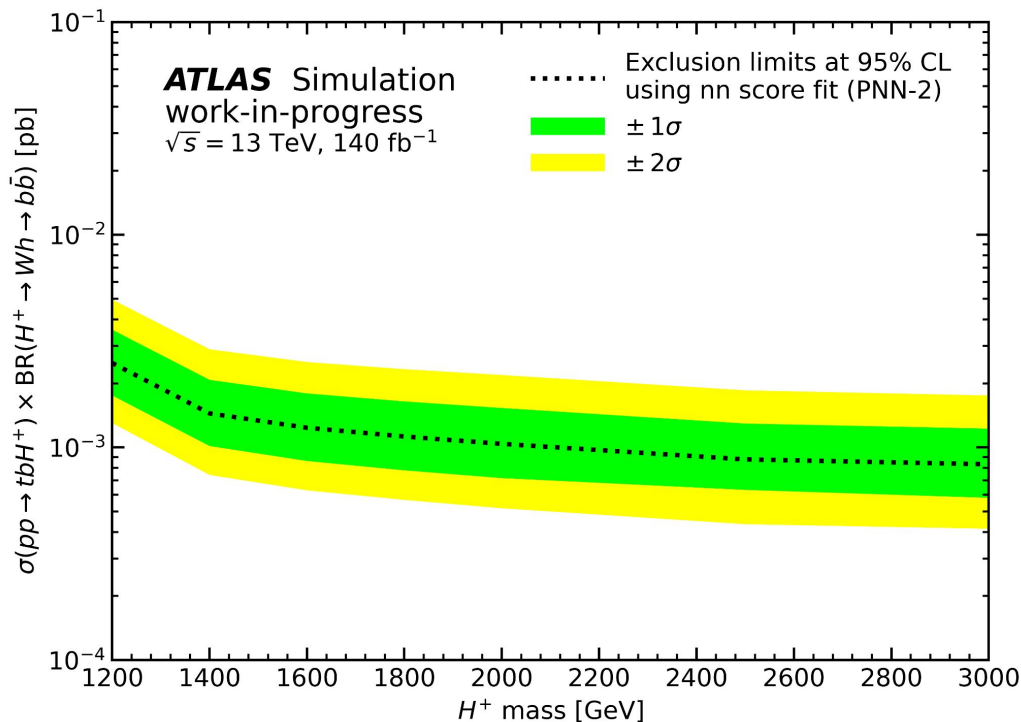
## Expected Limits (Statistics-only)



- ❖ DNN with  $M_{Wh}$  fit compared to PNN-1 and PNN-2 with nn score fit.
- ❖ Major improvement in sensitivity using PNN-2 with nn score fit with: 33% (1.2 TeV), but only 9% (3 TeV) wrt to DNN.
- ❖ Using analysis sensitive variables and fitting on the PNN output score improved sensitivity.
- ❖ Better exclusion limits compared to  $M_{Wh}$  fit.

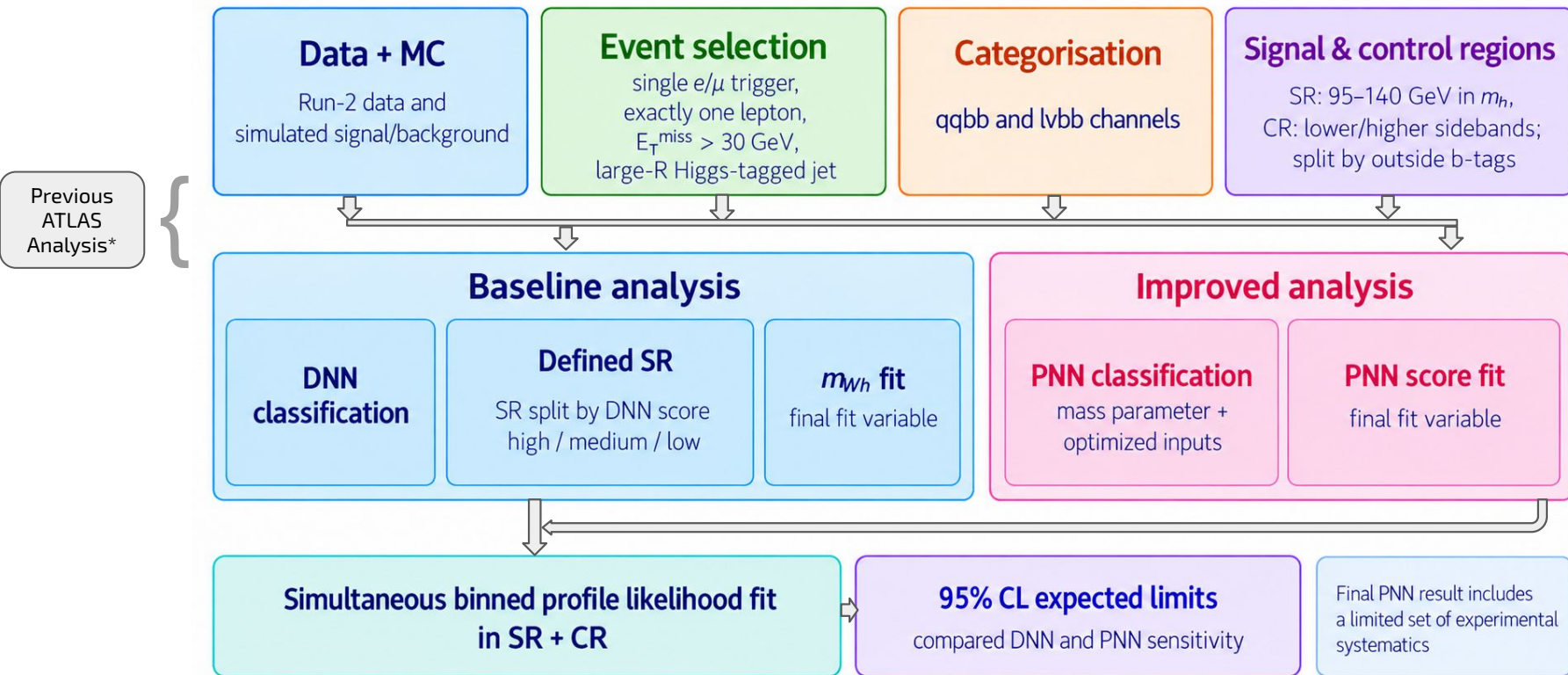
# PNN-2 limits including systematics

## Expected Limits



- ❖ Limited set of instrumental systematics included due to time constraint.
- ❖ **Integrated Luminosity** is the dominant one among all. N.F.s enter the fit as nuisance parameters.
- ❖ Expected limits worsen by only **0.8% at 1.2 TeV** and **0.1% at 3 TeV**.

# Summary



# Outlook

- ❖ Extend PNN-2 with the full set of relevant systematic uncertainties.
- ❖ Unblind the PNN-2 analysis and compare its performance directly with the previous ATLAS strategy.
- ❖ Explore more unified multivariate approaches across both channels.

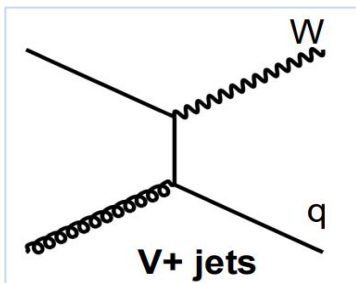
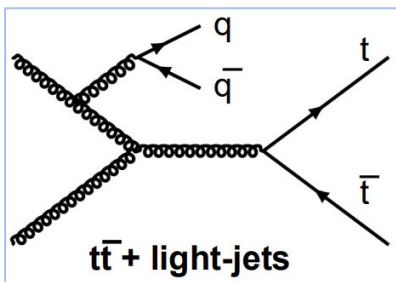
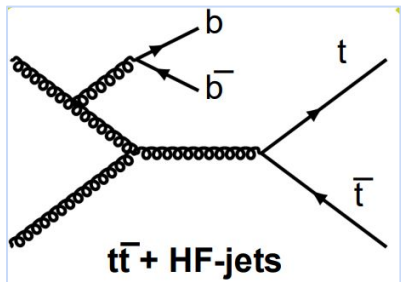
# Thank you!

