A Search for Triple Higgs Production at the ATLAS Detector

IoP APP, HEPP & NP 2024, Liverpool

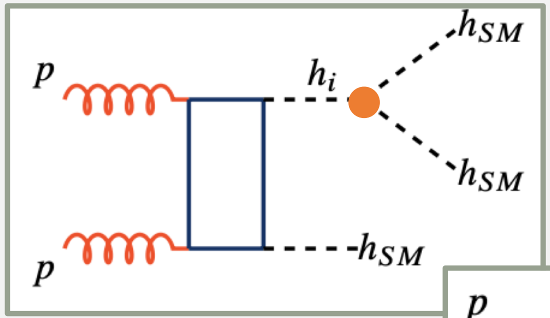
Maggie Chen, Todd Huffman
Holly Pacey, Bill Balunas

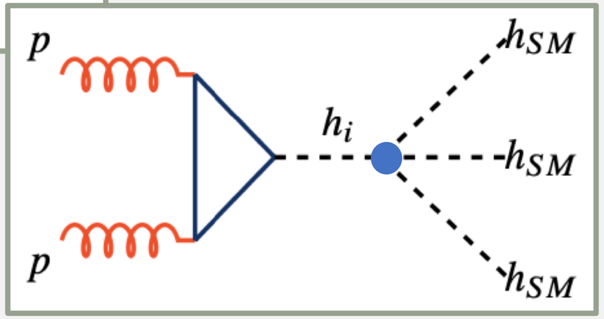
Triple Higgs Production in the Standard Model

- The Higgs potential at low-energy:

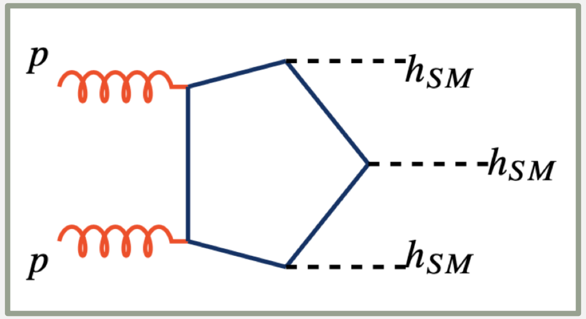
$$V(H) = \frac{1}{2}m_H^2H^2 + \lambda_3 vH^3 + \lambda_4 vH^4 + O(H^5)$$

  - SM tri-Higgs production is sensitive to tri-linear & quartic Higgs self-couplings, λ_3 and λ_4
  - Can constrain modifications to SM coupling values (κ_3 and κ_4)



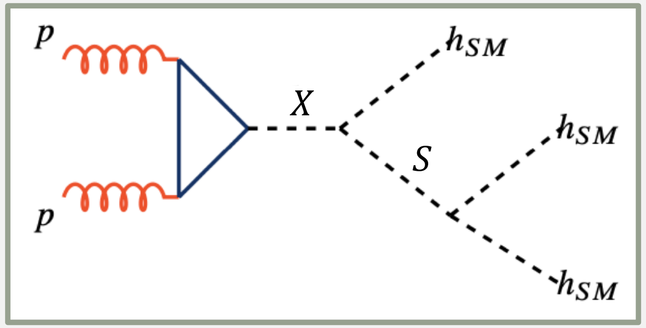


- However
  - SM tri-Higgs production suffers from small cross-section (~ 0.033 fb at LO)
  - Dominant SM production diagram – pentagon diagram

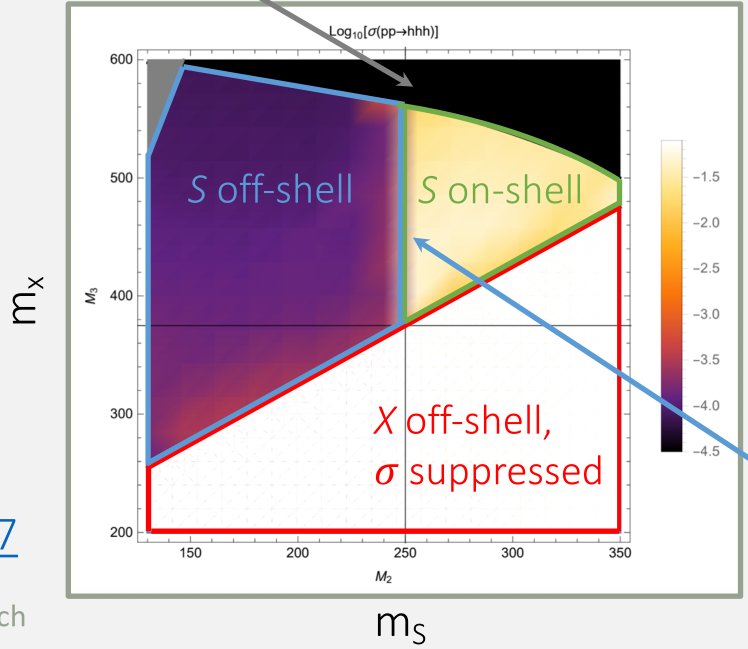


Triple Higgs Production Beyond the Standard Model

- Two Real Singlet Model (TRSM)
  - Extends SM by two real scalar fields that are singlets under all SM symmetries
  - New mass eigenstates: X and S
  - If X and S are sufficiently heavy – resonant production via cascade decay



