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

Satyajit Chakraborty

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

57th Herbstschule HEP, 2026

10th September, 2026

# Overview

- ❑ 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)
- ❑ Final PNN limit using Systematics

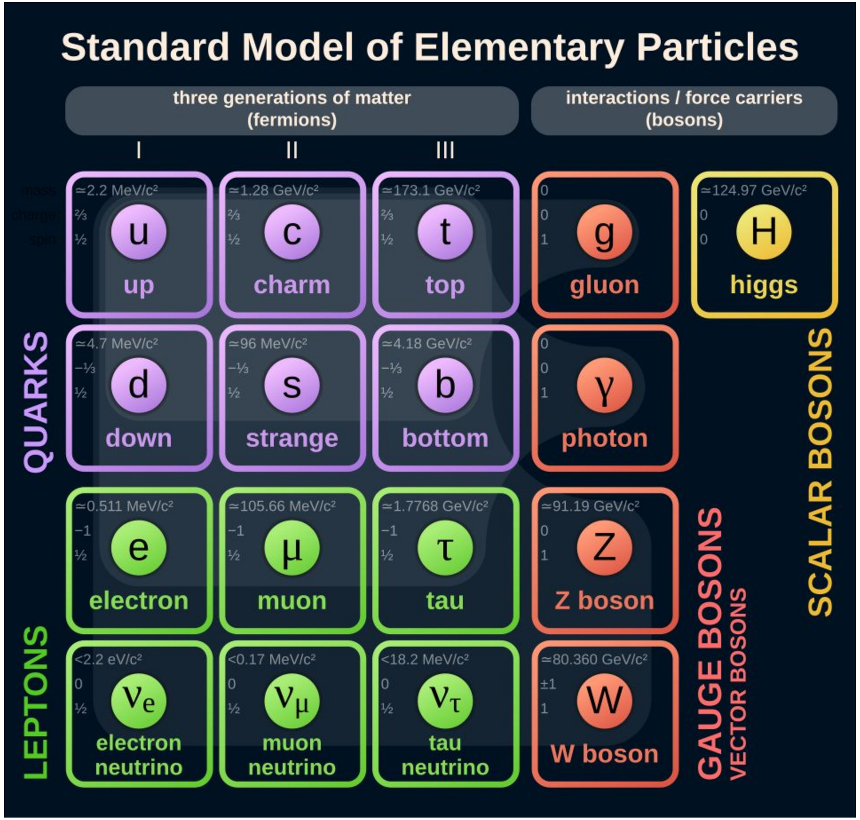
Following previous ATLAS analysis: arXiv:2411.03969

# 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) 

Limitations:

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

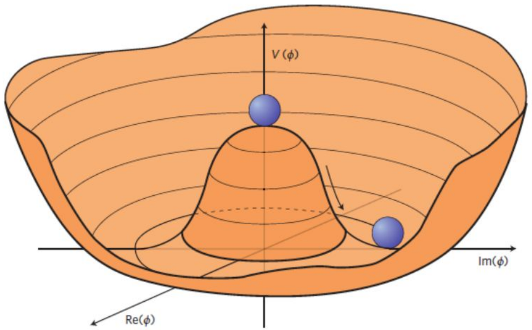



Wikimedia, Standard Model, (2016), url: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg

# Higgs and extended scalers

## Standard Model

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




$\mu$<0, <$\phi$>≠0
-> Spontaneous Symmetry breaking


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

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

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

## 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 $\phi_2$

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

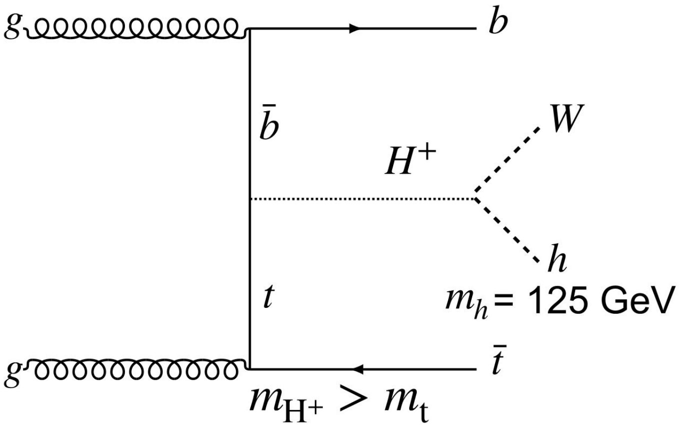
$$\langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_1 \end{pmatrix} \qquad \langle\Phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_2 \end{pmatrix}$$

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

$$\underbrace{v^2 = {v_1}^2 + {v_2}^2 = 246 \text{ GeV}}_{\text{CP-conserving one in this analysis}}$$

# 2HDM and charged Higgs

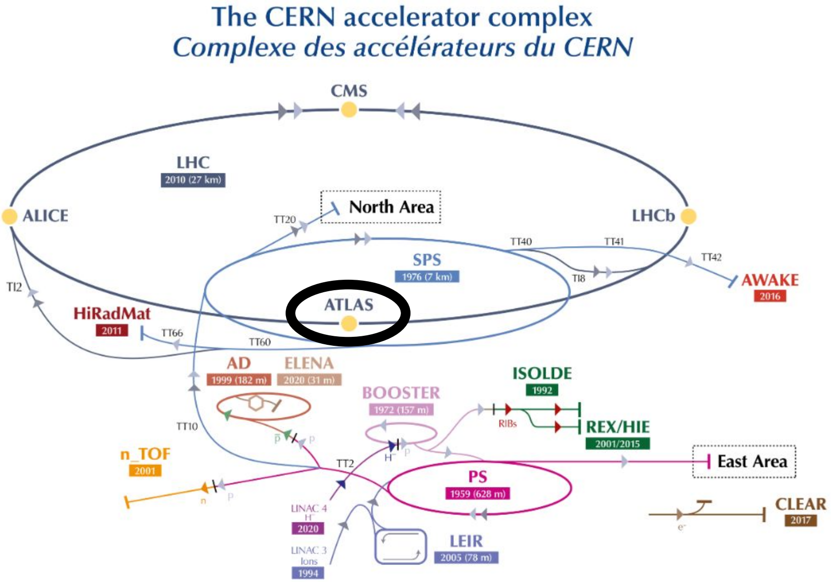
- 2HDM leading to 5 physical states after EWSB:
  - h(light),
  - H(heavy), } (CP-Even)
  - A (CP-Odd)
  - **H± (charged Higgs Bosons).**



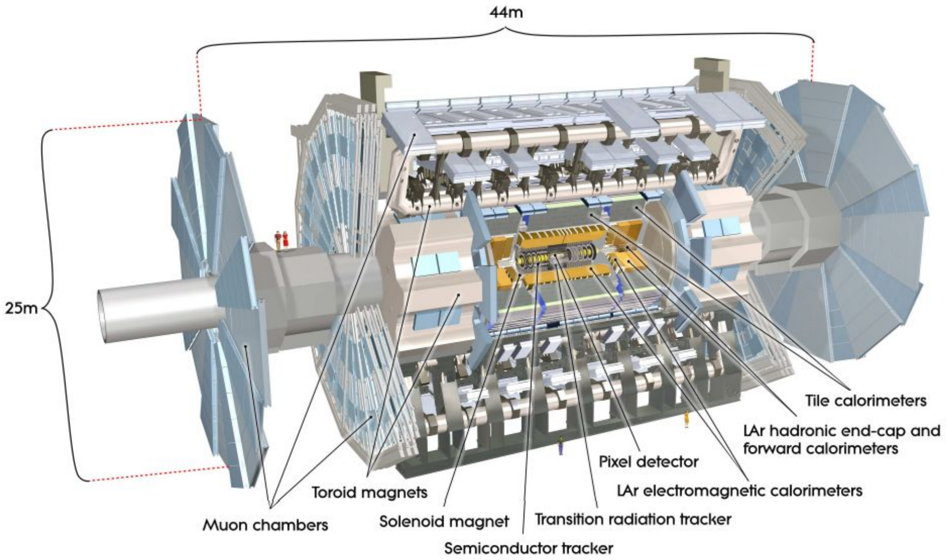

This search

# LHC and the ATLAS experiment

Cern Control Centre, url: https://home.cern/tags/control-centre



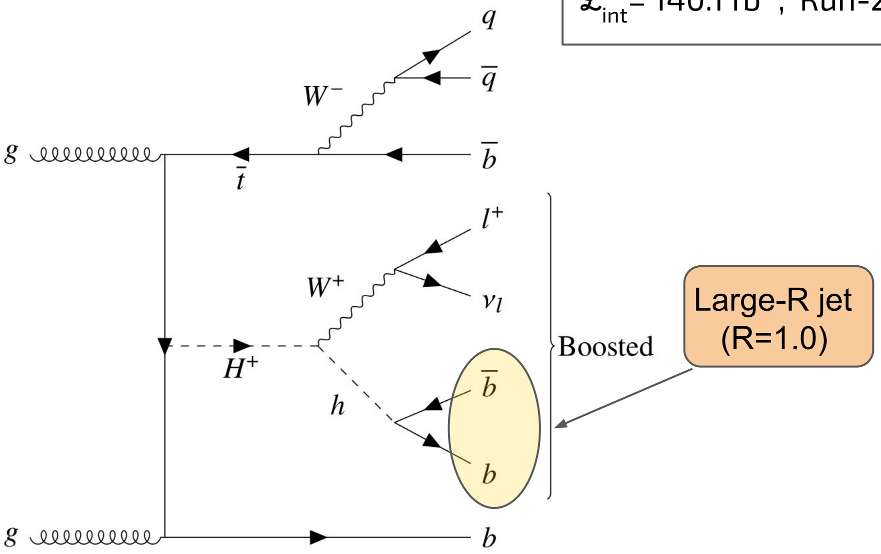

The ATLAS Collaboration *et al* 2008 *JINST* 3 S08003



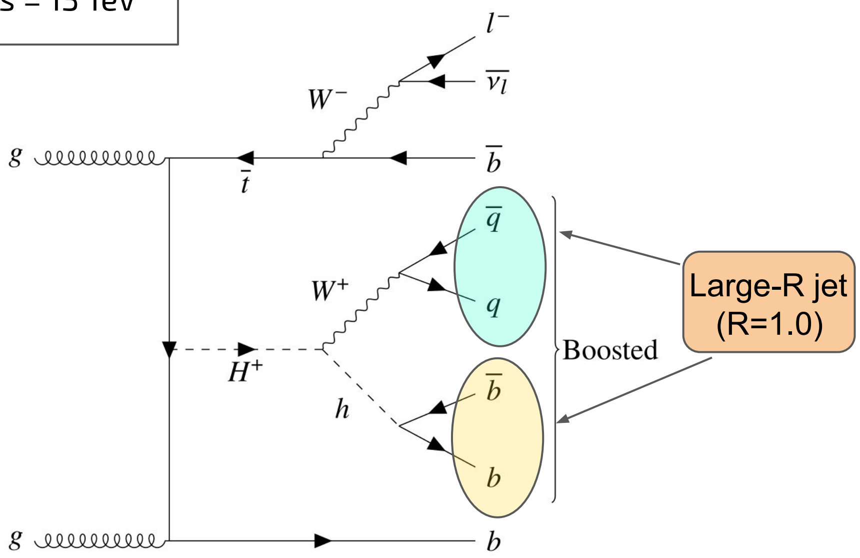

- ❖ The LHC is a circular particle collider with 27 km circumference.
- ❖ Proton-Proton and heavy ion collisions.
- ❖ The ATLAS experiment studies a wide range of physics phenomenons.
- ❖ Layered detector with inner trackers, calorimeters and muon spectrometer.

# Signal Topology

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




**ℓvbb channel**




**qqbb channel**

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

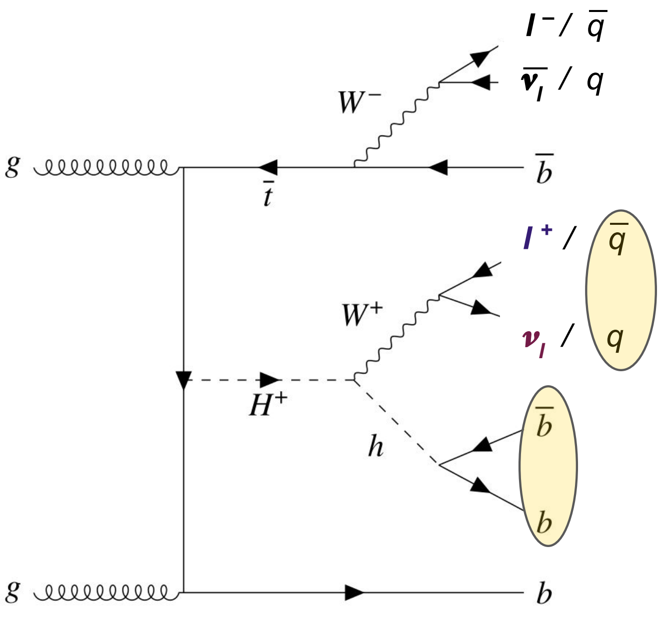
**Leading backgrounds: $t\bar{t}$ + LF jets, $t\bar{t}$ + HF jets, W/Z + jets**

# Event selection and categorization

## Event Selection:

B-tagging working point: 77%

- **Trigger:** pass at least one of the single e/μ triggers.
- **Missing transverse momentum:** $E_T^{miss}$ > 30 GeV
- **Leptons:** Exactly one e or μ with $p_T$ > 27 GeV
- **Large-R jets** ≥ 1
- **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 ΔR>1.4 from the **Large-R jet**