Particle physicists when  
their latest measurement  
once again confirms The  
Standard Model

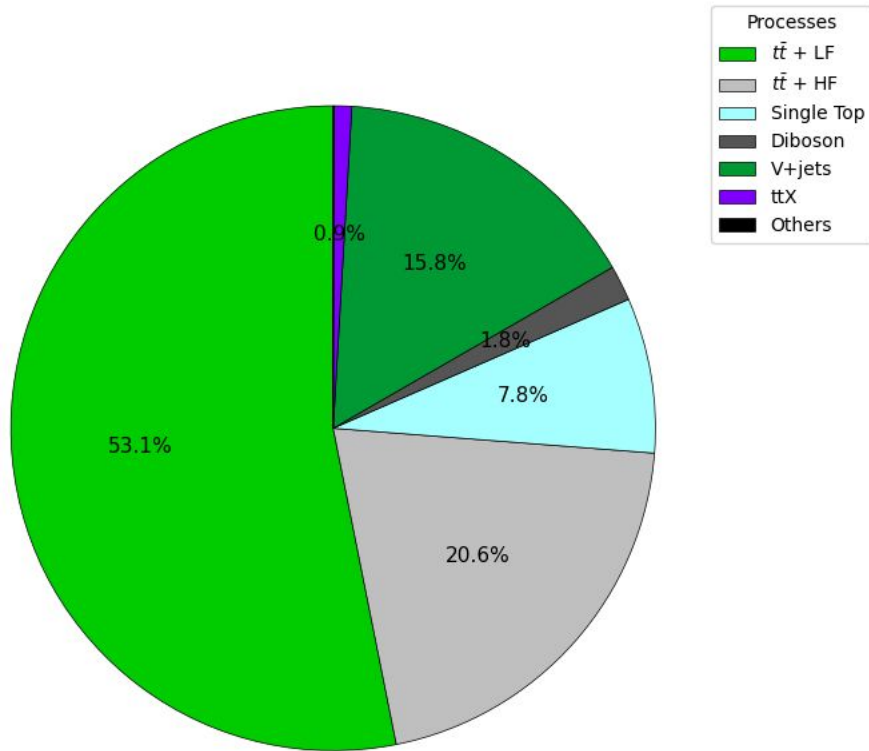


# Back-up

# Backgrounds



Background Composition



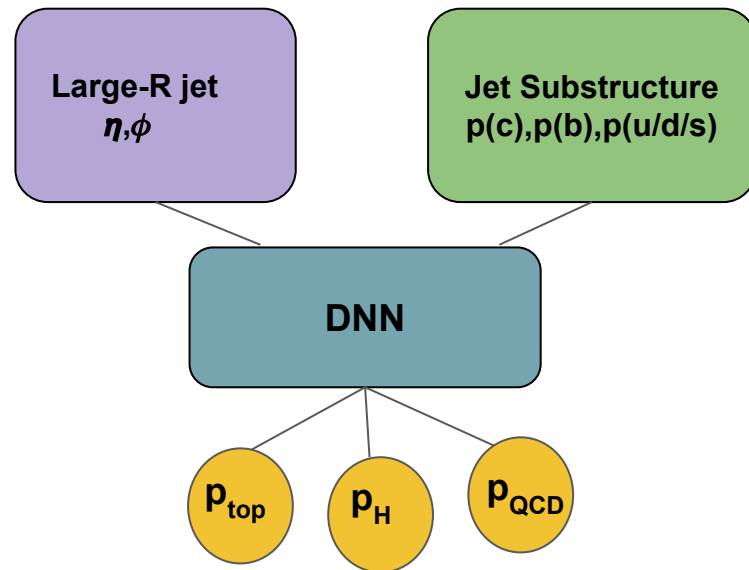
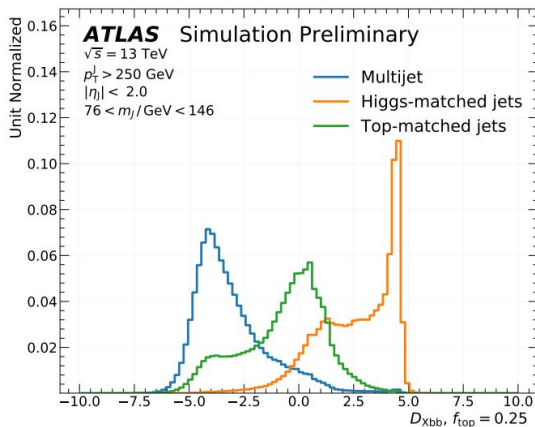
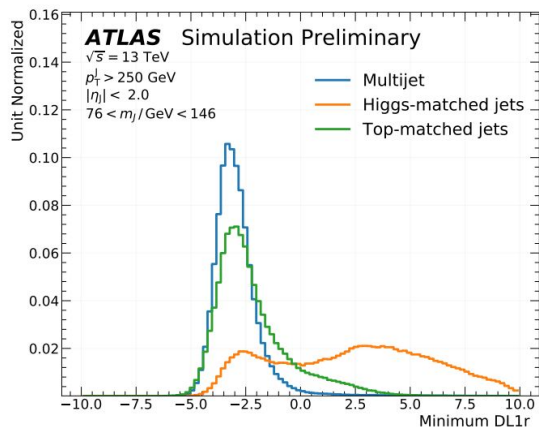
| Process          | qqbb weighted   | $\ell\nu bb$ weighted | qqbb events | $\ell\nu bb$ events | total events |
|------------------|-----------------|-----------------------|-------------|---------------------|--------------|
| Signal 1.2 TeV   | $4938 \pm 104$  | $15759 \pm 183$       | 29880       | 94806               | 124686       |
| Signal 1.4 TeV   | $4995 \pm 103$  | $14877 \pm 175$       | 31990       | 95576               | 127566       |
| Signal 1.6 TeV   | $4059 \pm 89$   | $11159 \pm 149$       | 26063       | 95545               | 121608       |
| Signal 1.8 TeV   | $5109 \pm 99$   | $13718 \pm 164$       | 33576       | 93854               | 127430       |
| Signal 2 TeV     | $4863 \pm 97$   | $13134 \pm 158$       | 34188       | 92409               | 126597       |
| Signal 2.5 TeV   | $4906 \pm 95$   | $12049 \pm 150$       | 34434       | 88119               | 122553       |
| Signal 3 TeV     | $4672 \pm 92$   | $11527 \pm 142$       | 33757       | 83396               | 117153       |
| V + jets         | $404 \pm 3$     | $5575 \pm 16$         | 336705      | 2215914             | 2552619      |
| diboson          | $59 \pm 2$      | $612 \pm 9$           | 1737        | 22597               | 24334        |
| others           | $2.03 \pm 0.05$ | $11.37 \pm 0.11$      | 19514       | 104074              | 123588       |
| singletop        | $293 \pm 6$     | $2658 \pm 19$         | 5173        | 48551               | 53724        |
| ttX              | $36.6 \pm 0.3$  | $301 \pm 0.8$         | 32308       | 291182              | 323490       |
| ttbar + HF jets  | $695 \pm 5$     | $7134 \pm 17$         | 35329       | 358682              | 394011       |
| ttbar + LF jets  | $1195 \pm 6$    | $18968 \pm 39$        | 243233      | 1383625             | 1626858      |
| Total background | $2683 \pm 11$   | $35260 \pm 51$        | 673999      | 4424625             | 5098624      |
| Data             | 2612            | 34793                 | 2612        | 34793               | 37405        |

# X-bb Tagger

For passing 60% working point:

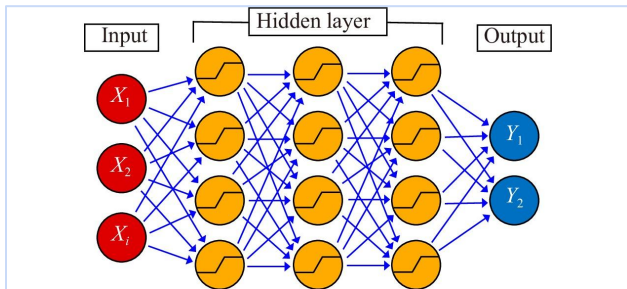
$$D = \ln \left( \frac{p_H}{f_{\text{top}} p_{\text{top}} + (1 - f_{\text{top}}) p_{\text{QCD}}} \right) > 2.44$$