Perturbative unitarity violated



S on-/off-shell border

arXiv:2101.00037

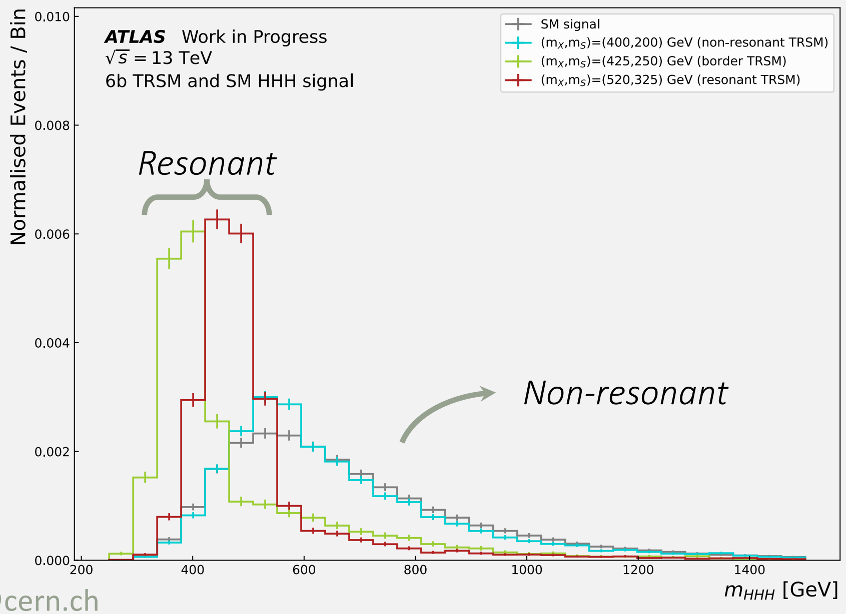
- Enhanced cross-section (~50 fb at LO)
  - Accessible at the LHC
- Maximises the decay to 3 SM Higgs bosons
- A broad region in [m_X, m_S] phase space to explore

Resonant & Non-resonant Interpretations

The event kinematics differ depending on the type of production – resonant vs. non-resonant

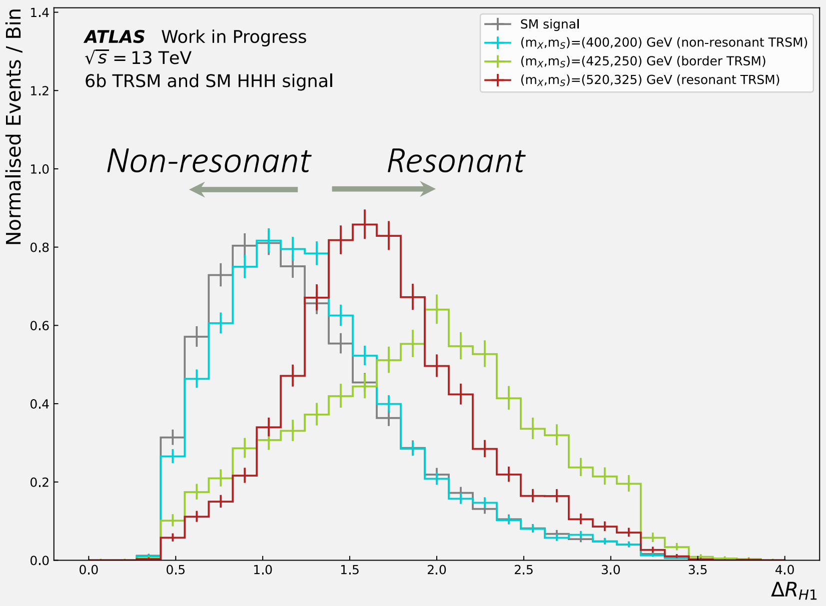
Resonant HHH production:

- When *X* and *S* in the TRSM model are on-shell
  - $m_S > 2\ m_H$ (250 GeV) – resonant TRSM
  - $m_S = 2\ m_H$ (250 GeV) – border TRSM



Non-resonant HHH production:

- When *S* in the TRSM model is off-shell, i.e.
  - $m_S < 2\ m_H$ (250 GeV) – non-resonant TRSM
- SM HHH production