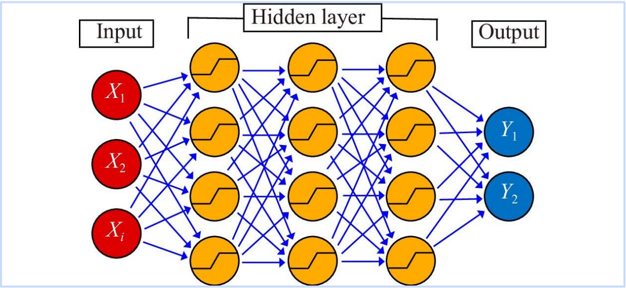

## Categorization:

| Channel | Large-R jets | W boson candidate | Higgs candidate |
| --- | --- | --- | --- |
| ***qqbb*** | **≥ 2** | ***Fails $X \rightarrow b\bar{b}$ tagging 50 GeV ≤ $M_J$ ≤ 110 GeV*** | $X \rightarrow b\bar{b}$ |
| ***ℓvbb*** | **≥ 1** | $W_{lep}$ | $X \rightarrow b\bar{b}$ |

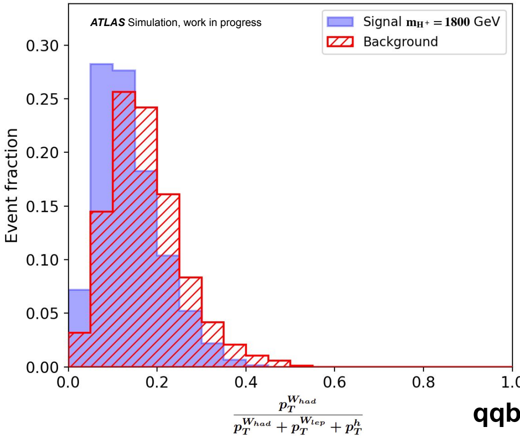
* $M_J$: Jet Mass

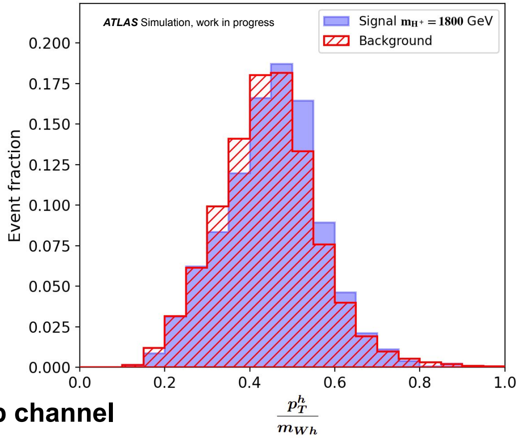
*ATL-PHYS-PUB-2020-019

# DNN for event classification

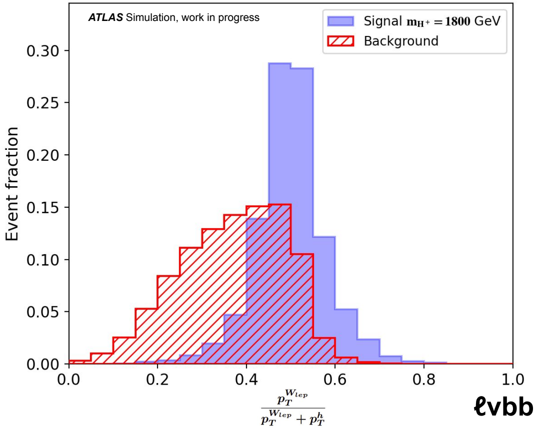


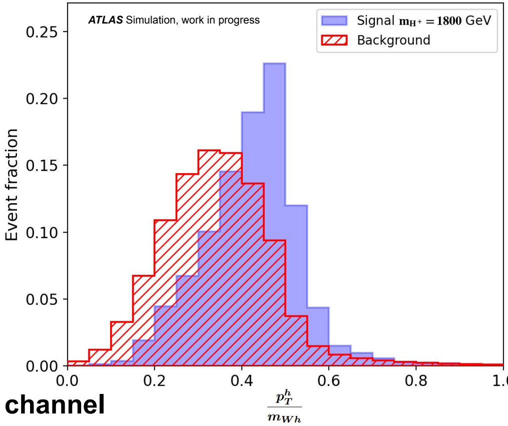

- DNN for binary classification using 3 dense layers.
- Separate networks for qqbb and ℓvbb.
- All signal mass points and background trained together.
- Use of variables for example, $p_{T,h}/m_{wh}$ and $p_{T,W}/m_{wh}$, to remove **H±** mass dependence.
- 2-fold cross-validation using odd-even split.




**qqbb channel**



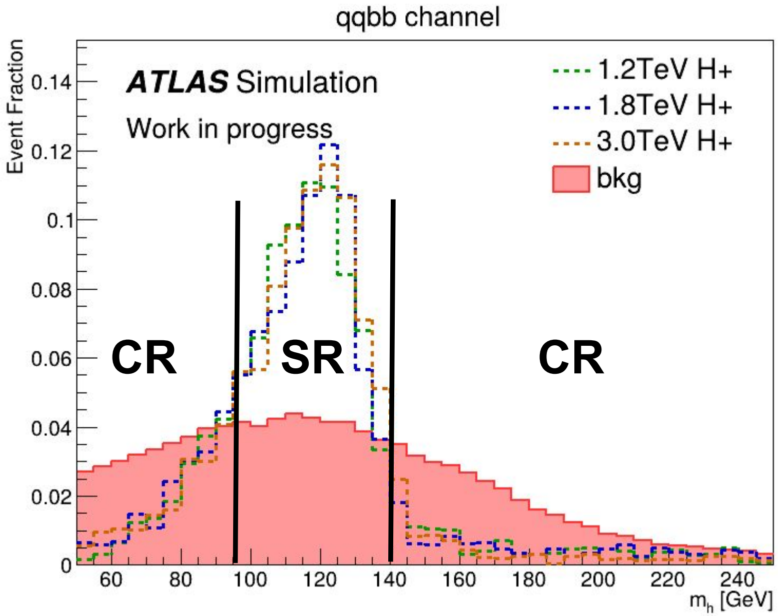



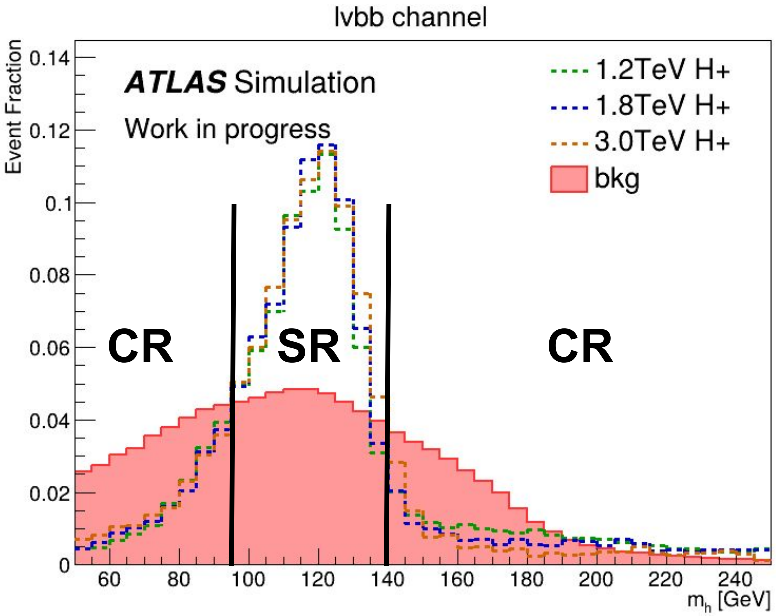

**ℓvbb channel**


Shao et al., IJMMM 31 (2024) 106-117

# Signal and Control Regions

- **Signal region**: 95 GeV ≤ $m_h$ ≤ 140 GeV
- **Control region**:
  - **Lower sideband**: $m_h$ < 95 GeV
  - **Higher sideband:** $m_h$ > 140 GeV



# Signal and Control Regions

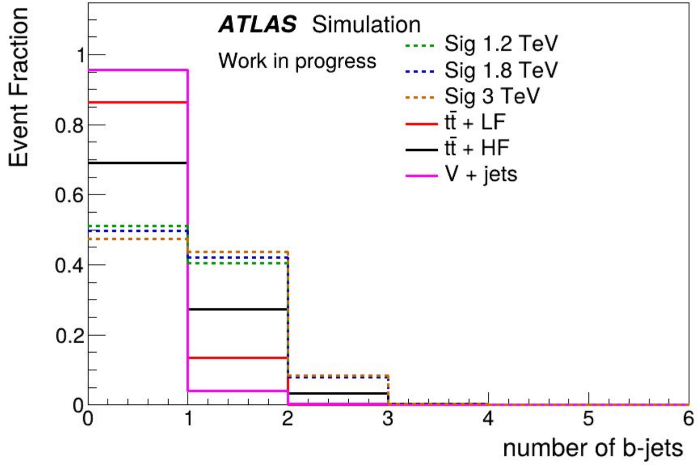
- Further division based on the number of small radius b-tagged jets outside Large-R jets.
- Additional division into two regions: **0 b-tagged, ≥1 b-tagged.**
- **Signal regions** are further divided into
  - High nn score
  - Medium nn score
  - Low nn score

**Total**

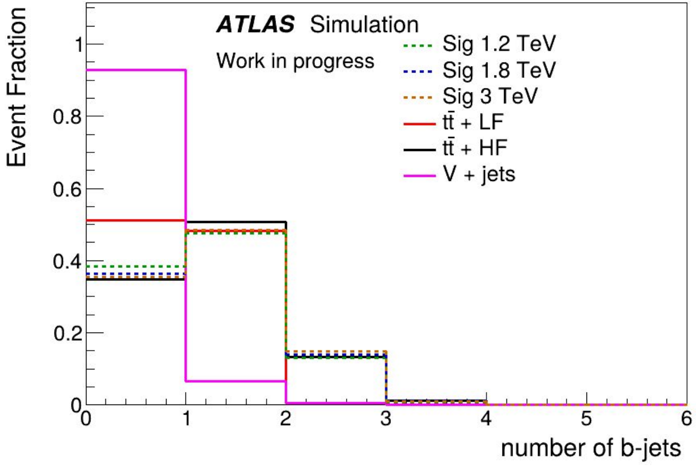
**Control regions** = 8

**Signal regions** = 10

qqbb

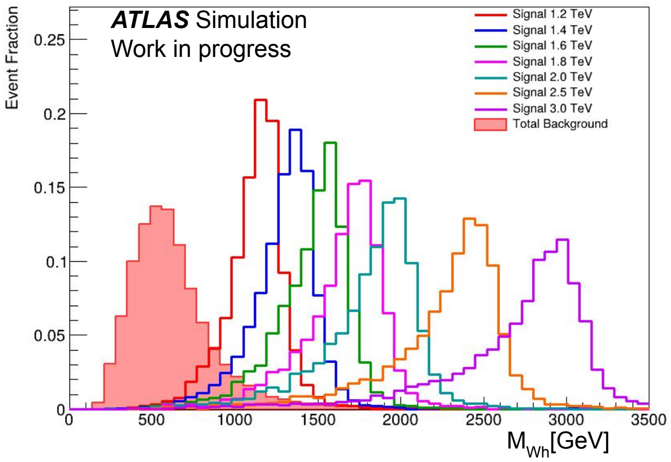



lvbb

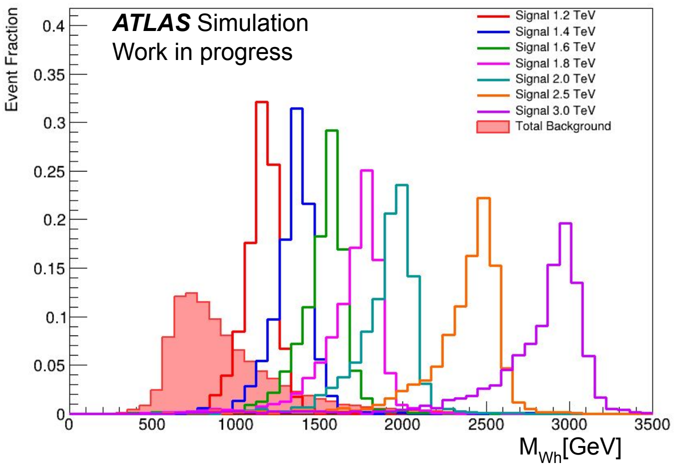


# $M_{Wh}$ as a fit variable

**ℓvbb channel**




**qqbb channel**



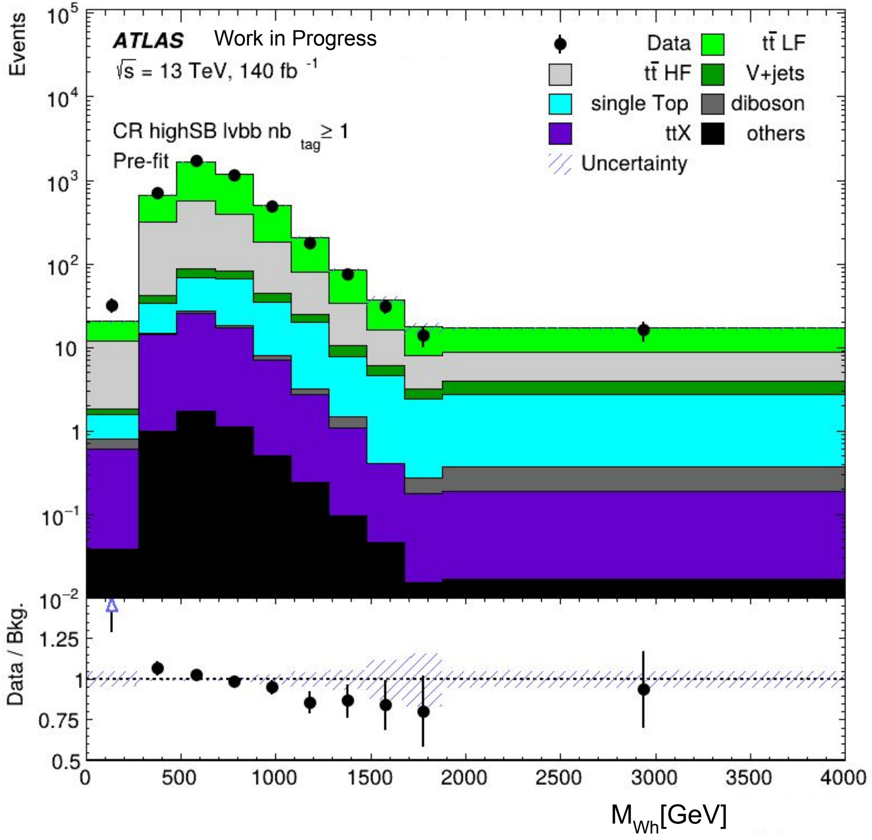

**Clear separation between the signal distributions and the total background**

- Normalization factors for **$t\bar{t}$ + LF, $t\bar{t}$ + HF , W/Z + 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 ($\sigma \times$ BR).**
- Both **qqbb** and **ℓvbb** channels included in the fit setup.