$$f_{\text{top}} = 0.25$$

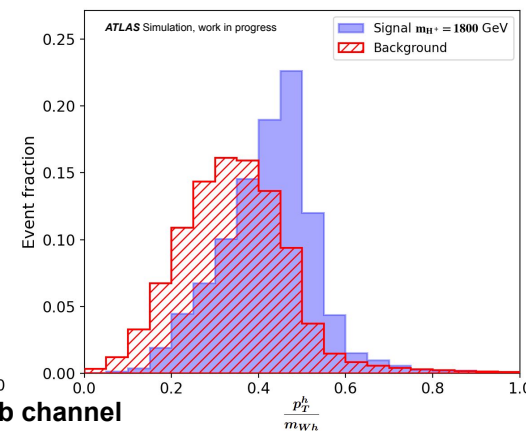
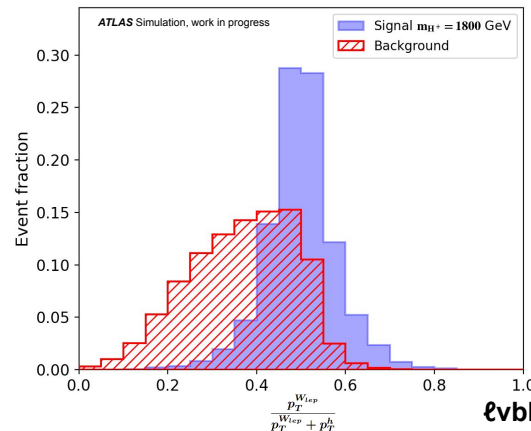
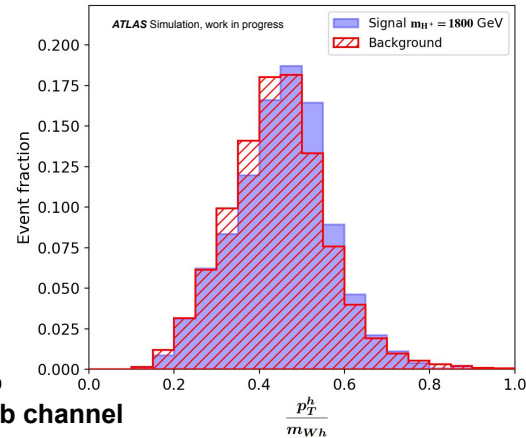
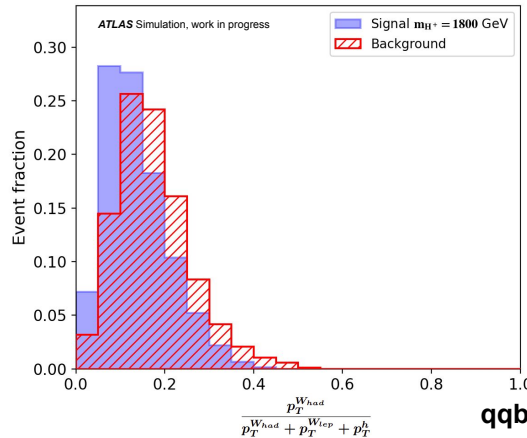


ATL-PHYS-PUB-2020-019

# DNN for event classification



- ❖ DNN for binary classification using 3 dense layers.
- ❖ Separate networks for qqbb and  $\ell$ vbb.
- ❖ All signal mass points and background trained together.
- ❖ Use of variables for example,  $p_{T,h}/m_{wh}$  and  $p_{T,W}/m_{wh}$ , to remove  $H^\pm$  mass dependence.
- ❖ 2-fold cross-validation using odd-even split.



# DNN input variables and setup

## qqbb variables

$$\frac{p_T^{W_{had}}}{p_T^{W_{had}} + p_T^{W_{lep}} + p_T^h}, \quad \Delta R(\ell, W_{had}), \quad \Delta R(\ell, h), \quad \Delta\eta(h, W_{had}), \quad \Delta\phi(h, W_{had})$$

$$\frac{p_T^{W_{had}}}{m_{WH^+}}, \quad \frac{p_T^h}{m_{WH^+}}$$

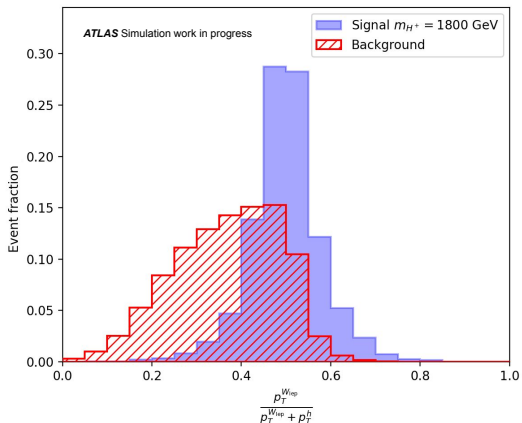
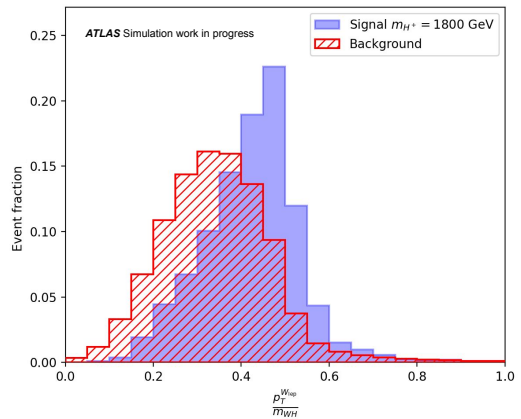
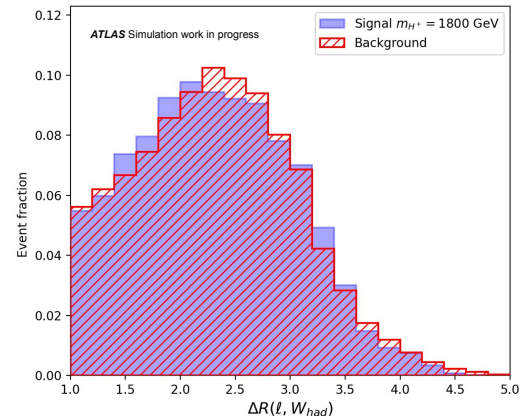
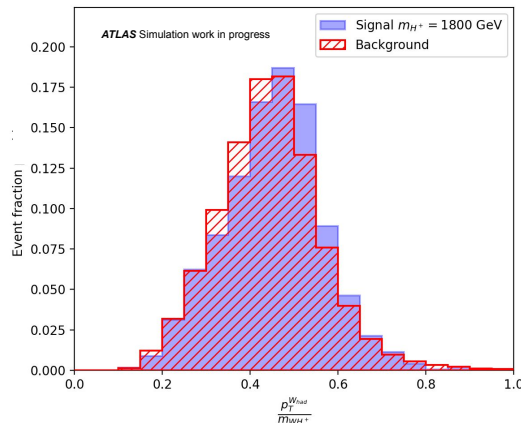
## ℓvbb variables

$$\frac{p_T^{W_{lep}}}{p_T^{W_{lep}} + p_T^h}, \quad m_{top}, \quad \Delta R(\ell, h), \quad \Delta\eta(h, W_{lep}), \quad \Delta\phi(h, W_{lep})$$

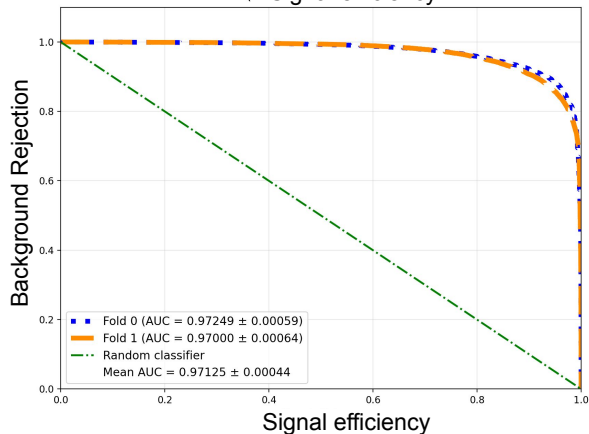
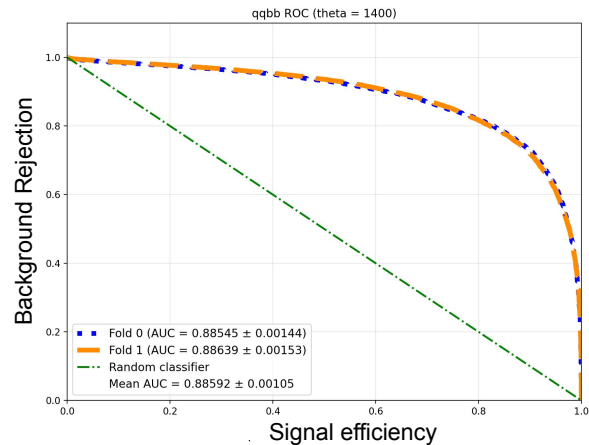
$$\frac{p_T^{W_{lep}}}{m_{WH}}, \quad \frac{p_T^h}{m_{WH}}$$