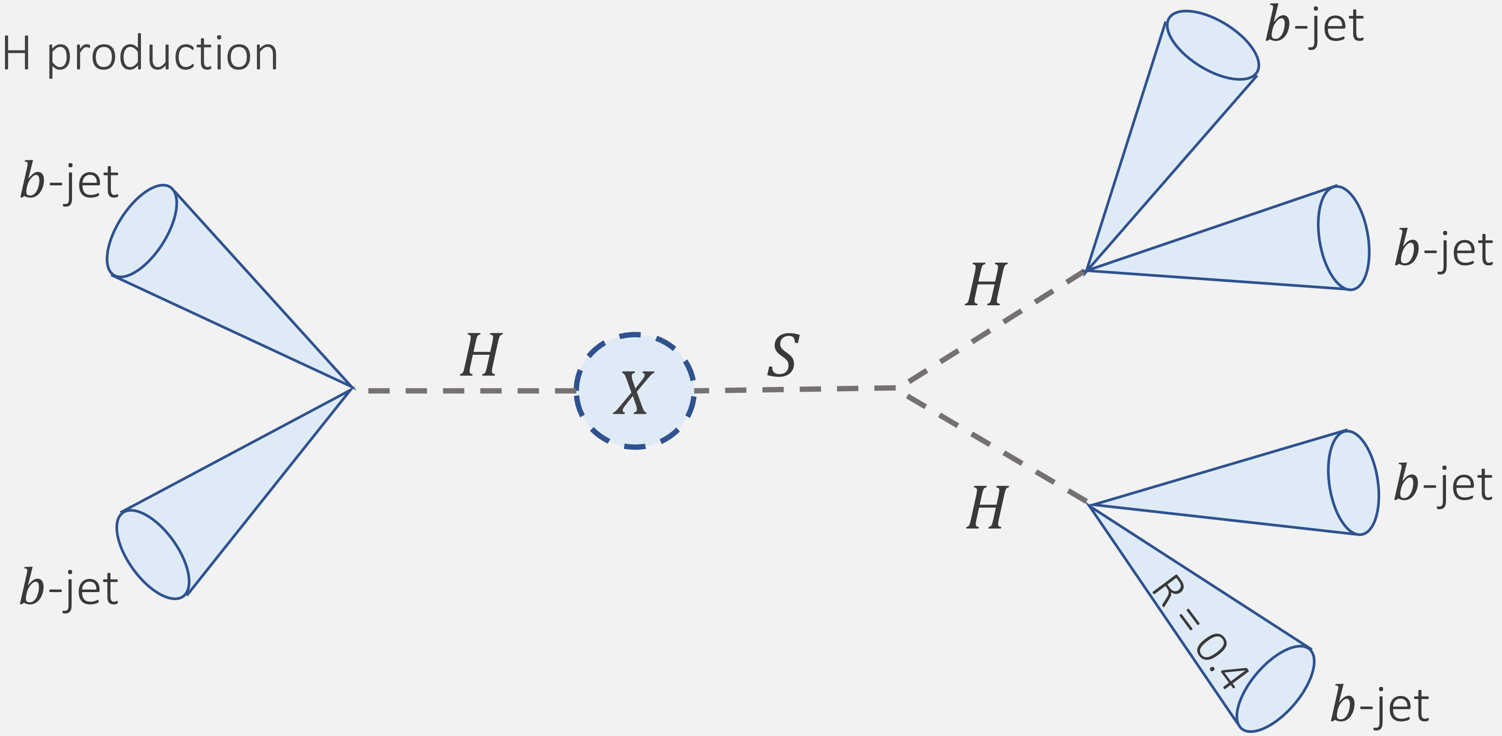


Event Final State & Topology

The 6*b* final state

- Exploiting the large branching ratio of $H \to b\bar{b}$ (~58%)
- *b*-quarks hadronise to form *b*-hadrons, and are detected as *b*-jets
- Identified by an ML-based *b*-tagging algorithm (DL1d)