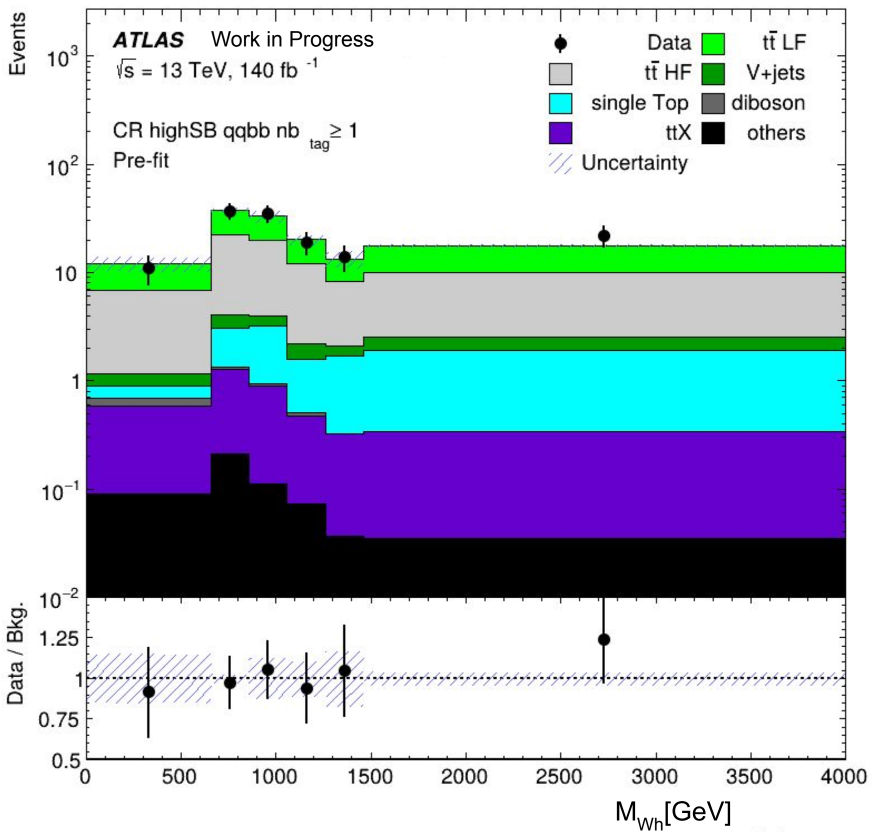
**Fit Setup**

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

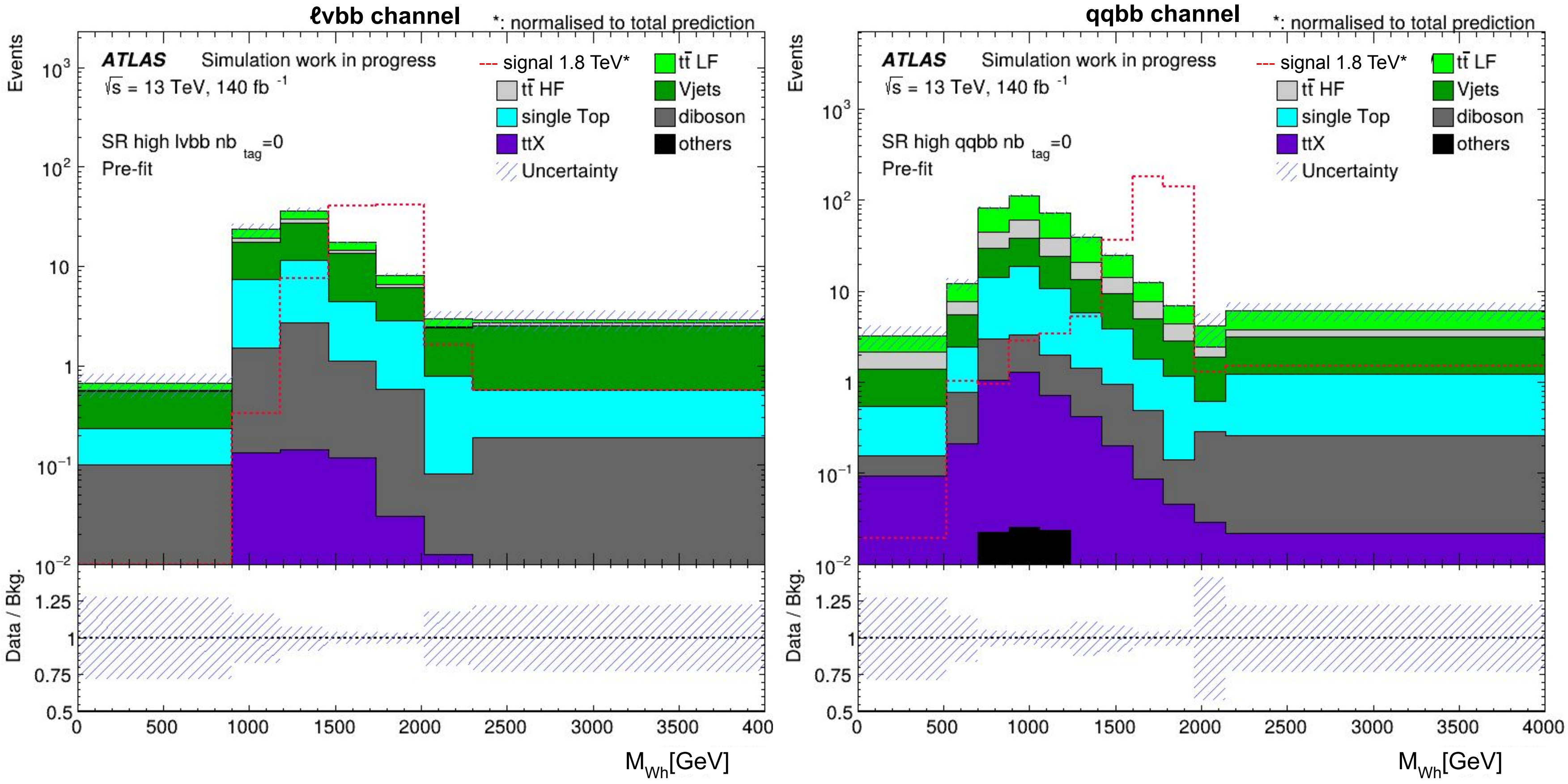
**ℓvbb channel**



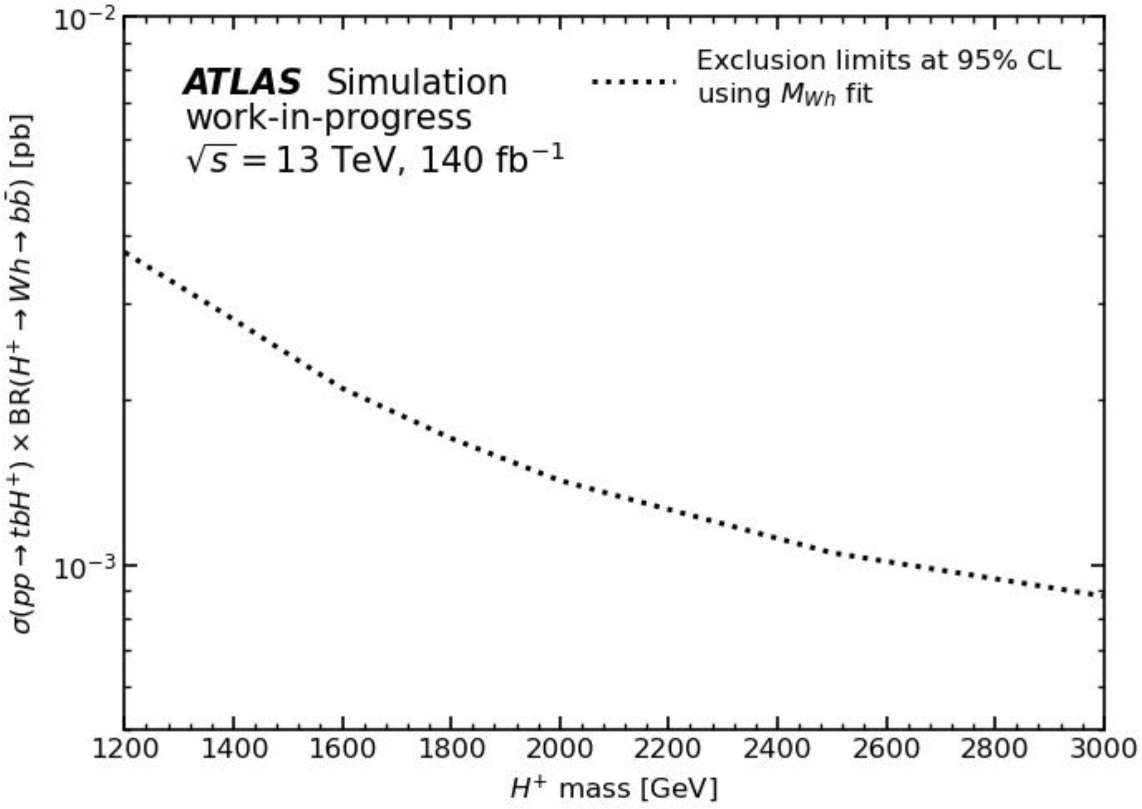

**qqbb channel**



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

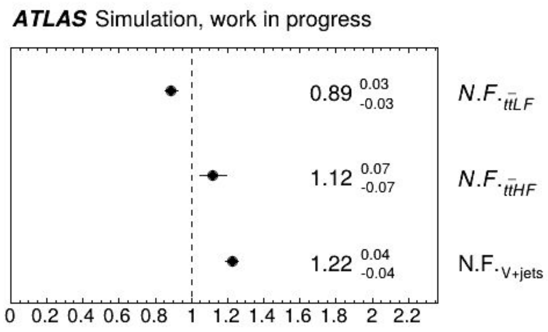


# Expected exclusion limits (Stat-only)




- Expected exclusion limit on signal cross sections obtained from a simultaneous S+B fit in SR and CR.
- Limits evaluated separately for each mass hypothesis.

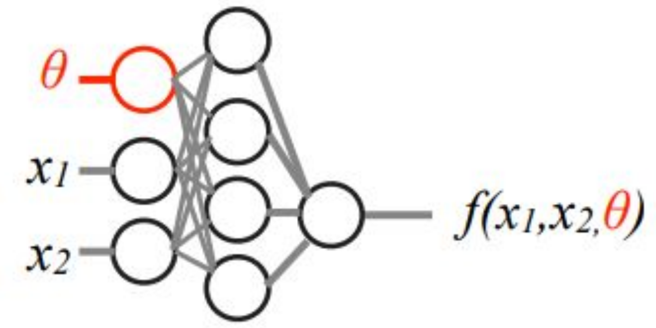
**From background-only fit**



# 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




Baldi, Pierre, et al. arXiv:1601.07913 (2016).

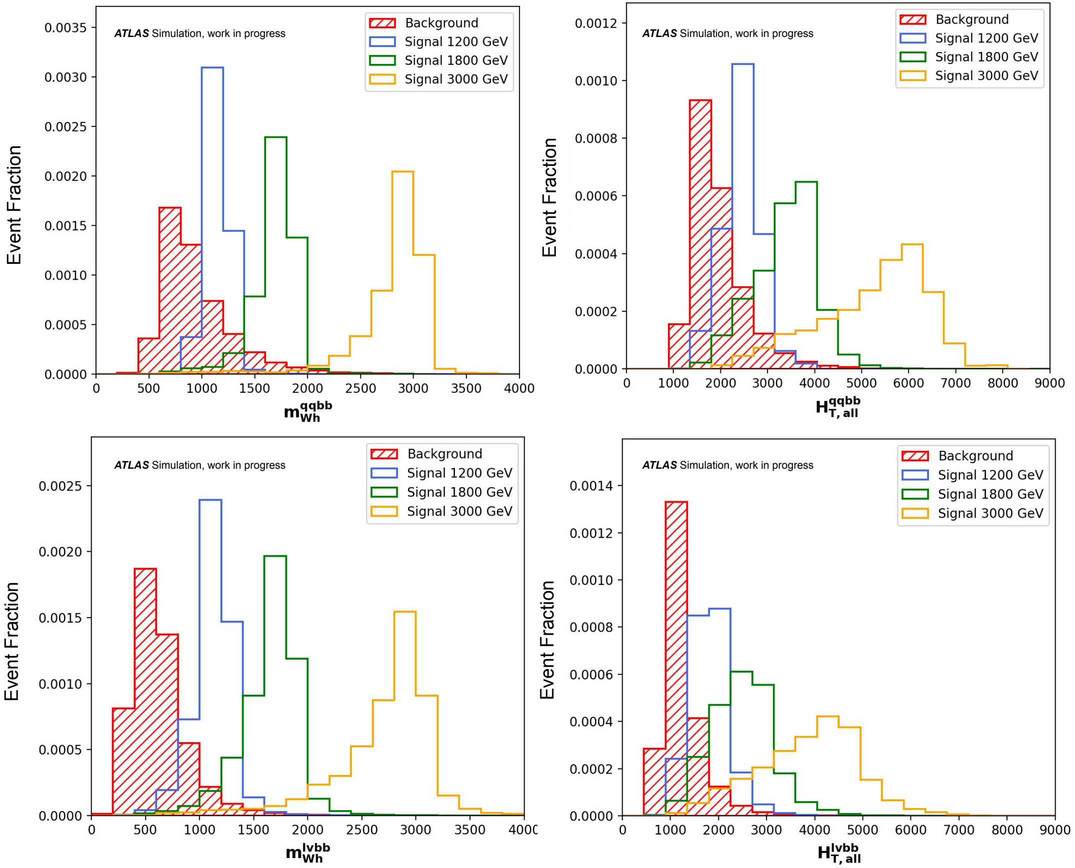
**Two Approaches chosen:**

- ❖ 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

Caveat: the PNN score is mass-dependent and is therefore evaluated separately at each mass hypothesis for both signal and background events.