## Training setup:

- Keras
- Hidden layers:
  - lvbb:  $3 \times 128$  nodes
  - qqbb:  $2 \times 128$  nodes
- ReLU activation + sigmoid output (binary classification)
- L2 regularisation
- Dropout
- Adam optimiser
- Batch size: 128
- Early Stopping



# DNN performance

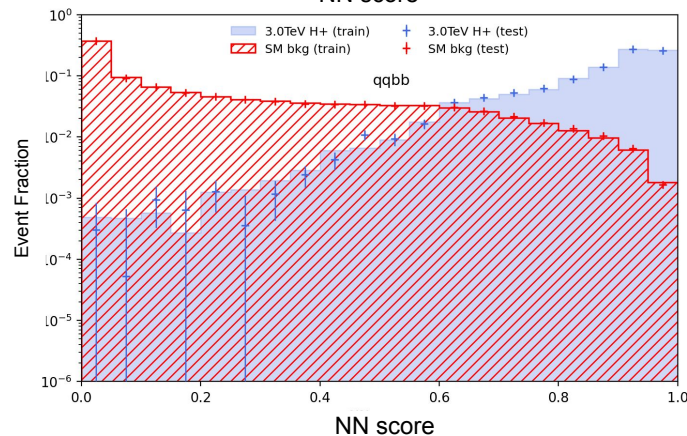
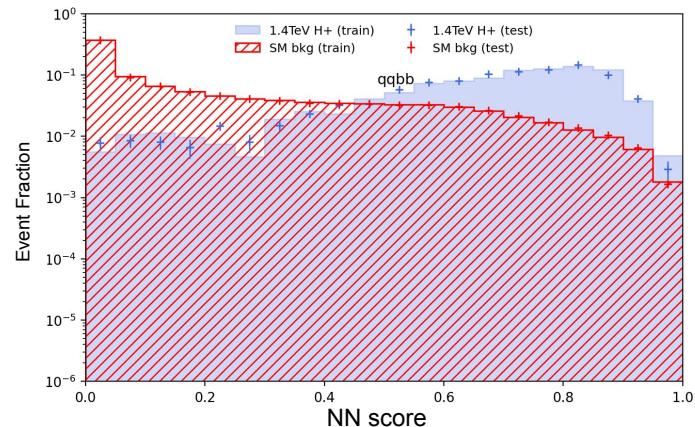


qqbb 1.4 TeV

- Performance becomes better with higher signal mass.
- Overtraining checked.

qqbb 3 TeV

Performance metric:  
Area under the curve (AUC)



\*AUC errors are bootstrapped errors

# Signal and Control Regions

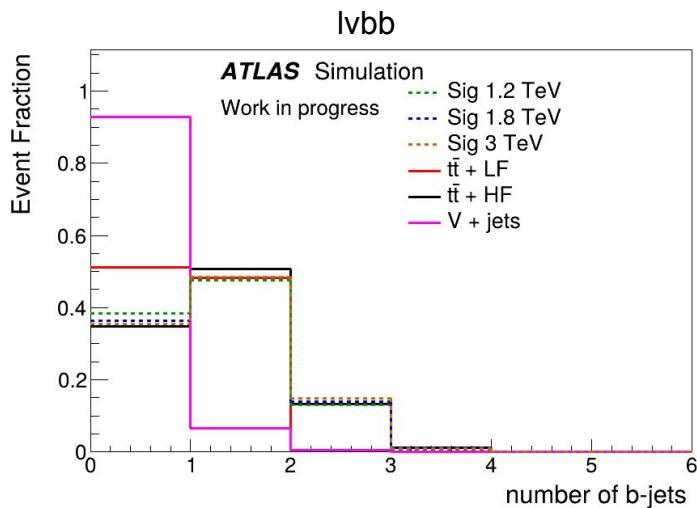
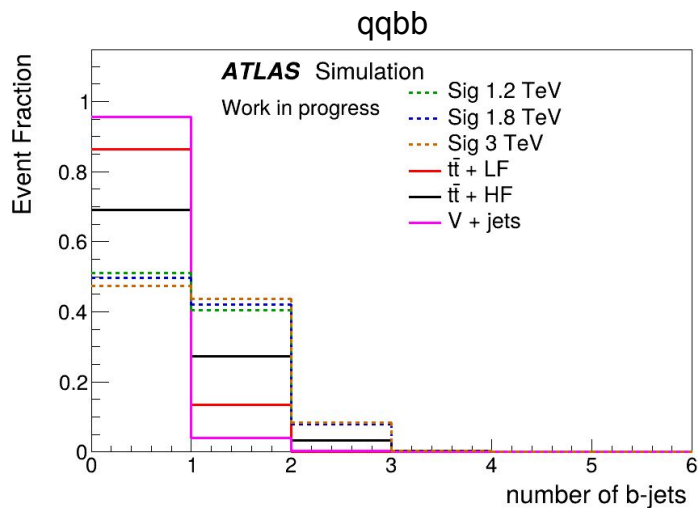
❖ Further division based on the number of small radius b-tagged jets outside Large-R jets.

❖ Additional division into two regions: **0 b-tagged,  $\geq 1$  b-tagged**.

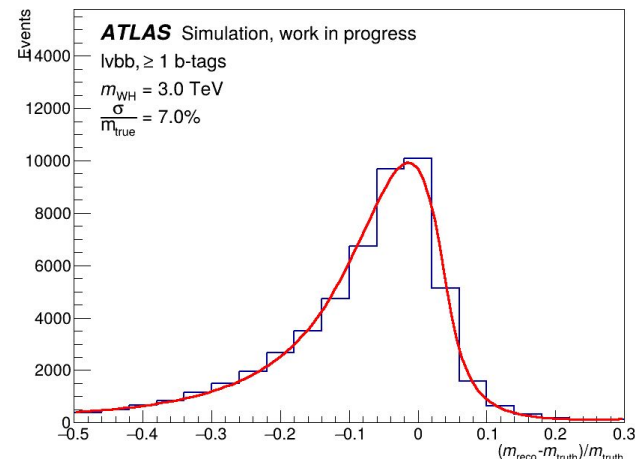
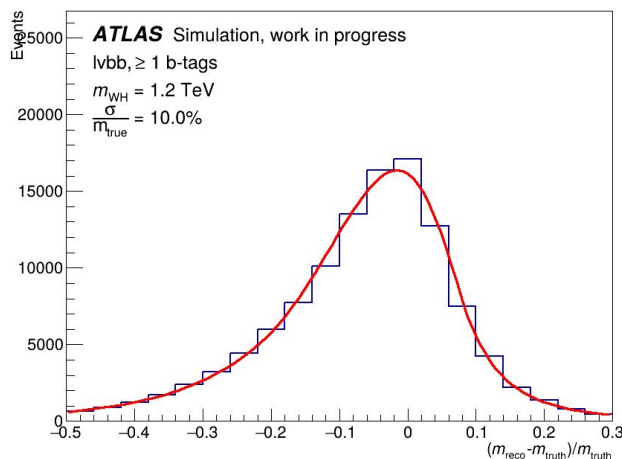
❖ **Signal regions** are further divided into