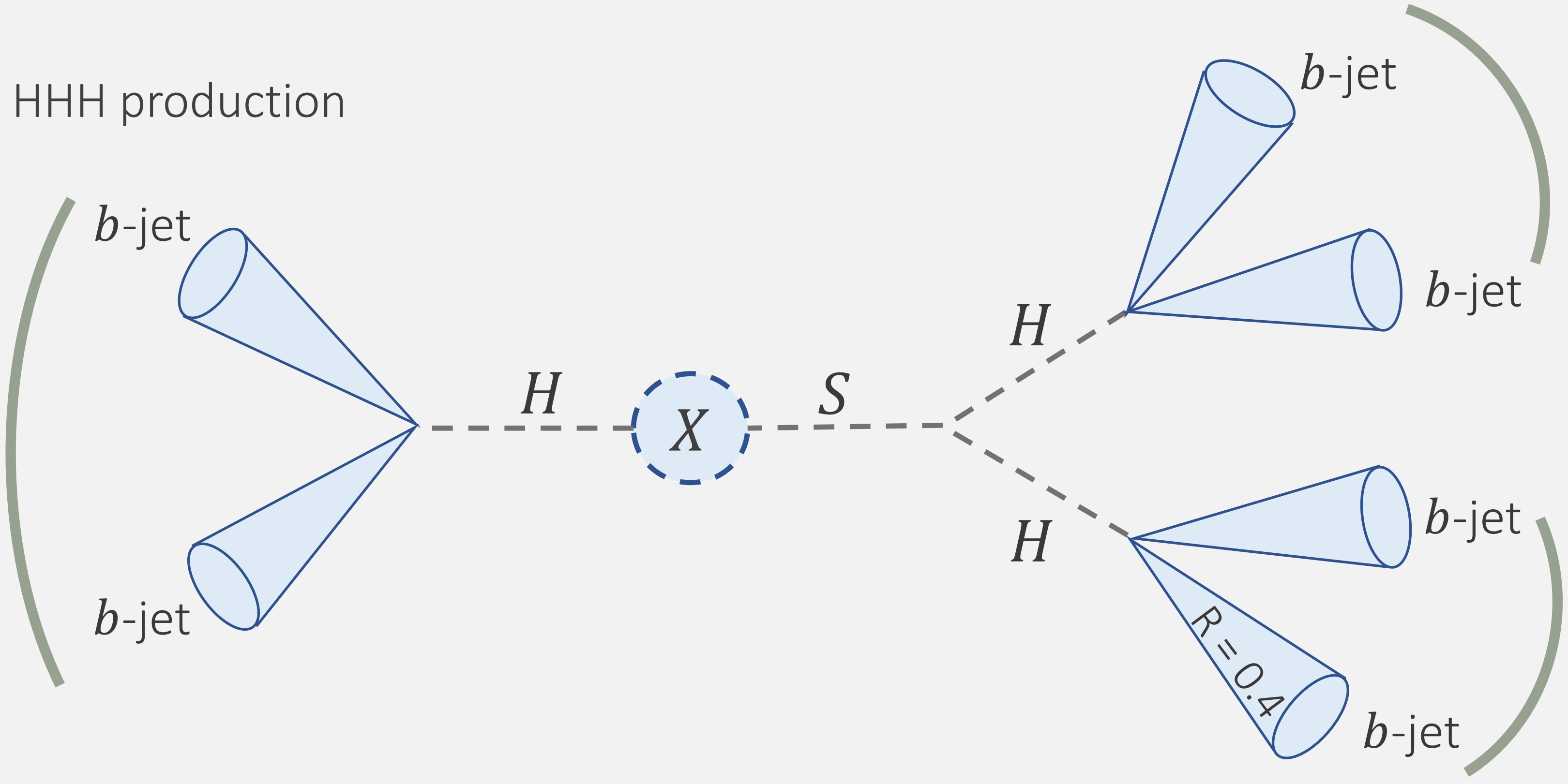
Resonant TRSM HHH production



Jet Pairing

- Jet pairing
  - b-jets are paired into 3 Higgs candidates
  - Such that the invariant mass of each pair is close to the SM m_H
  - Interesting Higgs-level kinematics can be reconstructed

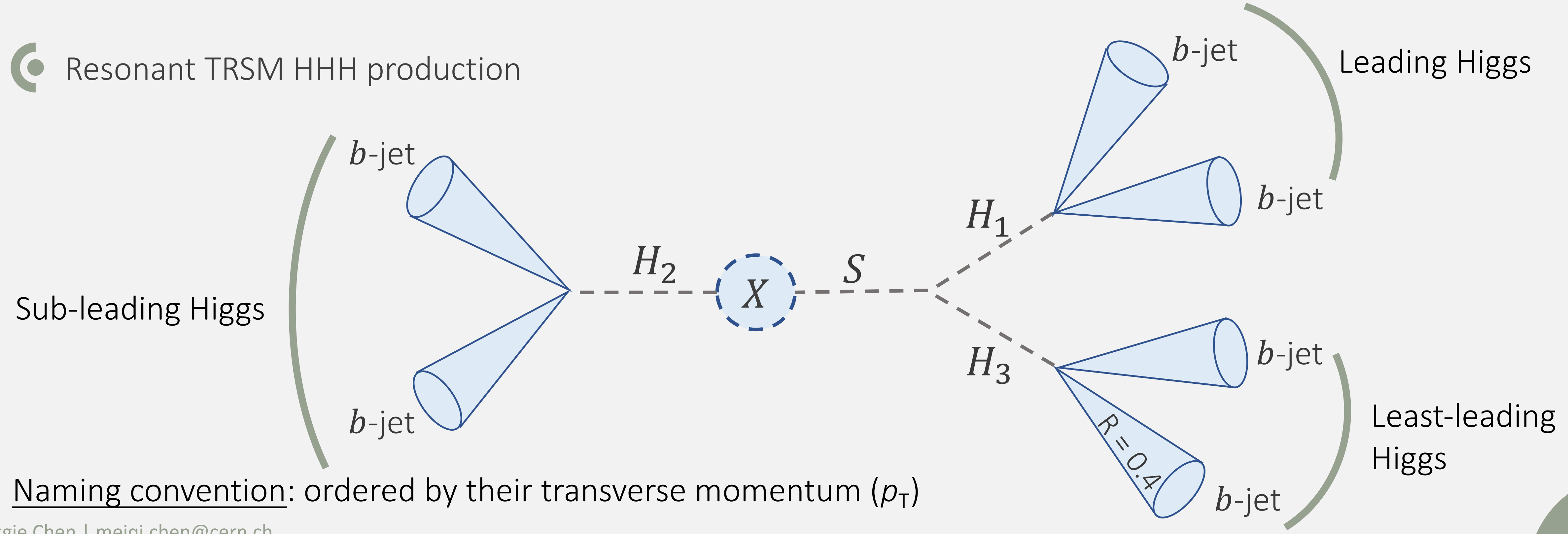
- Resonant TRSM HHH production



Jet Pairing

- Jet pairing
  - *b*-jets are paired into 3 Higgs candidates
  - Such that the invariant mass of each pair is close to the SM m_H
  - Interesting Higgs-level kinematics can be reconstructed

- Resonant TRSM HHH production



Naming convention: ordered by their transverse momentum (p_T)