# PNN-2 input features and optimization

- 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.

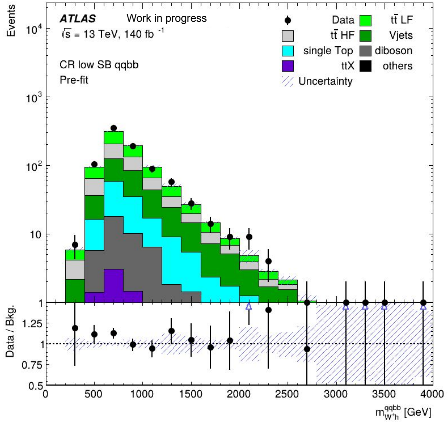


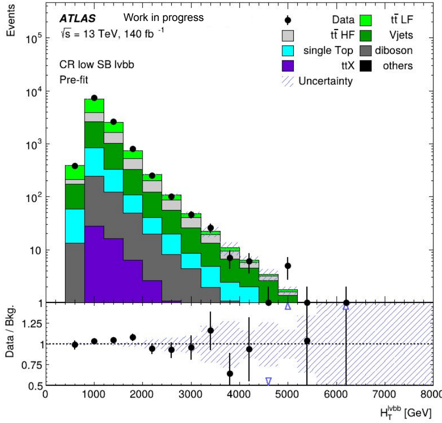
UNIVERSITÄT BONN

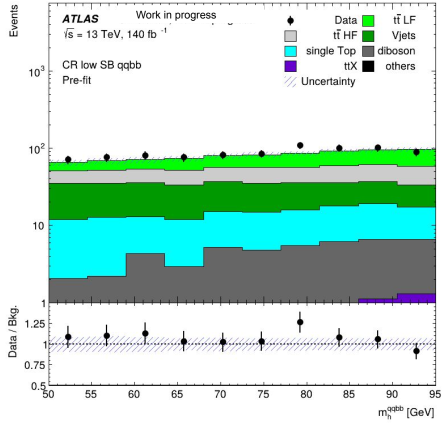
# Data-MC comparison in CR (PNN-2 inputs)