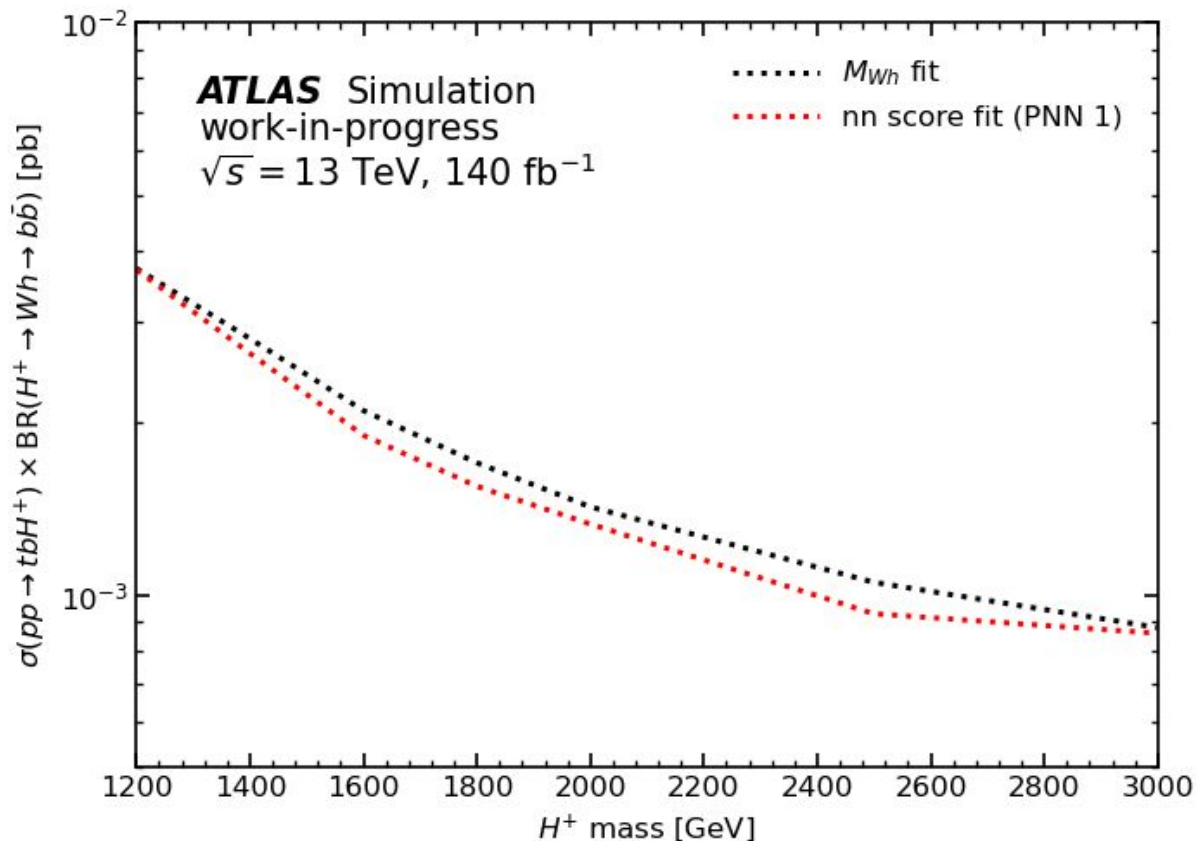
- High nn score
  - Medium nn score
  - Low nn score
- } after using a DNN for event S vs B classification



- Reconstructed charged higgs mass is used for binning
- Bukin fits applied to derive relative mass resolution using FWHM.
- Made sure that each bin was not finer than the corresponding resolution.



# Expected exclusion Limits (PNN 1)



- Slight improvement in sensitivity with **3%** (1.2 TeV) and **2%** (3 TeV).

# PNN-2 input feature strategy

- Proposed a large set of input variables (**30 for each Channel**).
- Permutation importance of each variable:  $\Delta\text{AUC} = \text{AUC}_{\text{Baseline}} - \text{AUC}_{\text{permuted}}$
- For many low-ranked features, the ranking kept changing across mass points.
- **Normalized  $\Delta\text{AUC}$**  for a feature across all features per mass point.

$$f(m, f) = \frac{\Delta\text{AUC}(m, f)}{\sum_{f'} \Delta\text{AUC}(m, f')} \quad \text{with} \quad \sum_f f(m, f) = 1$$

- Compared the **normalized  $\Delta\text{AUC}$**  of the feature across all mass points.

Baseline AUC [1.2 TeV,  
3 TeV]:

qqbb: [0.926, 0.956]

ttbb: [0.943, 0.971]

# PNN-2 input feature strategy

- Proposed a large set of input variables (**30 for each Channel**).

- Perm

- For r

- Norr**

- Com

| Quantity (variable: $m_{H^+}$ )     | 1.2 TeV | 1.8 TeV | 3 TeV  |
|-------------------------------------|---------|---------|--------|
| $\Delta\text{AUC}(m, f)$            | 0.13    | 0.06    | 0.006  |
| $\sum_{f'} \Delta\text{AUC}(m, f')$ | 0.30    | 0.107   | 0.0139 |
| $f(m, f)$                           | 43%     | 56%     | 43%    |

| Quantity (variable: $\eta_{W_{qq}}$ ) | 1.2 TeV              | 1.8 TeV              | 3 TeV                |
|---------------------------------------|----------------------|----------------------|----------------------|
| $\Delta\text{AUC}(m, f)$              | $1.9 \times 10^{-3}$ | $2.9 \times 10^{-5}$ | $5.1 \times 10^{-7}$ |
| $\sum_{f'} \Delta\text{AUC}(m, f')$   | 0.30                 | 0.107                | 0.0139               |
| $f(m, f)$                             | 0.064%               | 0.027%               | 0.0037%              |

Baseline AUC [1.2 TeV, 3 TeV]:

qqbb: [0.926, 0.956]

ℓℓbb: [0.943, 0.971]

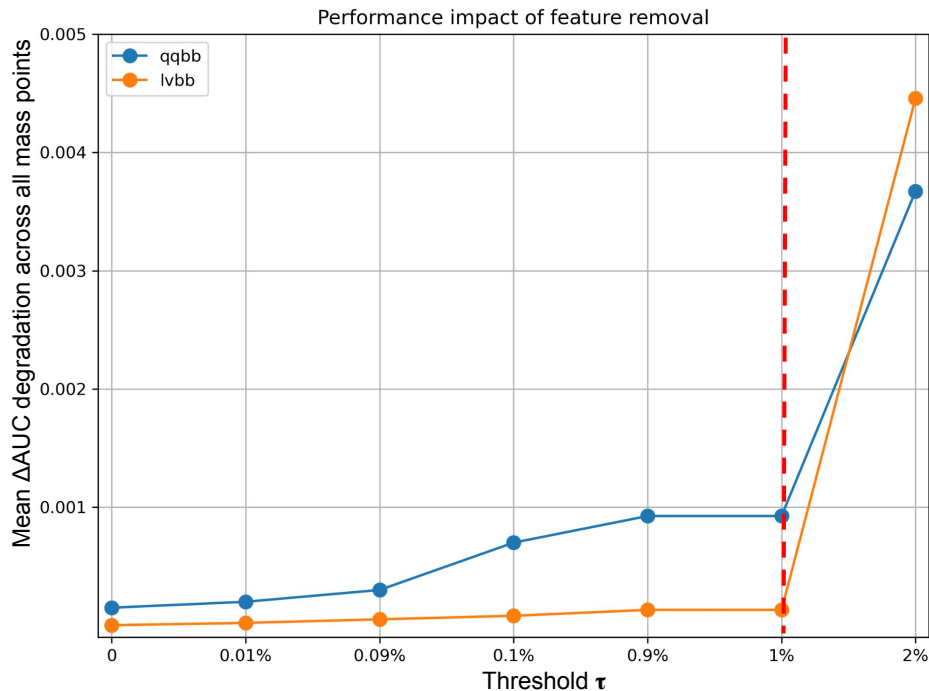
oints.

# Threshold scan

- Scanned grid of thresholds  $\tau$  (normalized  $\Delta\text{AUC}$ ) = {0.01%, 0.09%, 0.1%, 0.9%, 1%, 2%}
- Chosen threshold = 1% (Helped getting rid of many low ranked variables)