Event Categorisation

ATLAS Run-2 data

Selected by triggers requiring jets + *b*-jets

$\geqslant$ 6 jets

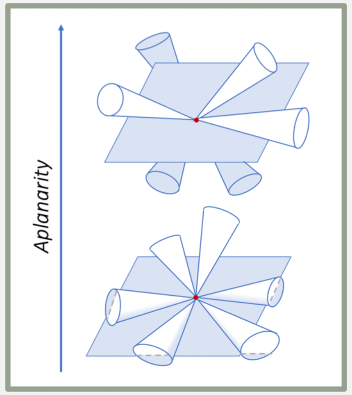
$\geqslant$ 4 jets with $p_{\mathrm{T}} > 40$ GeV

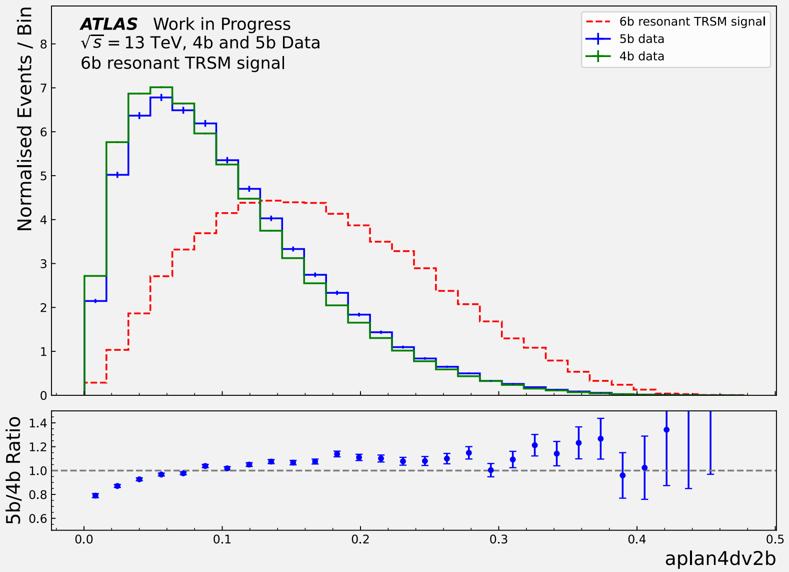
$\geqslant$ 4 *b*-tagged jets

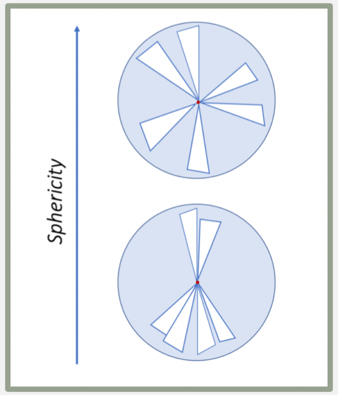
$\geqslant$ 6 *b*-tags	== 5 *b*-tags	== 4 *b*-tags
Signal Region (blinded)	Background estimation & validation	Background estimation & validation

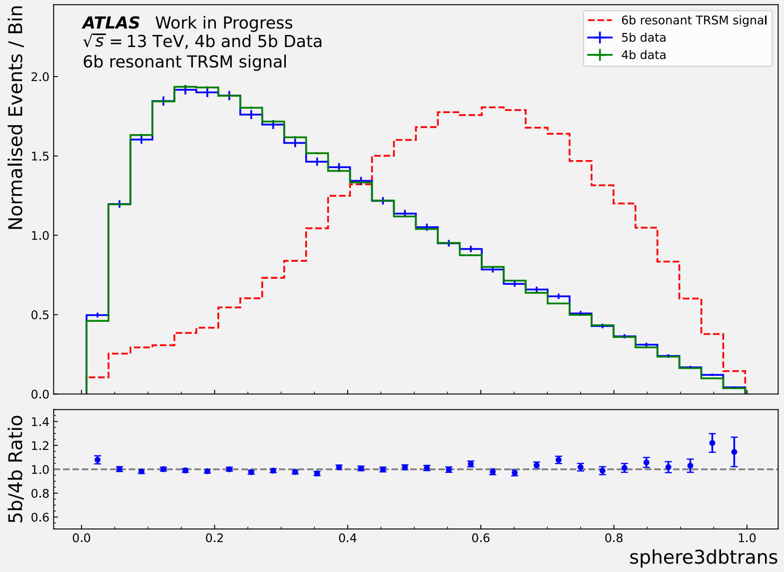
Signal / background classification

- Deep Neural Network (DNN)
  - Trained on signal (simulated 6*b* signal events) and background (5*b* data events)
  - Separate DNNs trained for the 2 signal interpretations (resDNN & nonresDNN)
- Input features
  - Pairing-independent features (e.g. H_T of the 6 jets, **aplanarity, sphericity**)
  - Pairing-dependent features (e.g. angular separation ΔR between the *b*-jet pair of a Higgs)