**Reconstructed $m_{Wh}$** | **$H_{T,all}$** | **Reconstructed $m_h$**

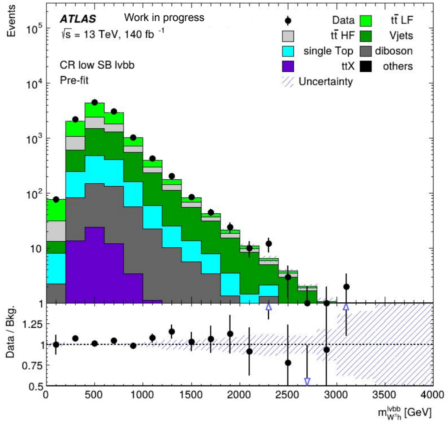
**qqbb**

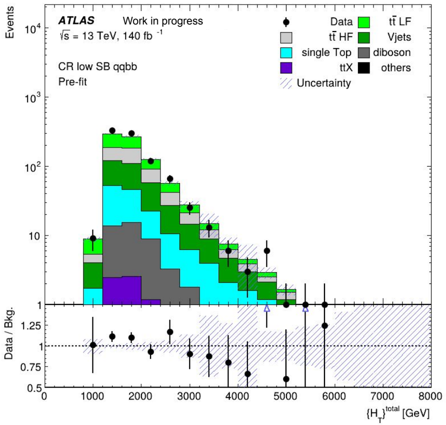


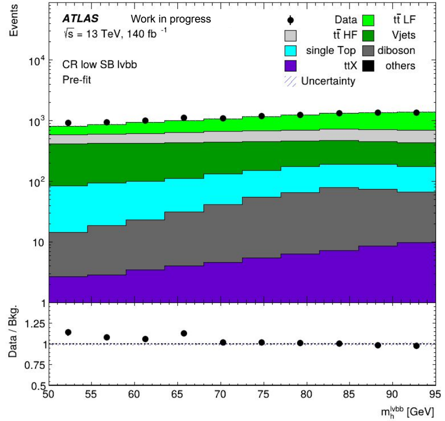




**ℓvbb**

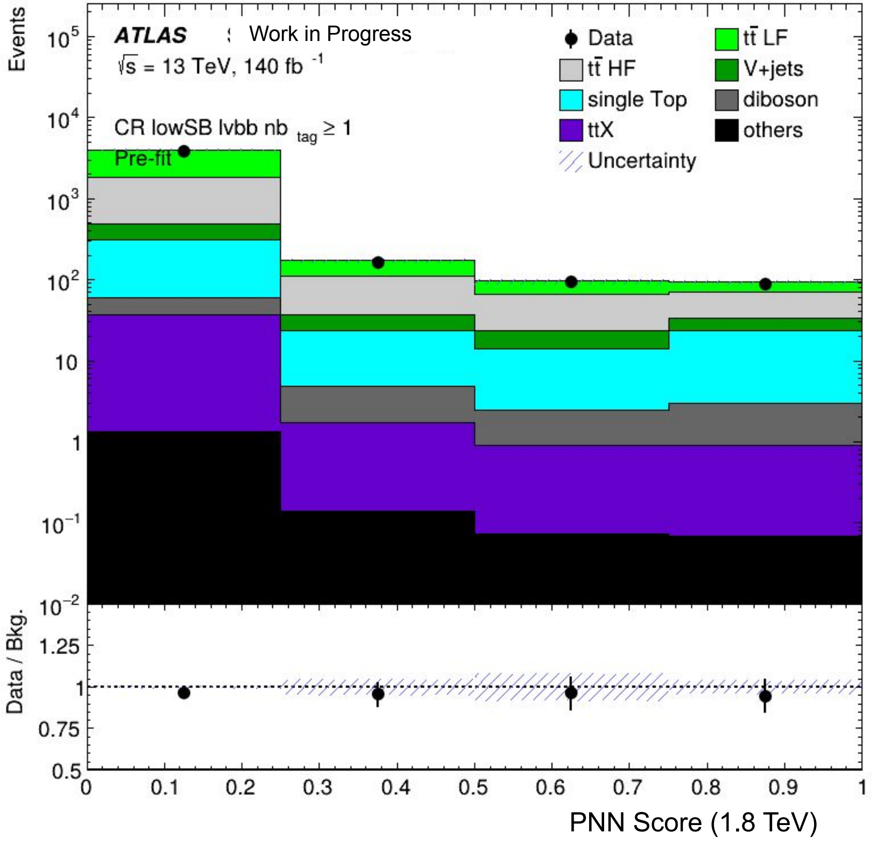


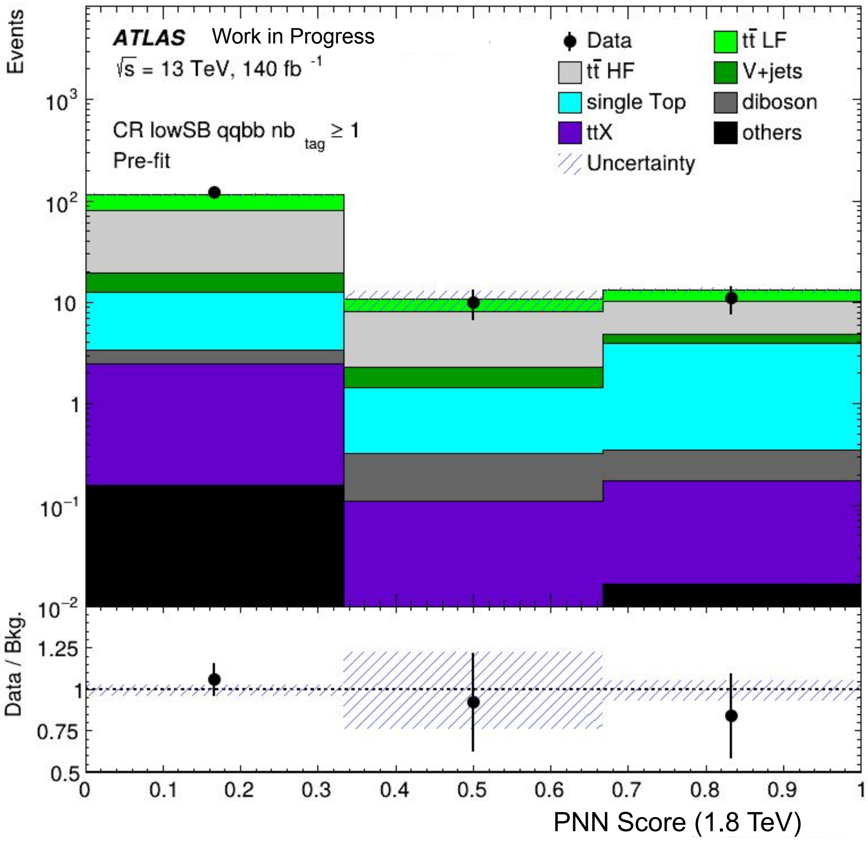



# PNN-2 score MC-data Agreement in CR

**ℓvbb channel**

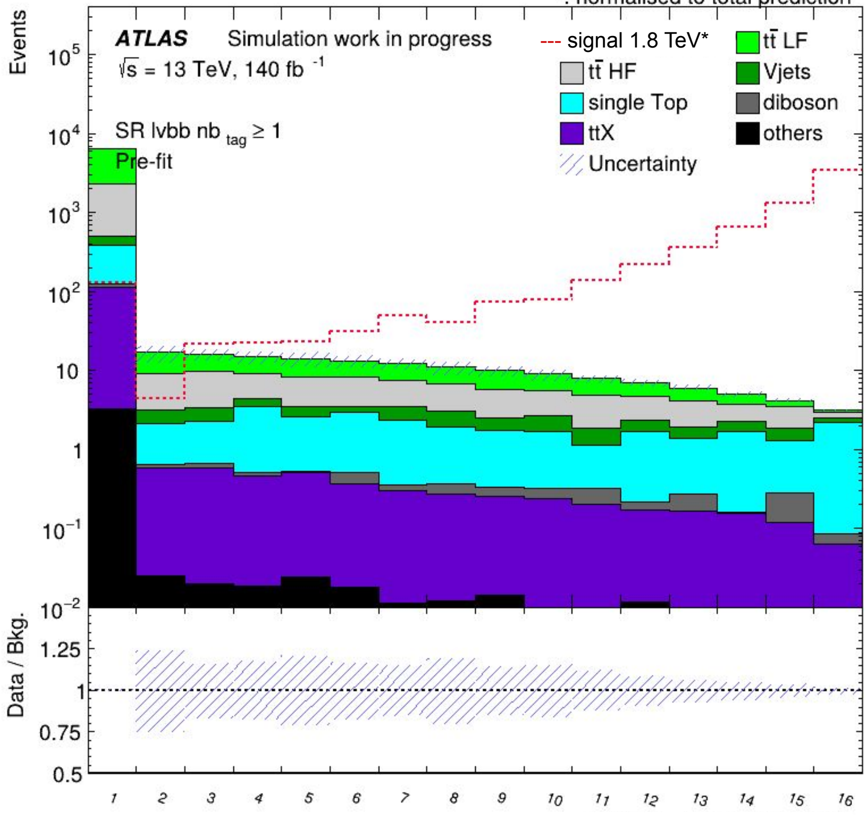



**qqbb channel**

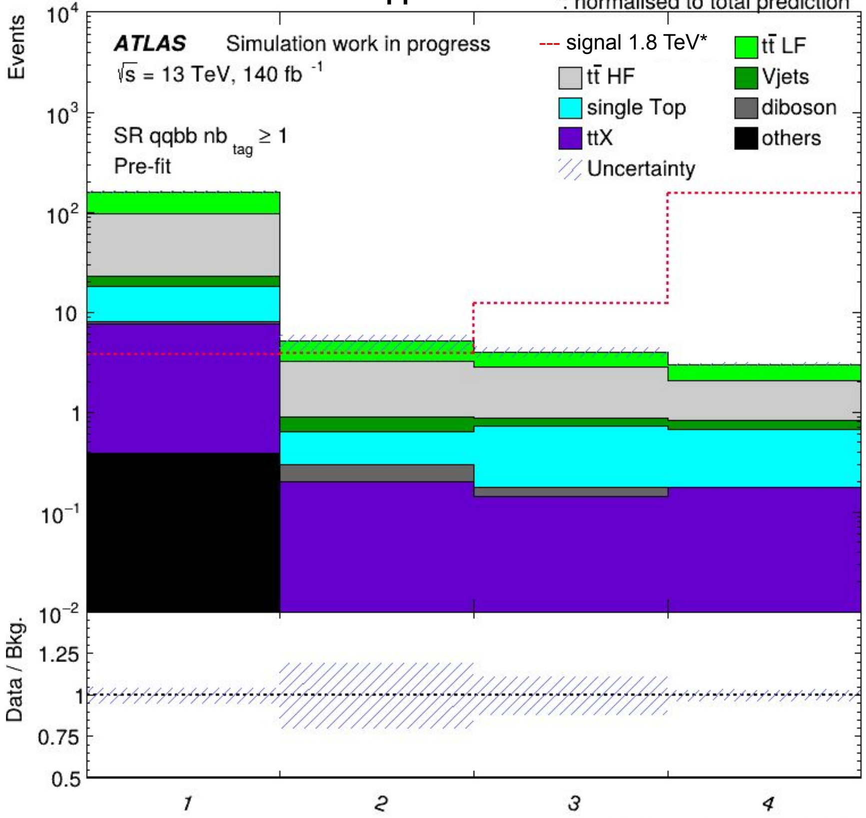


# PNN-2 score distribution in SR (Blinded)

**ℓvbb channel**

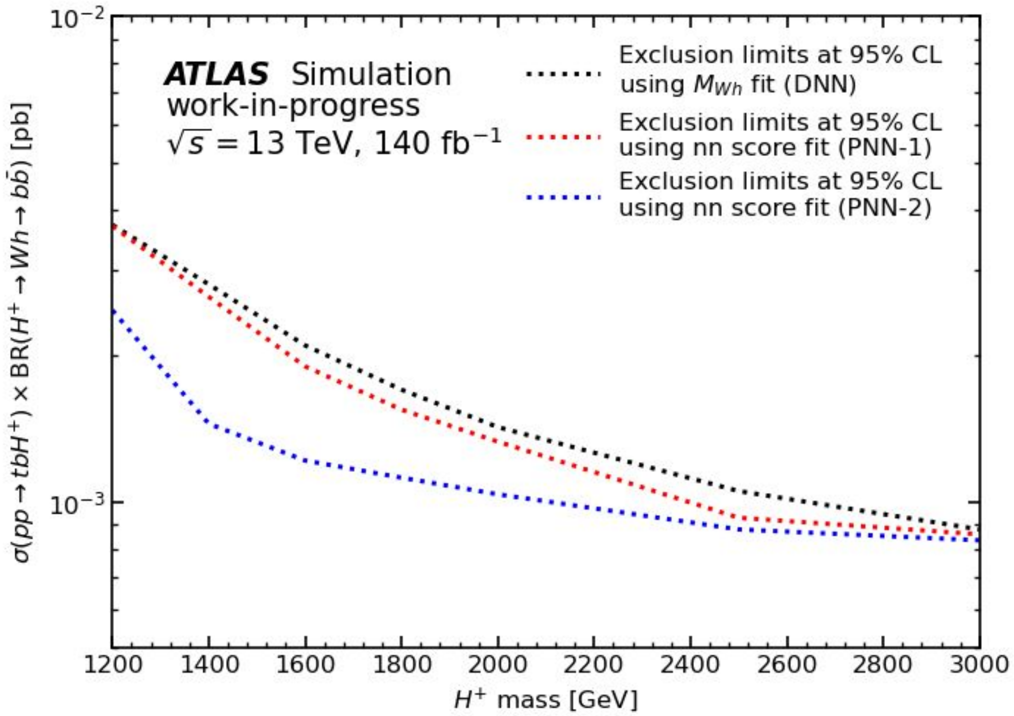



**qqbb channel**



# Comparison of exclusion limits

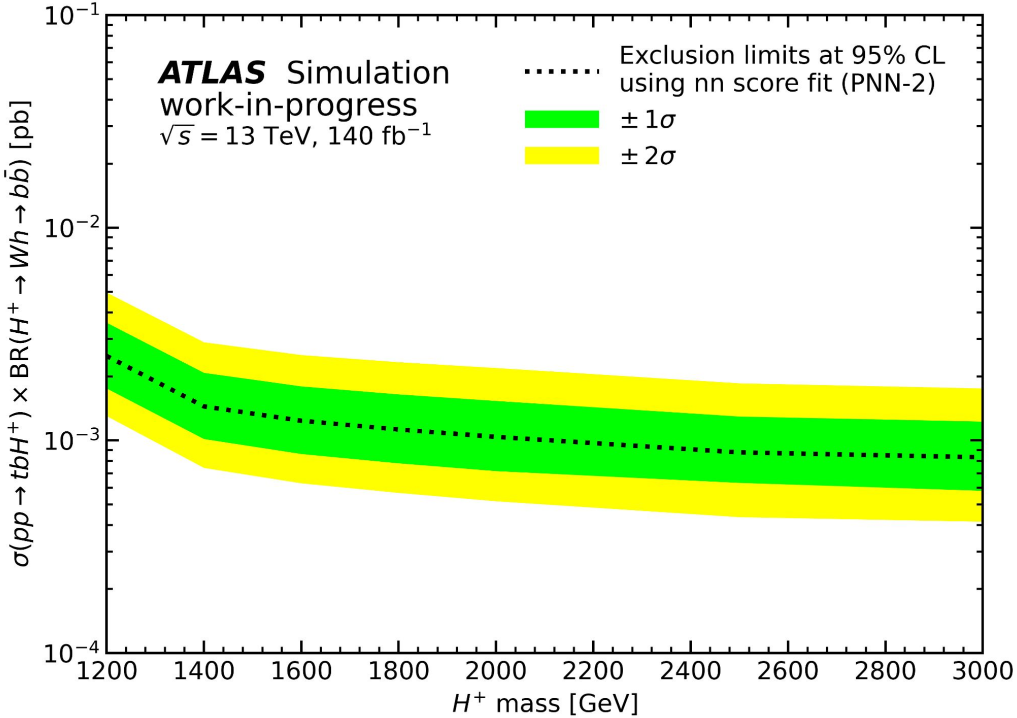
## Expected Limits (Statistics-only)




- **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 experimental 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

Previous ATLAS Analysis*

### Data + MC
Run-2 data and simulated signal/background

### Event selection
single e/$\mu$ trigger, exactly one lepton, $E_T^{miss} > 30$ GeV, large-R Higgs-tagged jet

### Categorisation
qqbb and lνbb channels

### Signal & control regions
SR: 95–140 GeV in $m_h$, CR: lower/higher sidebands; split by outside b-tags

## Baseline analysis

**DNN classification**

**Defined SR**
SR split by DNN score high / medium / low

**$m_{Wh}$ fit**
final fit variable

## Improved analysis

**PNN classification**
mass parameter + optimized inputs

**PNN score fit**
final fit variable

### Simultaneous binned profile likelihood fit in SR + CR

### 95% CL expected limits
compared DNN and PNN sensitivity

Final PNN result includes a limited set of experimental systematics

# Outlook

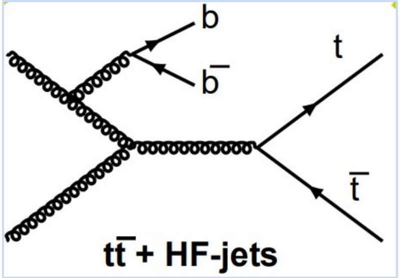
- Extend PNN-2 with the full set of relevant systematic uncertainties.
- Unblind the PNN-2 analysis and compare its performance directly with the previous ATLAS strategy.
- Explore more unified multivariate approaches across both channels.

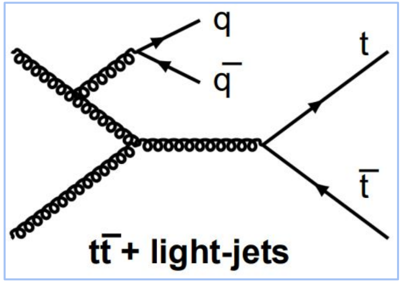
# Thank you!

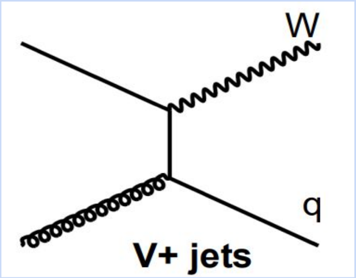
Particle physicists when
their latest measurement
once again confirms The
Standard Model

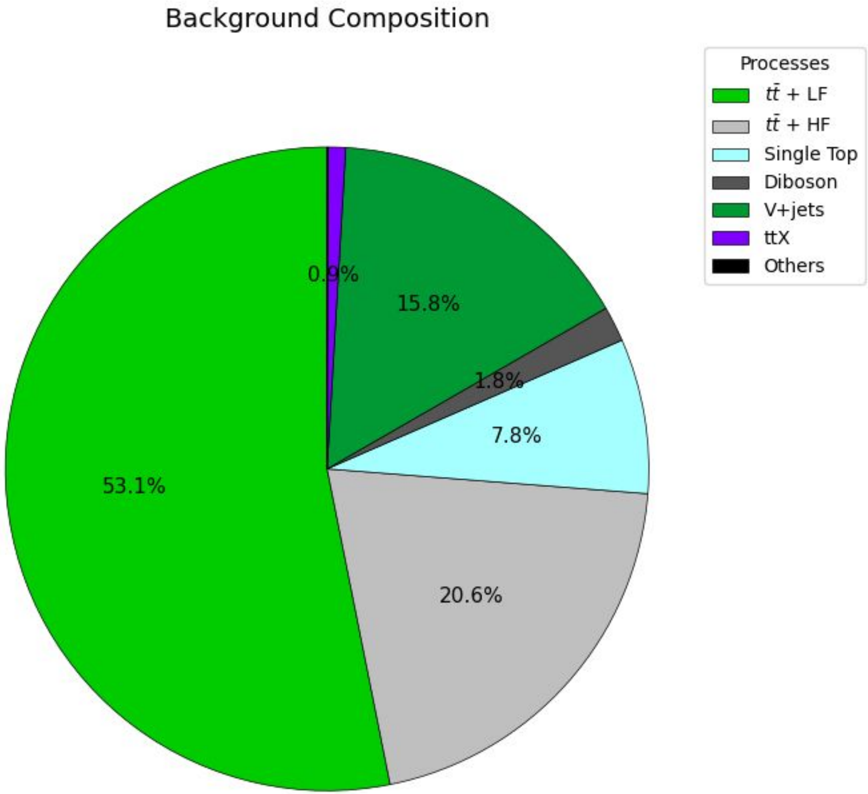
# Back-up

# Backgrounds



# Yields after preselection and categorisation

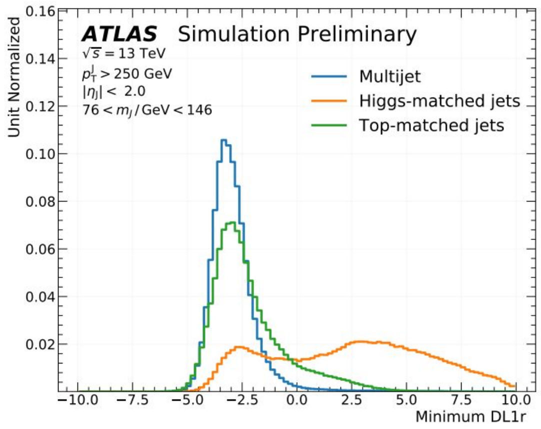
| Process | qqbb weighted | $\ell\nu bb$ weighted | qqbb events | $\ell\nu bb$ events | total events |
|---|---|---|---|---|---|
| Signal 1.2 TeV | 4938 ± 104 | 15759 ± 183 | 29880 | 94806 | 124686 |
| Signal 1.4 TeV | 4995 ± 103 | 14877 ± 175 | 31990 | 95576 | 127566 |
| Signal 1.6 TeV | 4059 ± 89 |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

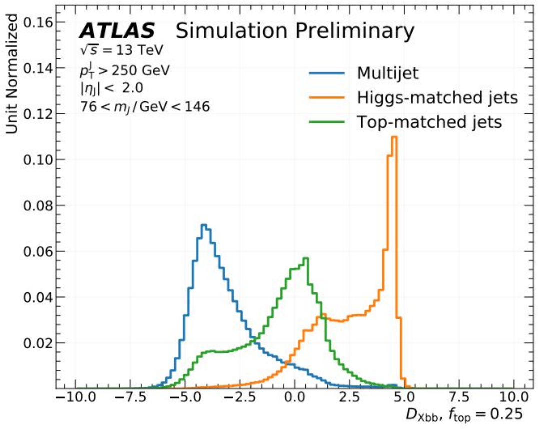
# X-bb Tagger

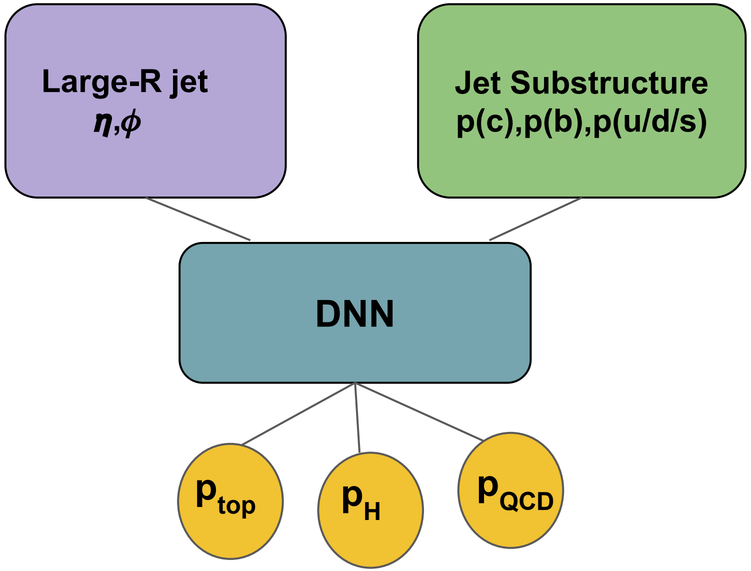
For passing 60% working point:

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

$f_{top} = 0.25$




ATL-PHYS-PUB-2020-019

# DNN input variables and setup

## qqbb variables

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

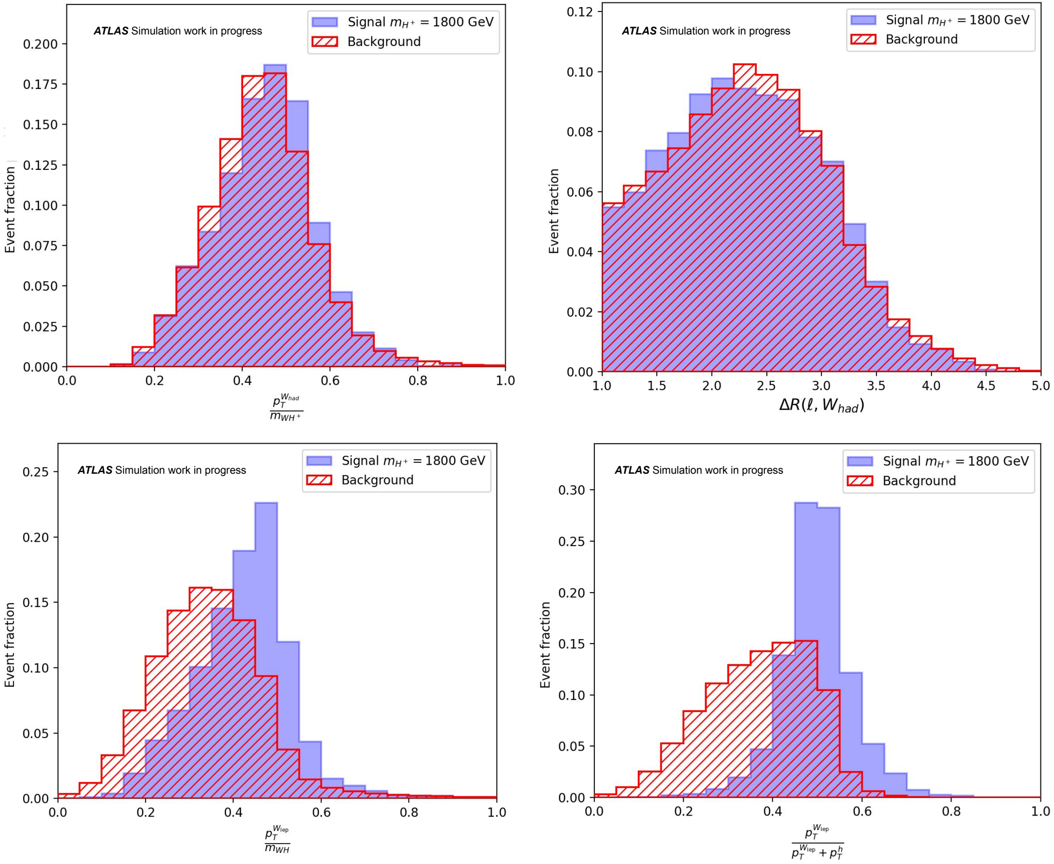
$$\frac{p_T^{W_{had}}}{m_{WH^+}}, \quad \frac{p_T^{h}}{m_{WH^+}}$$

## $\ell\nu$bb variables

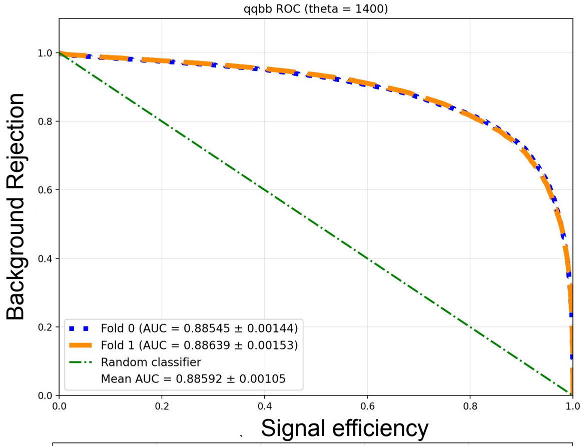
$$\frac{p_T^{W_{lep}}}{p_T^{W_{lep}} + p_T^{h}}, \quad m_{\text{top}}^{\text{lep}}, \quad \Delta R(\ell, h), \quad \Delta\eta(h, W_{\text{lep}}), \quad \Delta\phi(h, W_{\text{lep}})$$

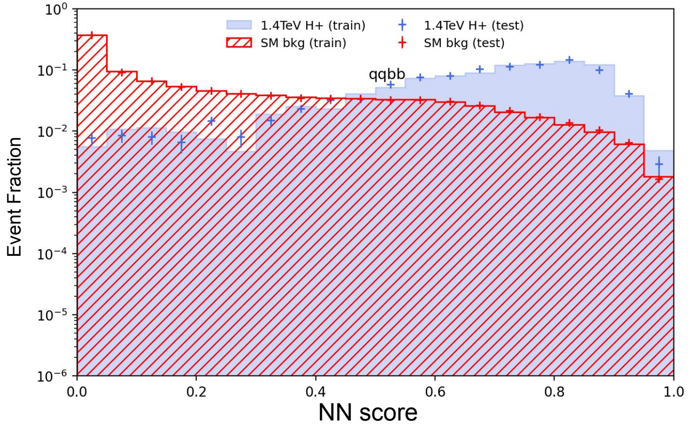
$$\frac{p_T^{W_{lep}}}{m_{WH}}, \quad \frac{p_T^{h}}{m_{WH}}$$

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



# DNN performance

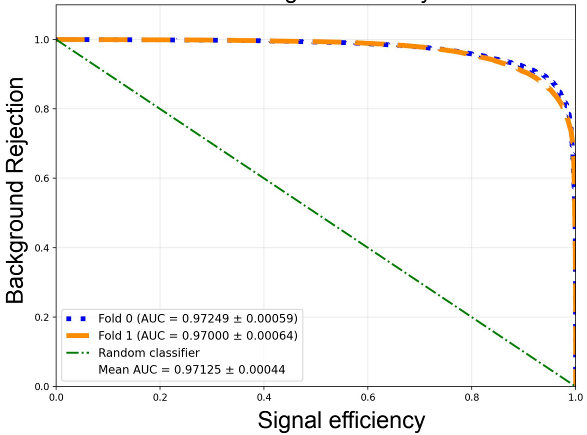


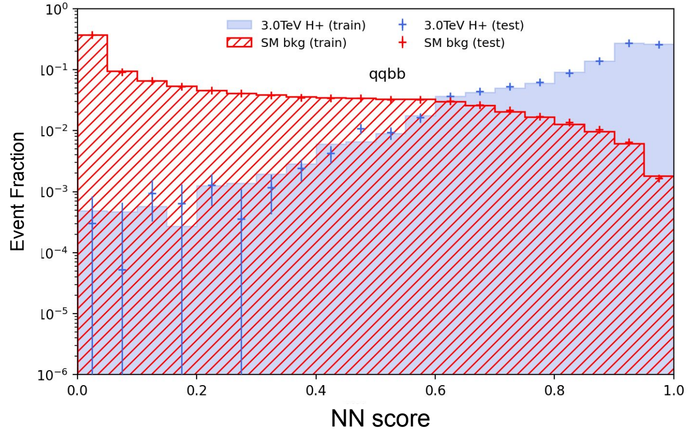




qqbb 1.4 TeV

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

qqbb 3 TeV



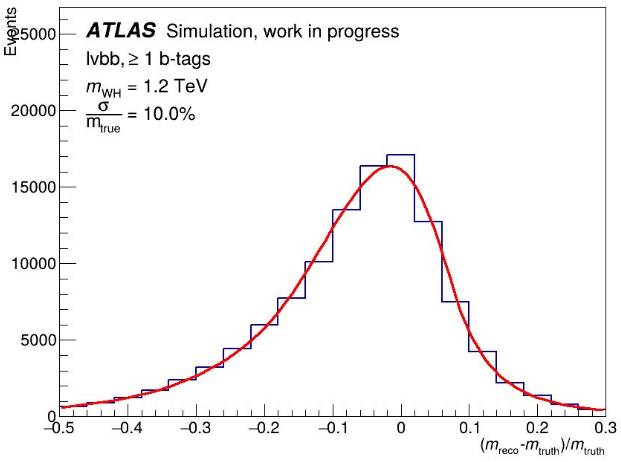



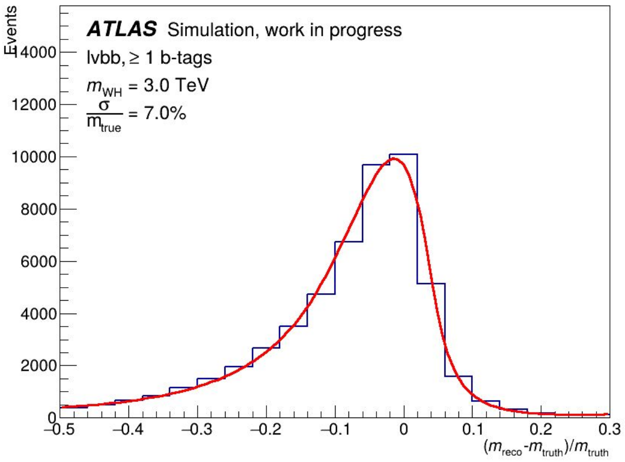

Performance metric:
Area under the curve (AUC)

*AUC errors are bootstrapped errors

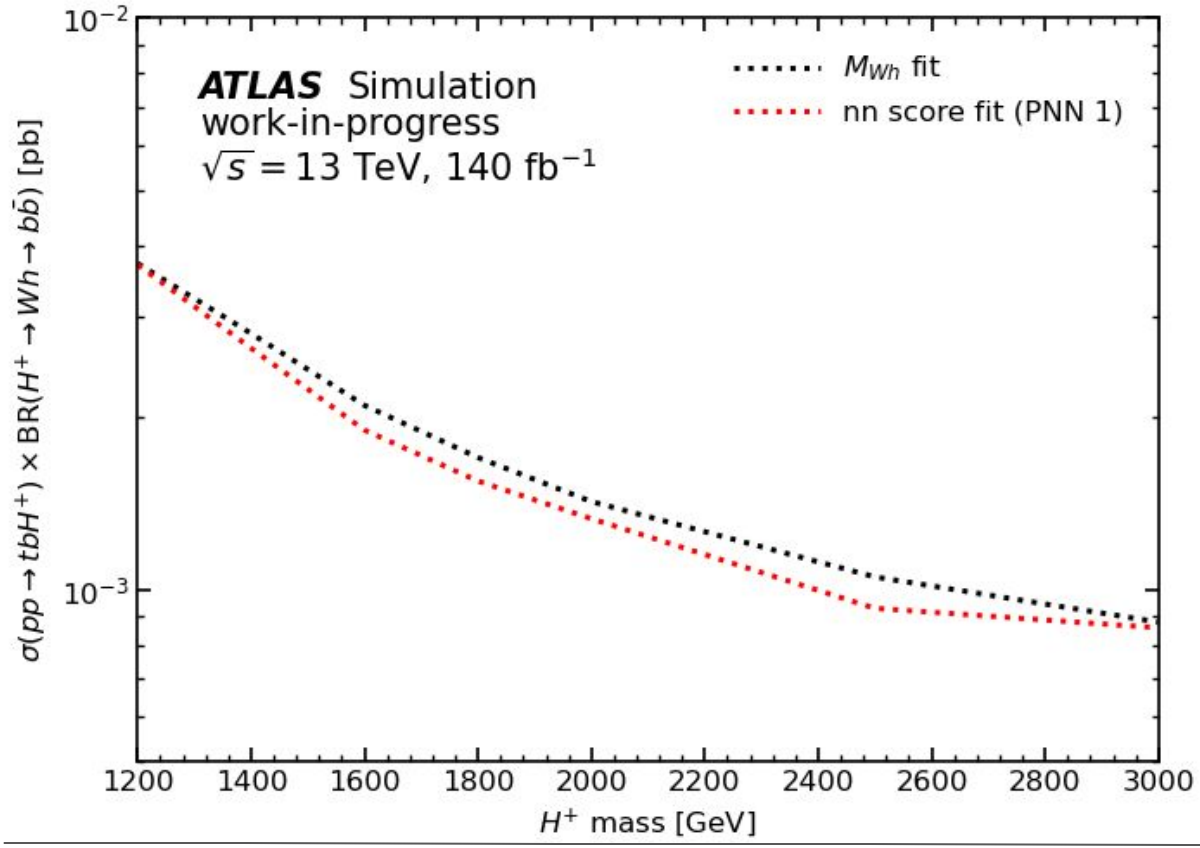
# $M_{Wh}$ Binning

- 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), **11%** (2.5 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: $\Delta AUC = AUC_{Baseline} - AUC_{permuted}$
- For many low-ranked features, the ranking kept changing across mass points.
- **Normalized ΔAUC** for a feature across all features per mass point.

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

- Compared the **normalized ΔAUC** of the feature across all mass points.