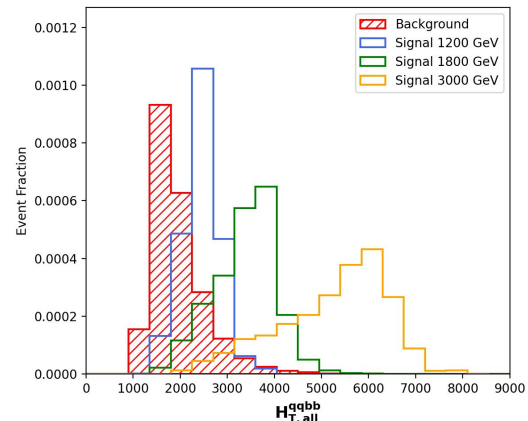
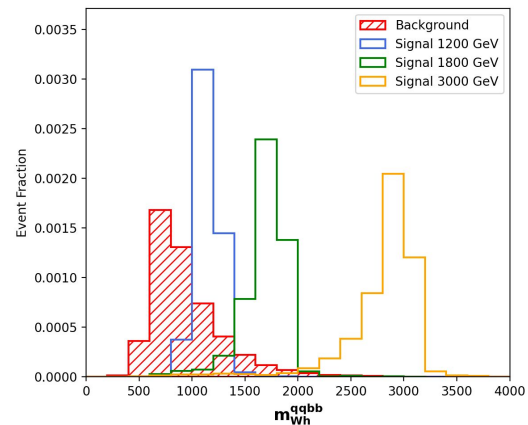
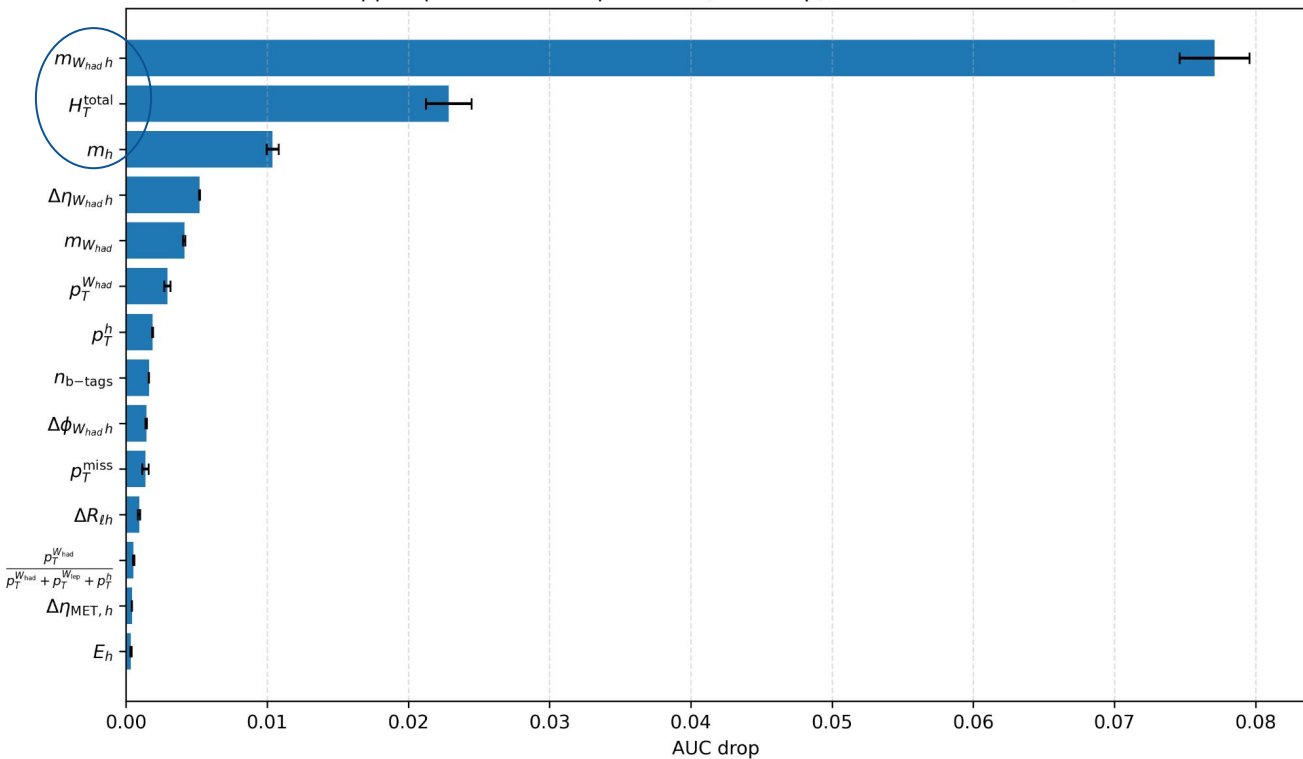
| Channel                      | Number of features dropped |
|------------------------------|----------------------------|
| <b>qqbb</b>                  | <b>15</b>                  |
| <b><math>\ell vbb</math></b> | <b>11</b>                  |

| Channel                      | Final number of features |
|------------------------------|--------------------------|
| <b>qqbb</b>                  | <b>14</b>                |
| <b><math>\ell vbb</math></b> | <b>19</b>                |



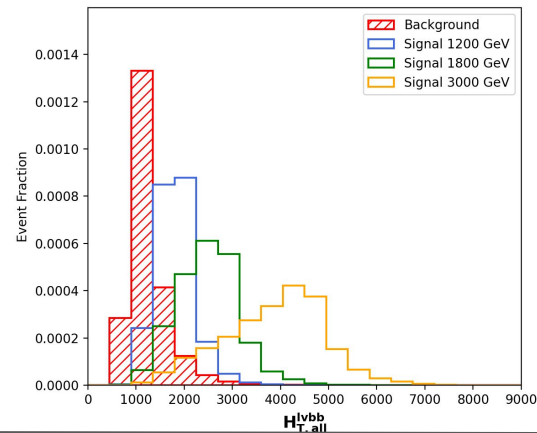
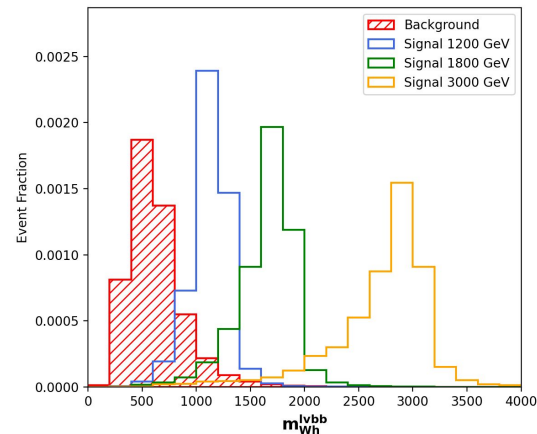
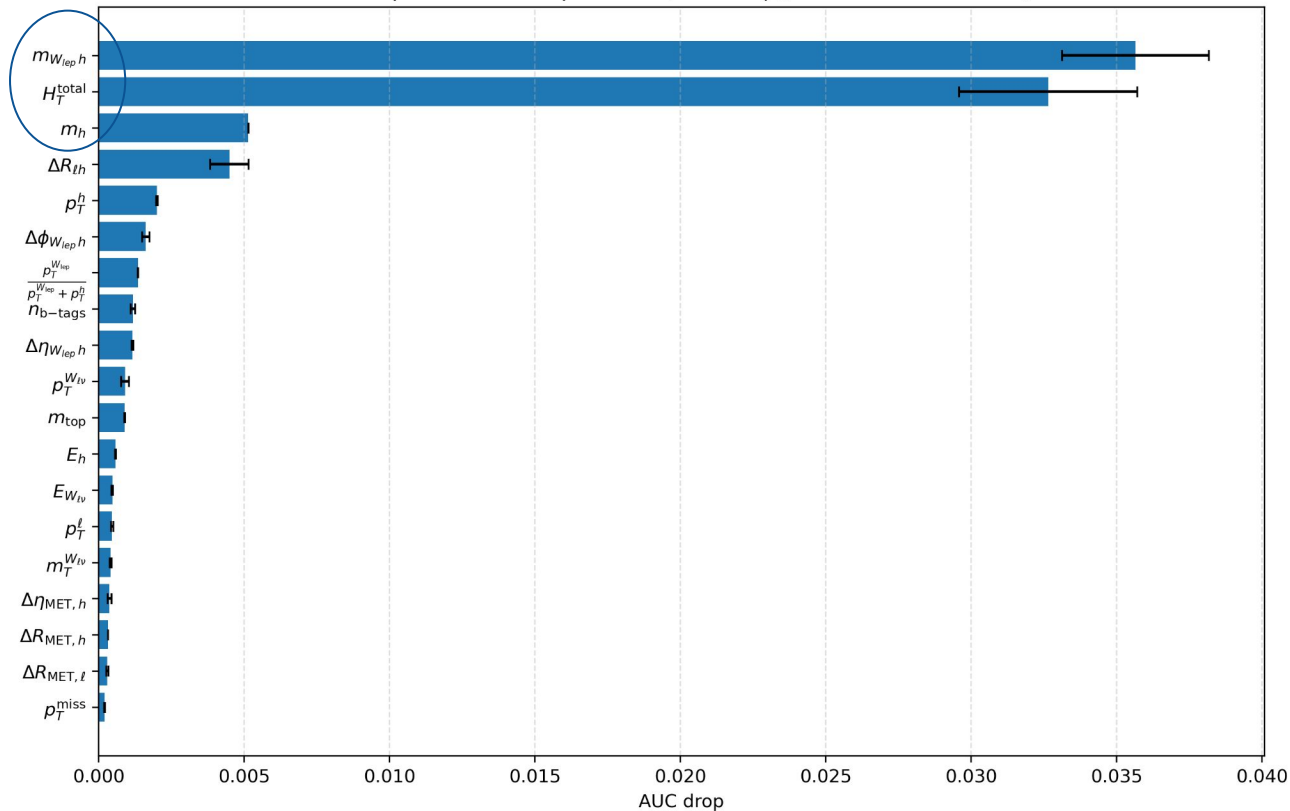
# PNN-2 Feature importance (qqbb)