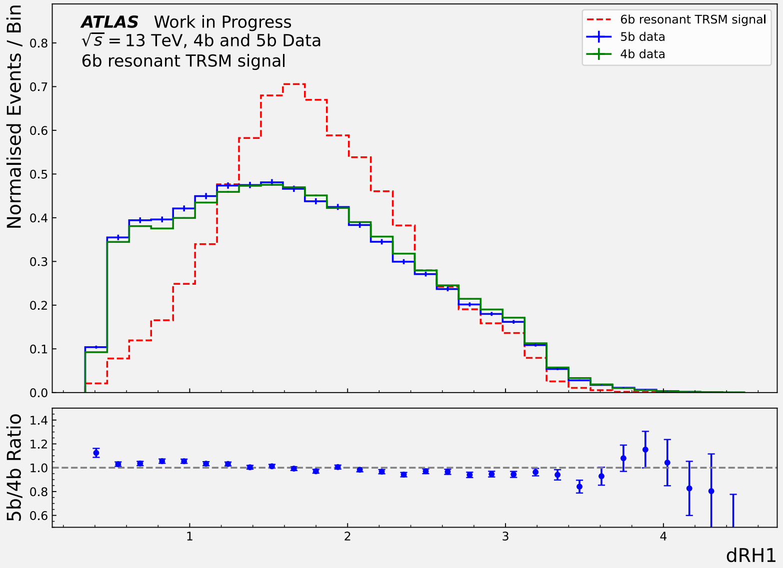


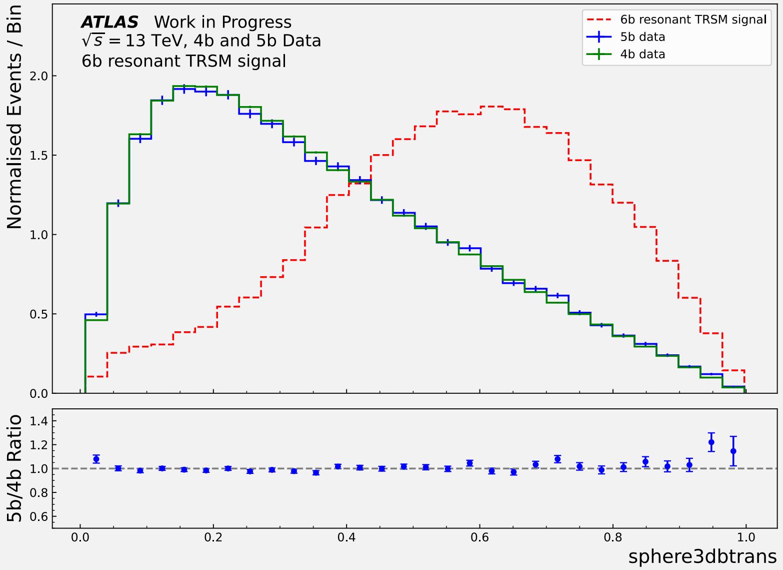


Signal / background classification

- Deep Neural Network (DNN)
  - Trained on signal (simulated 6*b* signal events) and background (5*b* data events)
  - Separate DNNs trained for the 2 signal interpretations (resDNN & nonresDNN)
- Input feature selection
  1. Feature Importance – maximises Shapley value ranking
  2. Minimses pair-wise correlations – avoid overlapping information
  3. **Minimises *b*-tag dependence** – smooth background extrapolation