**Baseline AUC [1.2 TeV, 3 TeV]:**

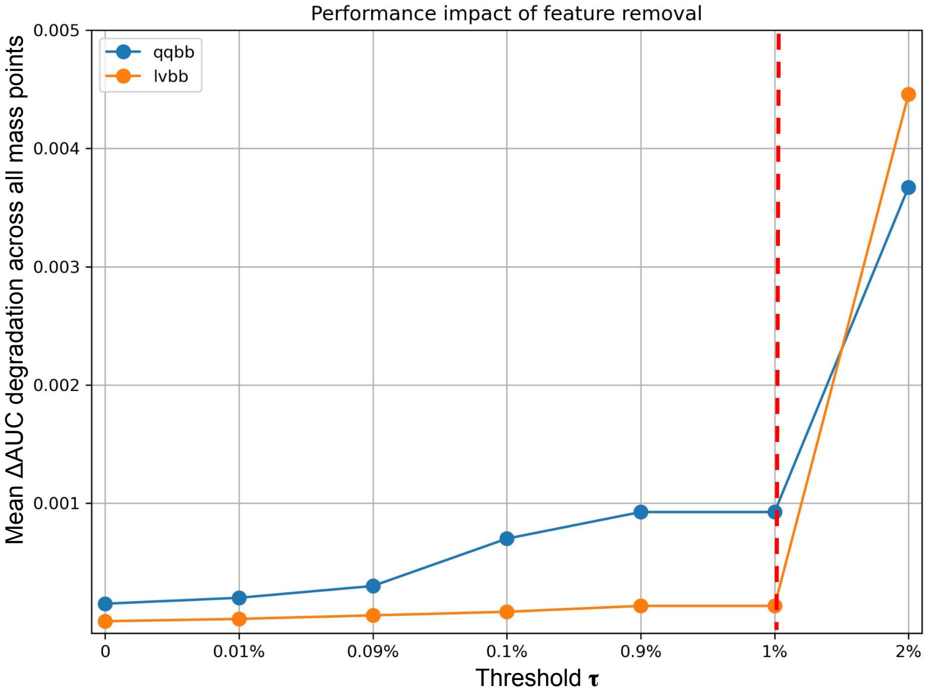
**qqbb: [0.926, 0.956]**
**ℓνbb: [0.943, 0.971]**

# Threshold scan

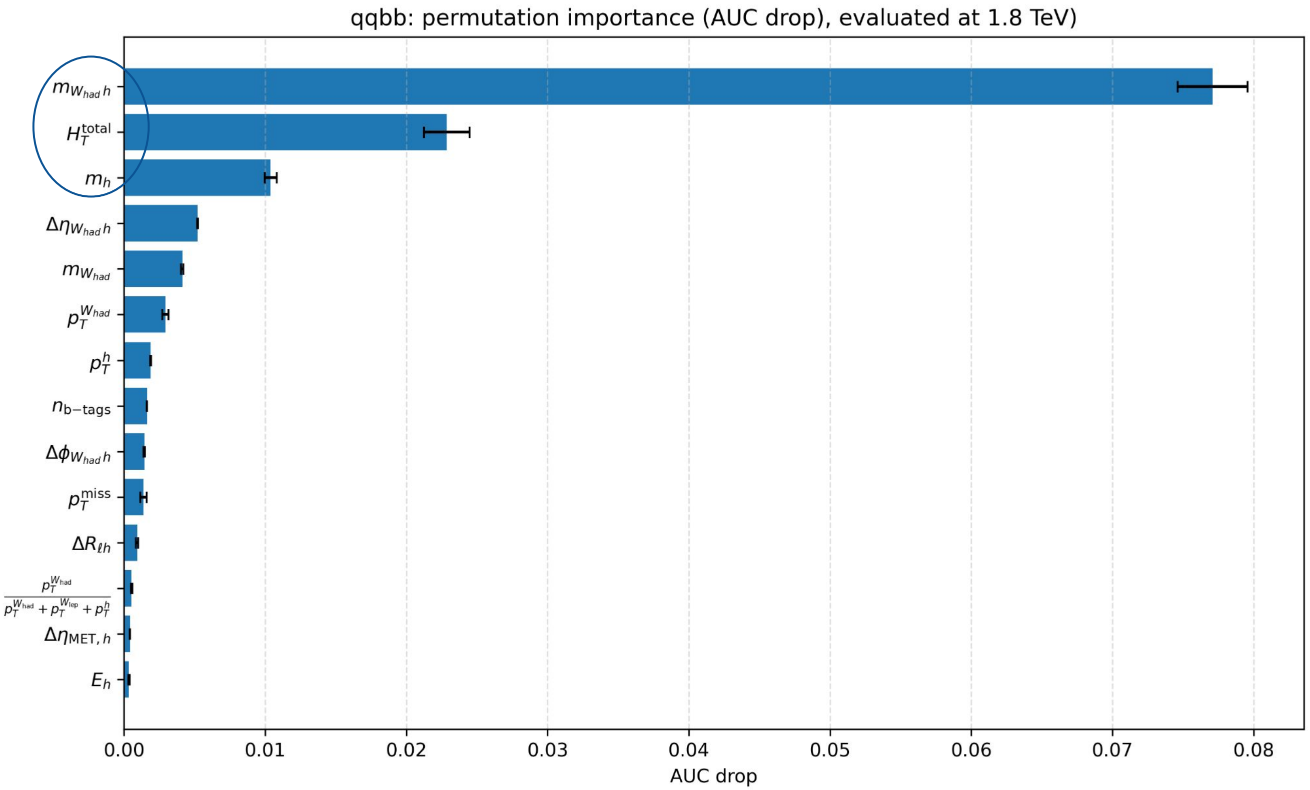
- Scanned grid of thresholds $\tau$ (normalized ΔAUC) = {0.01%, 0.09%, 0.1%, 0.9%, 1%, 2%}
- Chosen threshold = 1% (Helped getting rid of many low ranked variables)

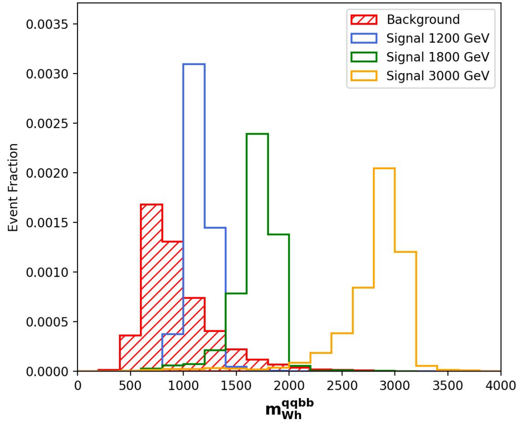
| Channel | Number of features dropped |
|---|---|
| qqbb | 15 |
| $\ell$vbb | 11 |

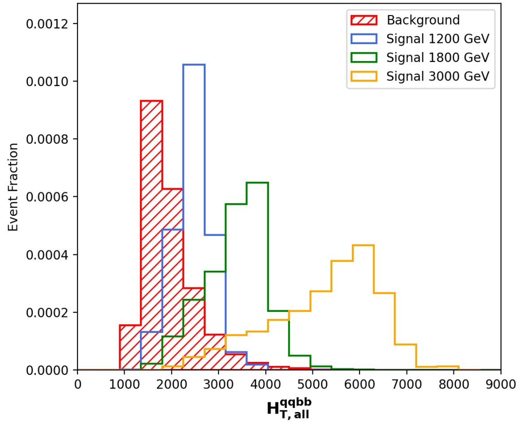
| Channel | Final number of features |
|---|---|
| qqbb | 14 |
| $\ell$vbb | 19 |



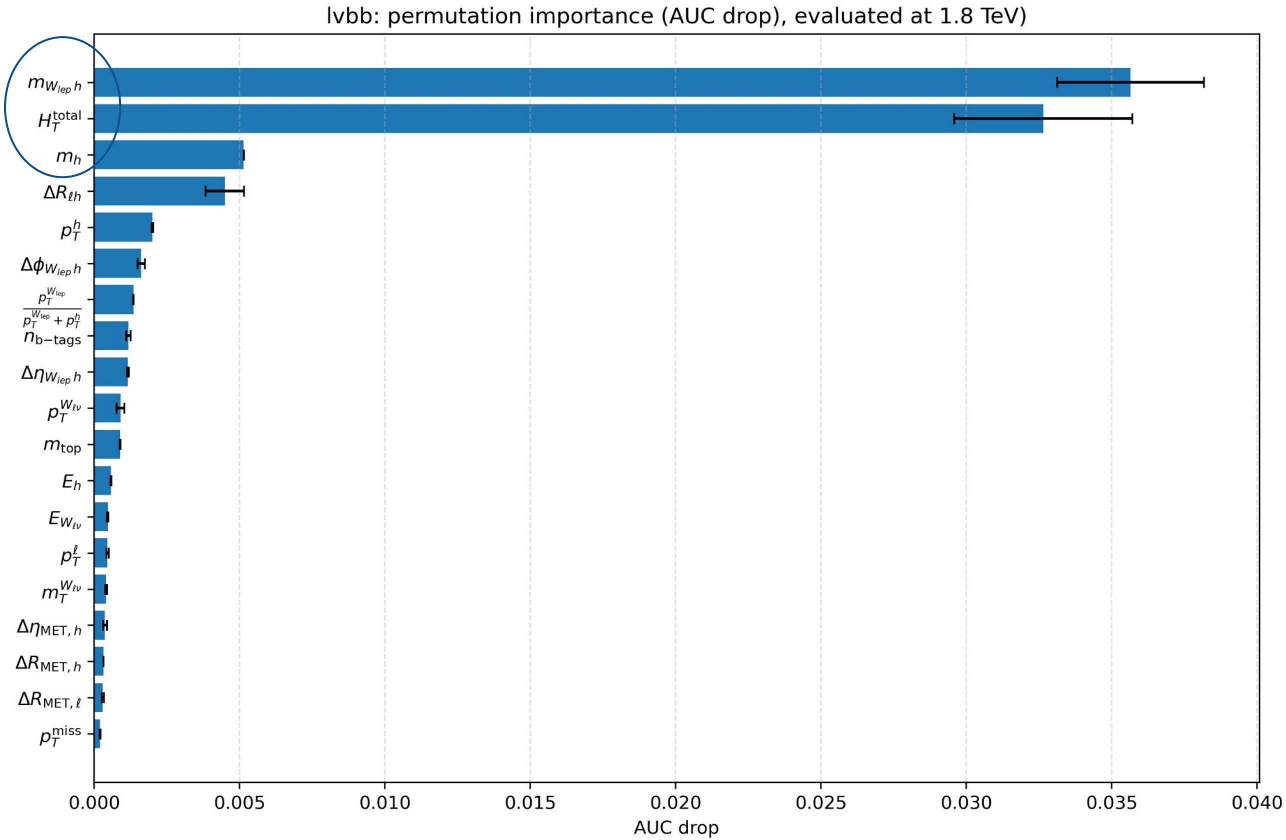
# PNN-2 Feature importance (qqbb)

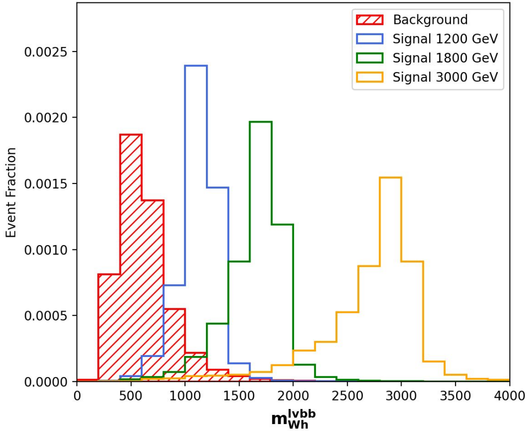


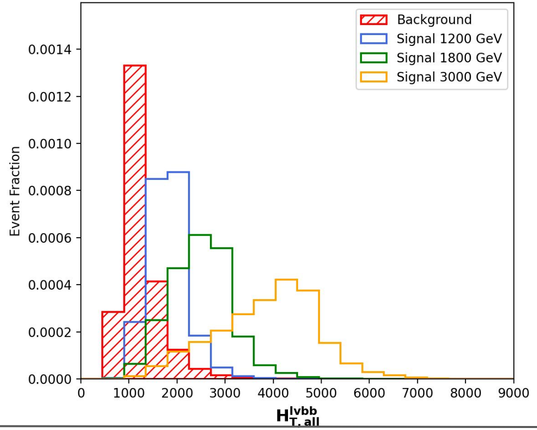



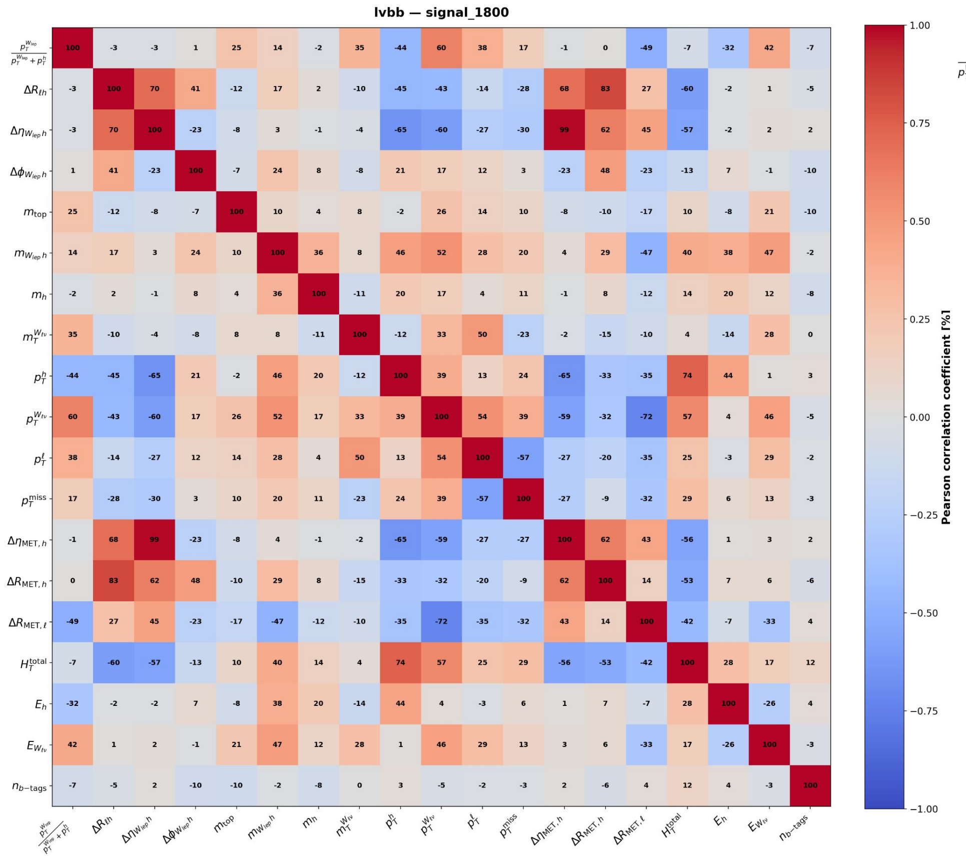
# PNN-2 Feature importance (ℓvbb)