qqbb: permutation importance (AUC drop), evaluated at 1.8 TeV

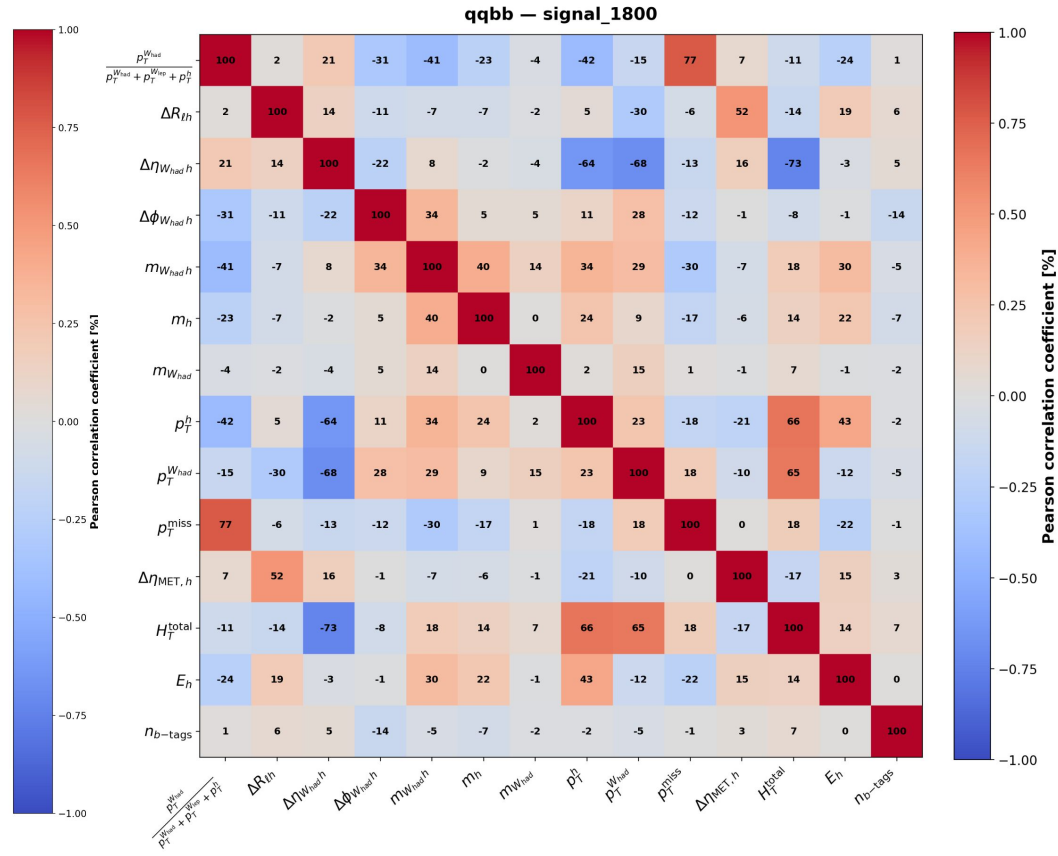
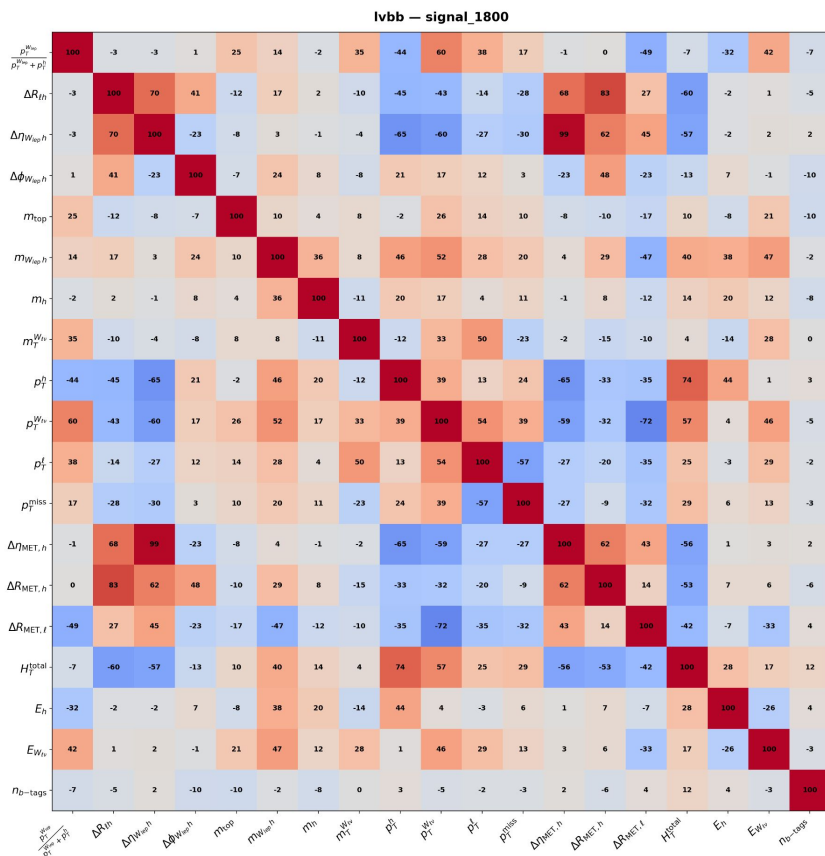


# PNN-2 Feature importance ( $\ell\nu b\bar{b}$ )

lvbb: permutation importance (AUC drop), evaluated at 1.8 TeV



# PNN-2 Correlation Matrices (1.8 TeV)



# PNN-2 variables and optimized setup

## qqbb variables

$$\Delta R_{\ell h}, \quad \Delta \eta_{W_{had}h}, \quad \Delta \phi_{W_{had}h}, \quad m_{W_{had}h}, \quad m_h, \quad m_{W_{had}},$$

$$p_T^h, \quad p_T^{W_{had}}, \quad p_T^{\text{miss}}, \quad \Delta \eta_{\text{MET},h}, \quad H_T^{\text{total}}, \quad E_h, \quad n_{\text{b-tags}}, \quad \frac{p_T^{W_{had}}}{p_T^{W_{had}} + p_T^{W_{lep}} + p_T^h}$$

## ℓvbb variables

$$\Delta R_{\ell h}, \quad \Delta \eta_{W_{lep}h}, \quad \Delta \phi_{W_{lep}h}, \quad m_{\text{top}}, \quad m_{W_{lep}h}, \quad m_h, \quad m_T^{W_{\ell\nu}},$$

$$p_T^h, \quad p_T^{W_{\ell\nu}}, \quad p_T^\ell, \quad p_T^{\text{miss}}, \quad \Delta \eta_{\text{MET},h}, \quad \Delta R_{\text{MET},h}, \quad \Delta R_{\text{MET},\ell},$$

$$H_T^{\text{total}}, \quad E_h, \quad E_{W_{\ell\nu}}, \quad n_{\text{b-tags}}, \quad \frac{p_T^{W_{lep}}}{p_T^{W_{lep}} + p_T^h},$$