Flat ratios between 4b and 5b data

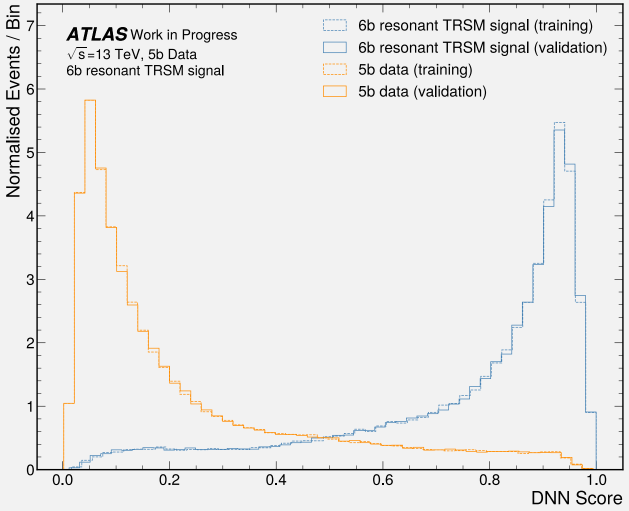


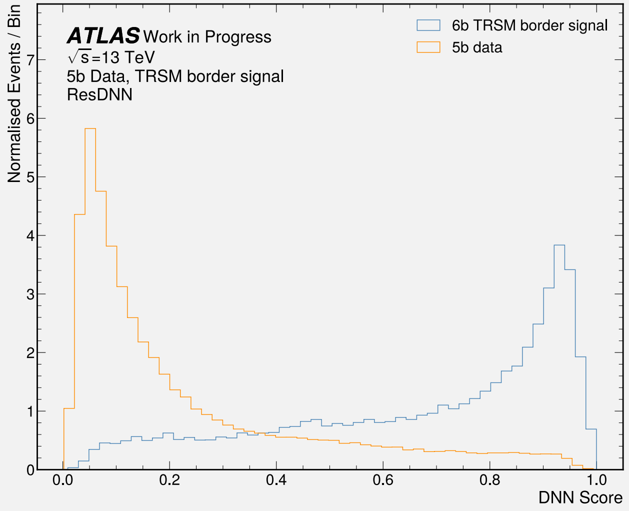


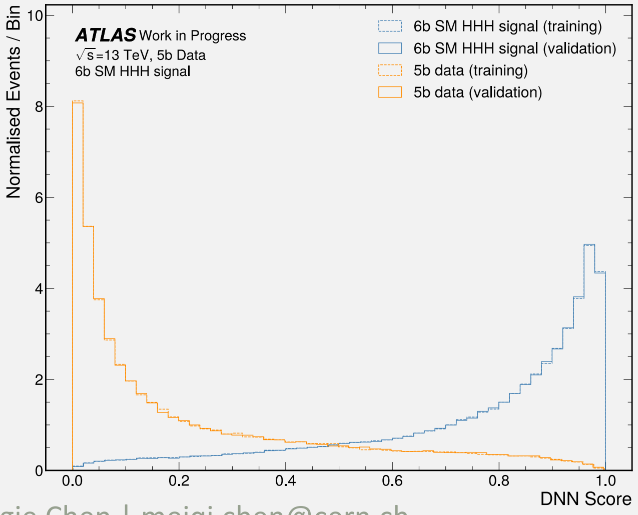
Signal / background classification

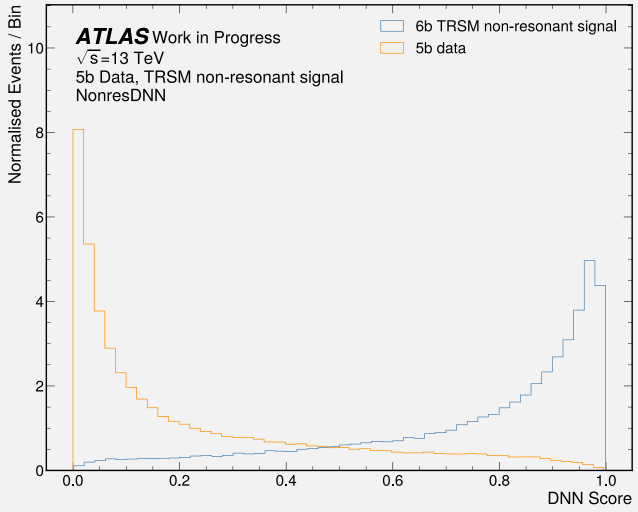
The DNN output distributions form the discriminating variables for statistical fitting

Resonant Interpretation – resDNN









Non-resonant Interpretation – nonresDNN

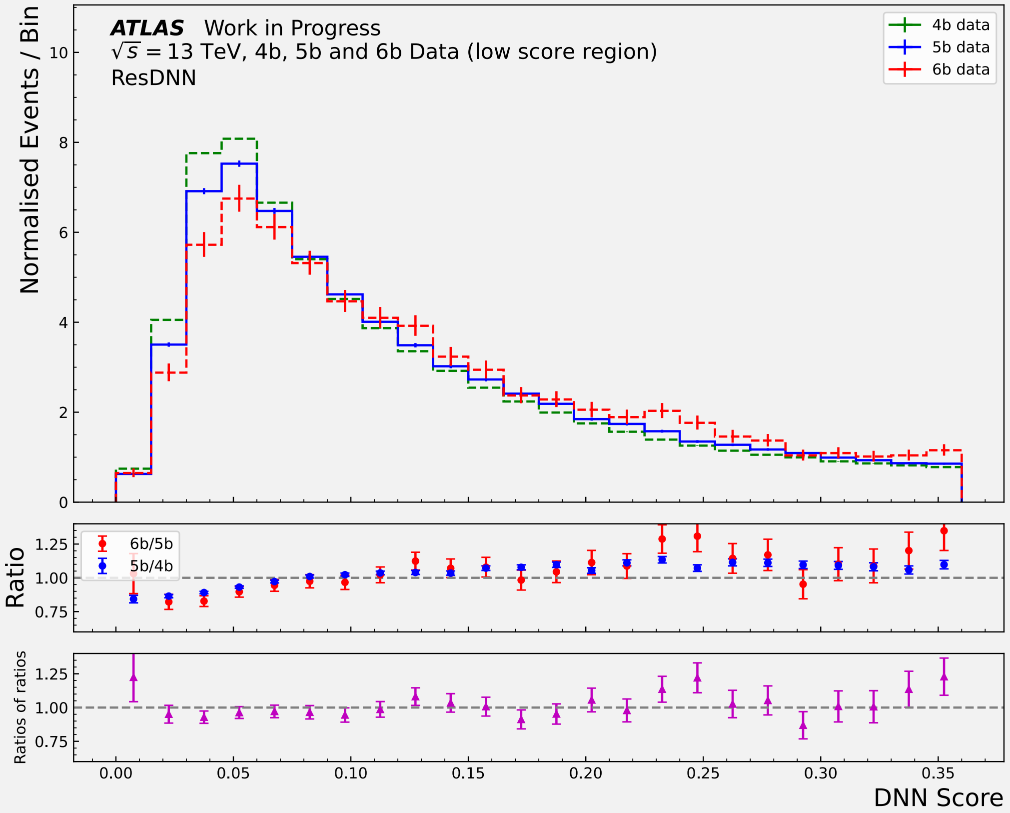
Background Estimation

What does the SM background look like in 6*b* events?