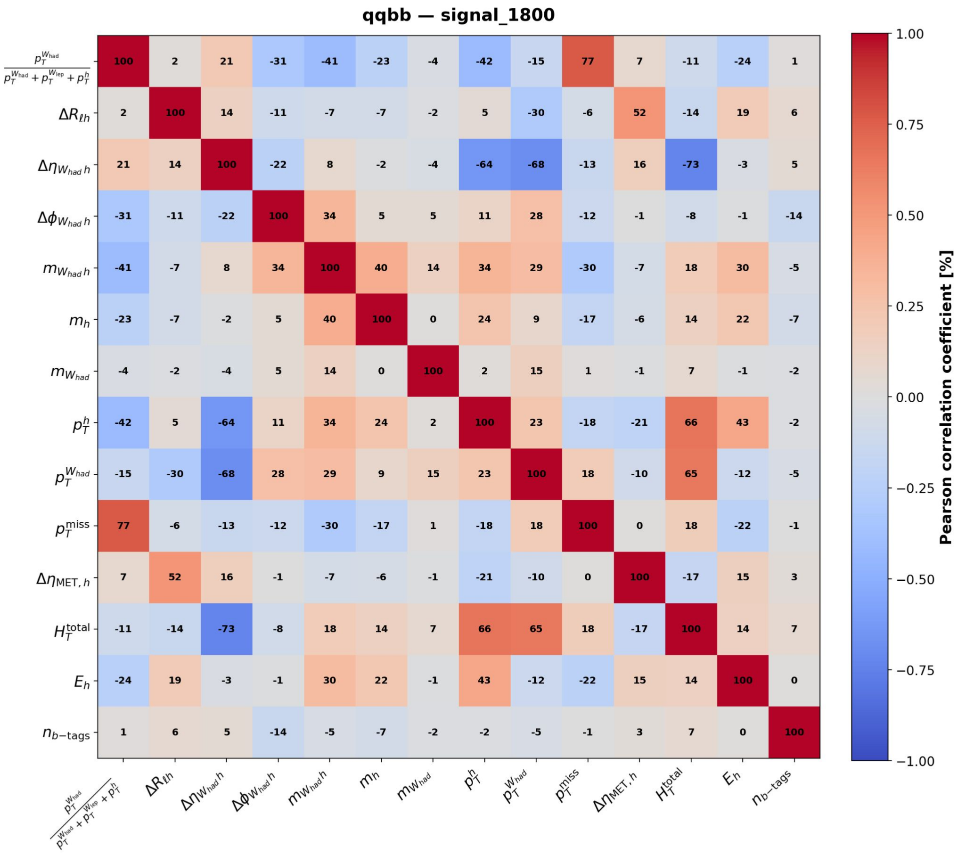


# PNN-2 Correlation Matrices (1.8 TeV)



# PNN-2 variables and optimized setup

**qqbb variables**

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

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

**ℓvbb variables**

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

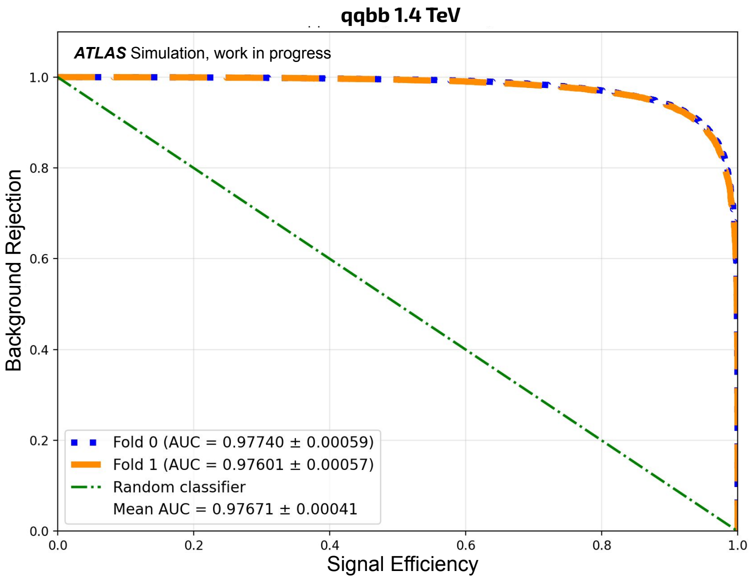
$$p_T^h,\quad p_T^{W_{\ell\nu}},\quad p_T^{\ell},\quad p_T^{\text{miss}},\quad \Delta\eta_{\text{MET},h},\quad \Delta R_{\text{MET},h},\quad \Delta R_{\text{MET},\ell},$$

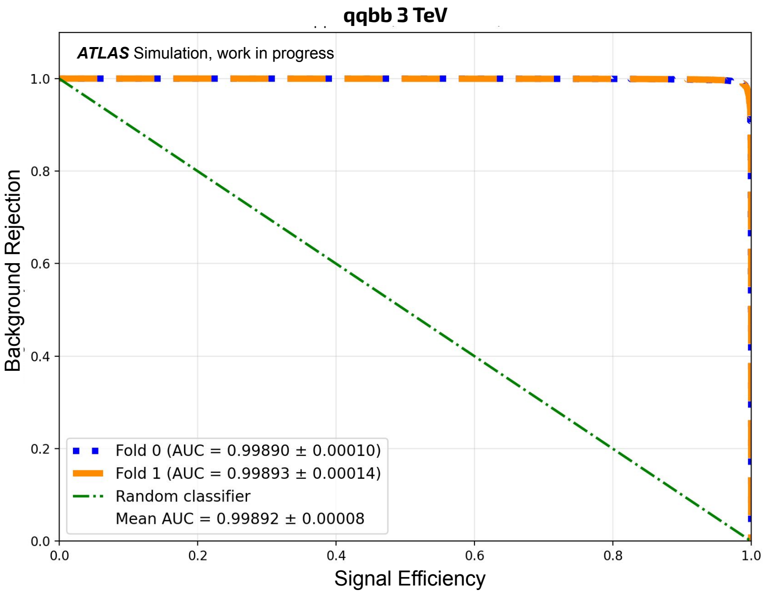
$$H_T^{\text{total}},\quad E_h,\quad E_{W_{\ell\nu}},\quad n_{\text{b-tags}},\quad \frac{p_T^{W_{\text{lep}}}}{p_T^{W_{\text{lep}}}+p_T^h},$$

- **Architecture (Best from Optuna scan)**
  - **qqbb:**
    - **4 hidden layers: 512 → 256 → 256 →256**
    - **Dropout: 0.3**
    - **L2: $0.7\times10^{-6}$**
    - **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: $1.05\times10^{-6}$**
    - **Learning rate: 0.0001**
    - **Batch size: 64**
    - **AUC: 0.986**

# PNN-2 Performance



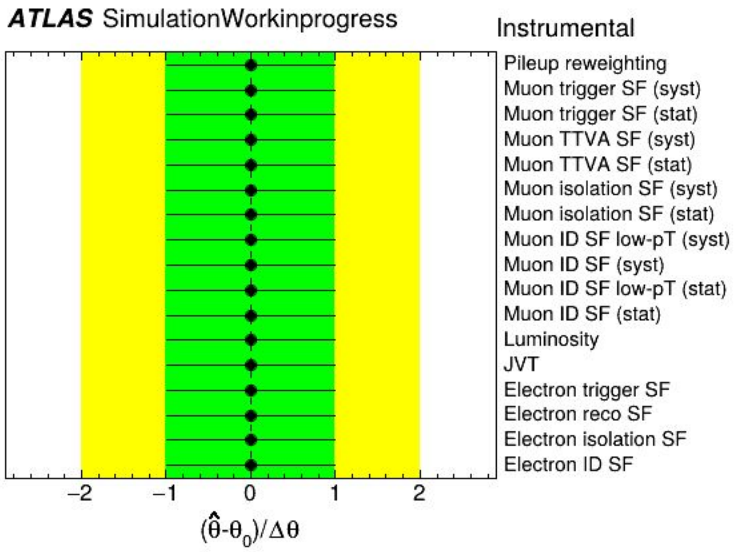

**Interpolation:**

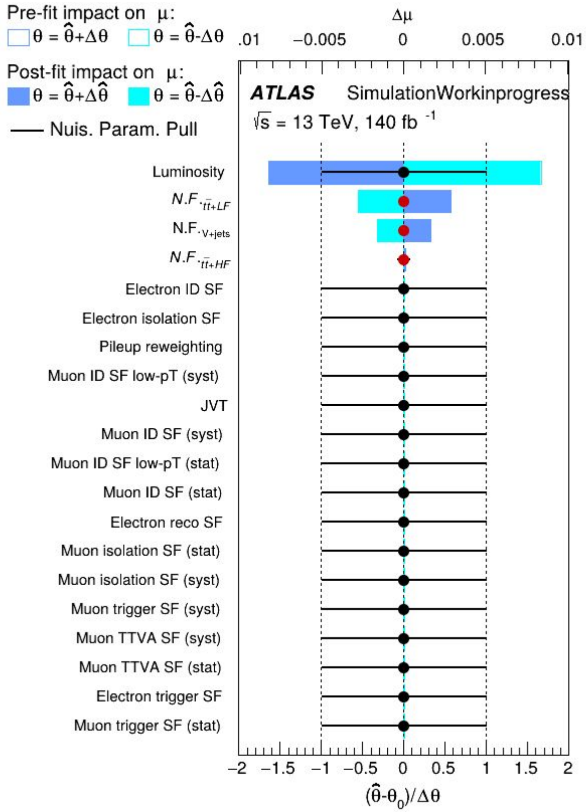
| Signal mass dropped | 1.8 TeV | 2.5 TeV |
|---|---|---|
| Drop in AUC | 0.001% | 0.001% |

NN Score

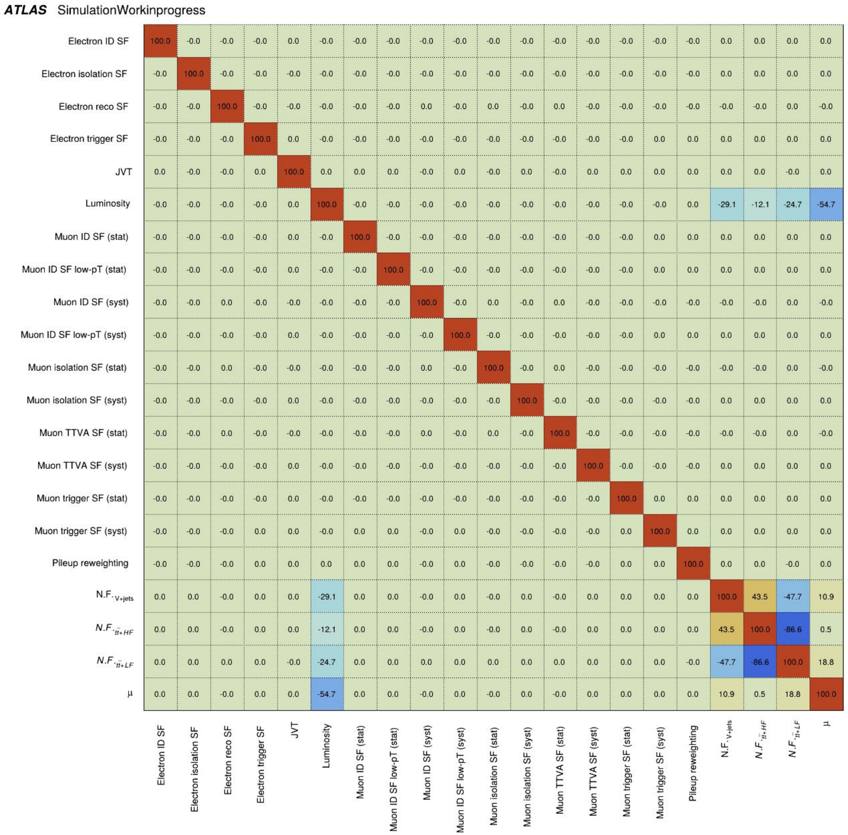
# PNN-2 systematics

After S+B blinded fit



# PNN-2 systematics



# ATLAS analysis Systematics

| Systematic uncertainty | Type | Components |
|---|---|---|
| **Experimental uncertainties** | | |
| Luminosity | N | 1 |
| Pileup modeling | SN | 1 |
| *Physics objects* | | |
| Electrons | SN | 7 |
| Muons | SN | 15 |
| Jet energy scale | SN | 31 |
| Jet energy resolution | SN | 9 |
| Jet vertex tagger | SN | 1 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ | SN | 3 |
| *b-tagging* | | |
| Efficiency | SN | 45 |
| Mis-tag rate (*c*) | SN | 20 |
| Mis-tag rate (light) | SN | 20 |
| high $p_{\mathrm{T}}$ extrapolation | SN | 3 |

# ATLAS analysis Systematics

| **Signal and background modeling** | | |
|---|---|---|
| *Signal* | | |
| PDF variations | N | 1 |
| Scales | SN | 2 |
| PS | N | 1 |
| *$t\bar{t}$ background* | | |
| PDF variations | SN | 30 |
| $t\bar{t}$ reweighting | SN | 24 |
| $t\bar{t}$ normalisation | N (free floating) | 1 ($t\bar{t}$ + HF) |
| $t\bar{t}$ modelling | SN | 21 ($t\bar{t}$+ light) and 57 ($t\bar{t} + HF$ ) |
| Transfer factor uncertainties | N | 6 |
| *Other backgrounds* | | |
| $t\bar{t}W$ cross-section | N | 2 |
| $t\bar{t}Z$ cross-section | N | 2 |
| $t\bar{t}W$ modelling | SN | 1 |
| $t\bar{t}Z$ modelling | SN | 1 |
| $t\bar{t}H$ cross-section | N | 1 |
| $t\bar{t}H$ modelling | SN | 2 |
| Single top cross-section | N | 6 |
| Single top modeling | SN | 12 |
| $W$+jets normalisation | N | 1 |
| $Z$+jets normalisation | N | 1 |
| Diboson normalisation | N | 1 |
| Small backgrounds cross-sections | N | 3 |