### • Architecture (Best from Optuna scan)

#### – qqbb:

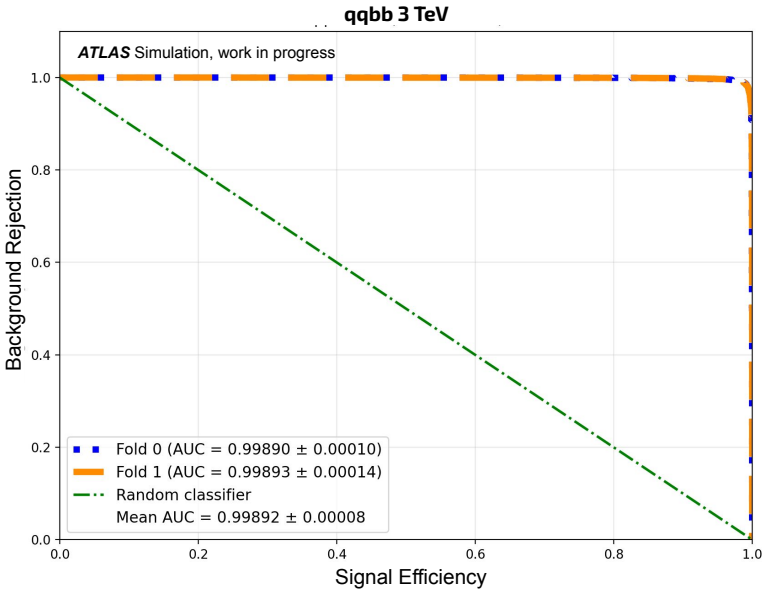
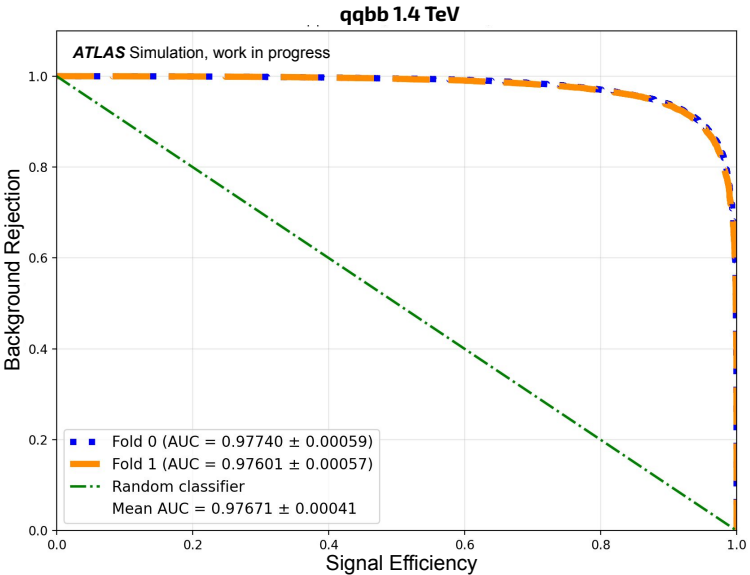
- 4 hidden layers: 512 → 256 → 256 → 256
- Dropout: 0.3
- L2:  $0.7 \times 10^{-6}$
- Learning rate: 0.0001
- Batch size: 64
- AUC: 0.966

### • Architecture (Best from Optuna scan)

#### – ℓvbb:

- 4 hidden layers: 512 → 256 → 256 → 256
- Dropout: 0.1
- L2:  $1.05 \times 10^{-6}$
- Learning rate: 0.0001
- Batch size: 64
- AUC: 0.986

# PNN-2 Performance



## Interpolation:

| Signal mass dropped | 1.8 TeV | 2.5 TeV |
|---------------------|---------|---------|
| Drop in AUC         | 0.0013% | 0.0018% |

NN Score

# Basics of Statistical Analysis

$$v_{ir}(\mu, \theta) = \mu s_{ir}(\theta) + b_{ir}(\theta)$$

Number of signal and background events in the bin



$$\mathcal{L}(\mu, \theta) = \prod_r \prod_i P(n_{ir} | \mu s_{ir}(\theta) + b_{ir}(\theta)) \prod_k \pi(\theta_k)$$

Likelihood function



$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}$$

Profile Likelihood ratio



$$q_\mu = -2 \ln \lambda(\mu)$$

Test Statistic



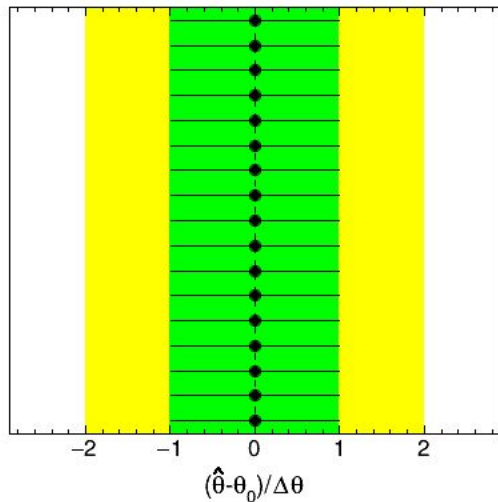
$$CL_s = \frac{CL_{s+b}}{1 - CL_b}$$

For BSM searches, first it is tested whether the data is compatible with SM background-only hypothesis. If no significant excess is seen, no discovery is claimed. Instead an upper limit on the allowed signal rate at 95% confidence Level excluding signal models that predict a rate above this limit

# PNN-2 systematics

After S+B blinded fit

**ATLAS** SimulationWorkinprogress



Instrumental

Pileup reweighting  
 Muon trigger SF (syst)  
 Muon trigger SF (stat)  
 Muon TTVA SF (syst)  
 Muon TTVA SF (stat)  
 Muon isolation SF (syst)  
 Muon isolation SF (stat)  
 Muon ID SF low-pT (syst)  
 Muon ID SF (syst)  
 Muon ID SF low-pT (stat)  
 Muon ID SF (stat)  
 Luminosity  
 JVT  
 Electron trigger SF  
 Electron reco SF  
 Electron isolation SF  
 Electron ID SF

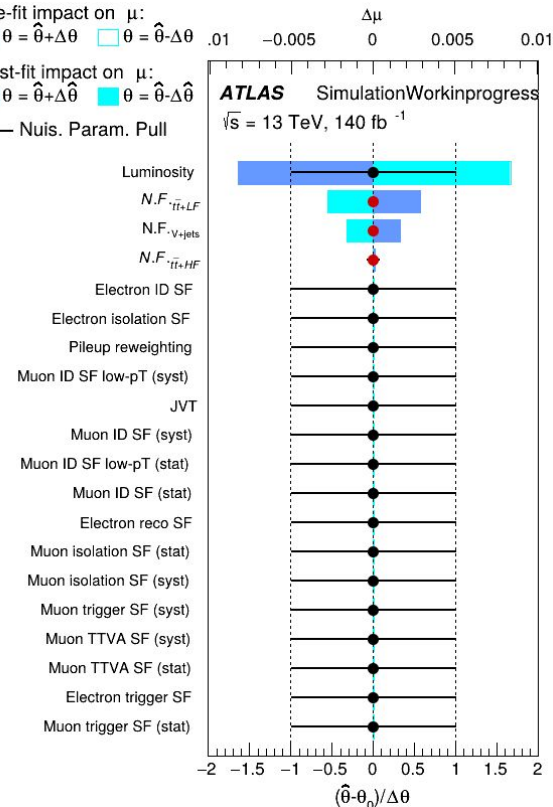
Pre-fit impact on  $\mu$ :

$\theta = \hat{\theta} + \Delta\theta$   $\theta = \hat{\theta} - \Delta\theta$

Post-fit impact on  $\mu$ :

$\theta = \hat{\theta} + \Delta\hat{\theta}$   $\theta = \hat{\theta} - \Delta\hat{\theta}$

— Nuis. Param. Pull



# PNN-2 systematics

ATLAS Simulation Working progress

|                          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Electron ID SF           | 100.0 | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   |
| Electron isolation SF    | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   |
| Electron reco SF         | -0.0  | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  |
| Electron trigger SF      | -0.0  | -0.0  | -0.0  | 100.0 | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   |
| JVT                      | 0.0   | -0.0  | -0.0  | 0.0   | 100.0 | 0.0   | 0.0   | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | -0.0  | 0.0   |
| Luminosity               | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -29.1 | -12.1 | -24.7 |
| Muon ID SF (stat)        | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon ID SF low-pT (stat) | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon ID SF (syst)        | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  |
| Muon ID SF low-pT (syst) | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon isolation SF (stat) | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | 100.0 | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  |
| Muon isolation SF (syst) | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon TTVA SF (stat)      | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | 100.0 | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  |
| Muon TTVA SF (syst)      | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon trigger SF (stat)   | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | 0.0   | 0.0   | 0.0   | 0.0   |
| Muon trigger SF (syst)   | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 100.0 | 0.0   | 0.0   | 0.0   |
| Pileup reweighting       | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | -0.0  | 100.0 | 0.0   | 0.0   |
| $N F_{V_{\text{jets}}}$  | 0.0   | 0.0   | -0.0  | 0.0   | 0.0   | -29.1 | 0.0   | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | 0.0   | 0.0   | 100.0 | 43.5  | -47.7 |
| $N F_{\text{ID, HF}}$    | 0.0   | 0.0   | -0.0  | 0.0   | 0.0   | -12.1 | 0.0   | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | -0.0  | 0.0   | 0.0   | 0.0   | 43.5  | 100.0 | -88.6 |
| $N F_{\text{ID, LF}}$    | 0.0   | 0.0   | 0.0   | 0.0   | -0.0  | -24.7 | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   | -0.0  | 0.0   | 0.0   | -0.0  | -47.7 | -88.6 | 100.0 |
| $\mu$                    | 0.0   | 0.0   | -0.0  | 0.0   | 0.0   | -54.7 | 0.0   | 0.0   | -0.0  | 0.0   | -0.0  | -0.0  | -0.0  | -0.0  | 0.0   | 0.0   | 10.9  | 0.5   | 18.8  |