- Dominantly QCD multijet production (*b*-quarks from gluon splitting etc.)
- We assume it looks similar to 4*b* and 5*b* events
- Any kinematic difference is the same between 4*b* → 5*b* → 6*b* events

Unblinding 6b data in the DNN low-score region:

Non-unity high-tag to low-tag ratios:
→ some kinematic differences



But unity in ratios of ratios:
→ kinematic differences between 4*b*, 5*b* and 6*b* data are the same

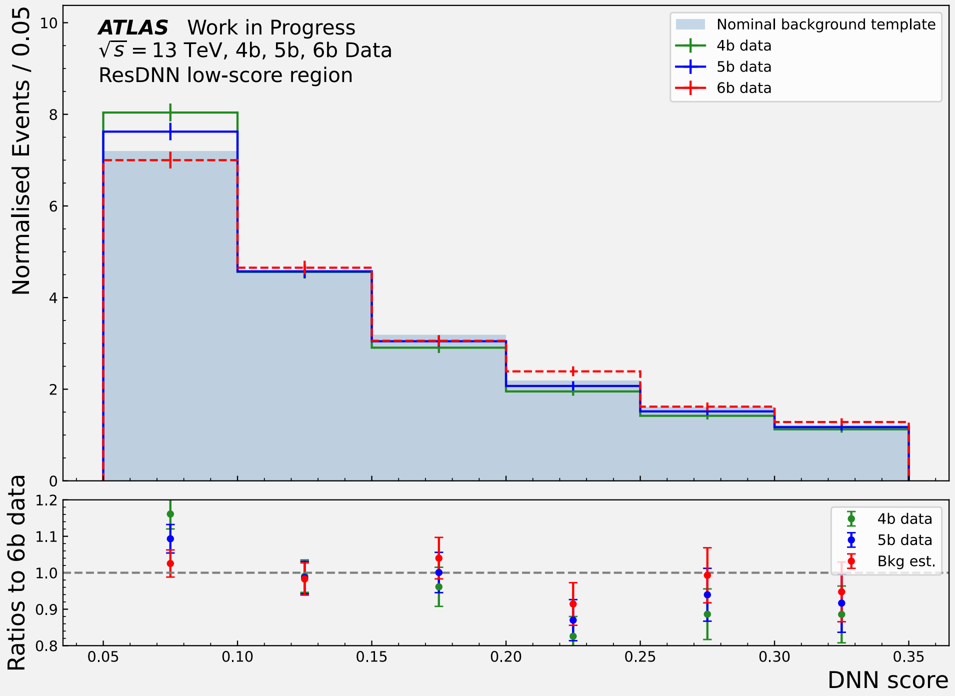
Background Estimation

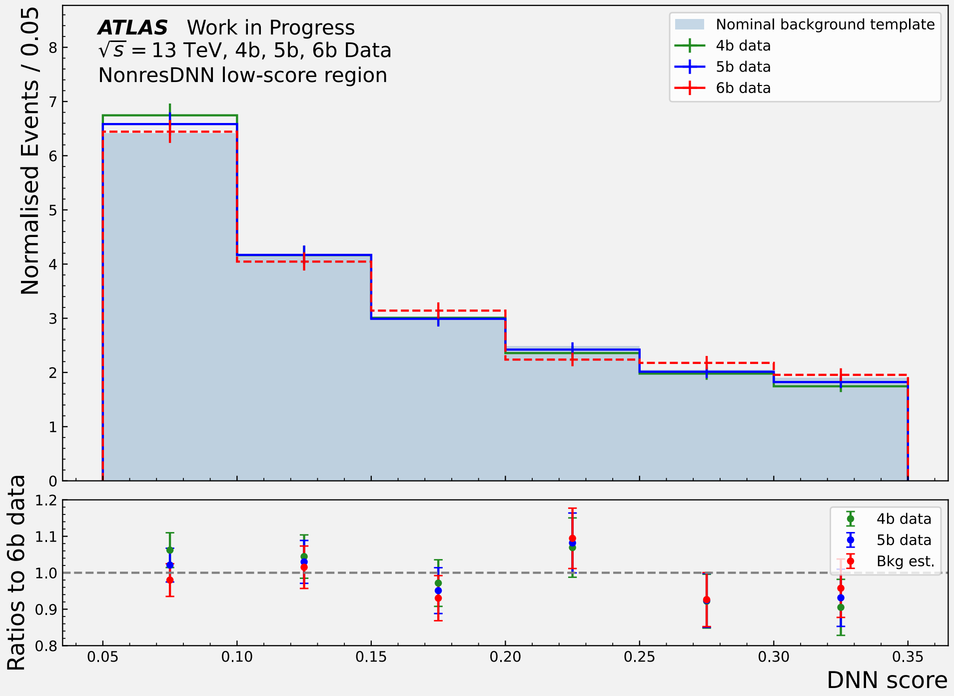
- Nominal background template:

$$Background_{6b} = \left[data_{5b} + \Delta shape_{data\ 4b/5b}\right] * R_{data\ 5b \to 6b}$$

Normalisation from $5b \to 6b$ data yields

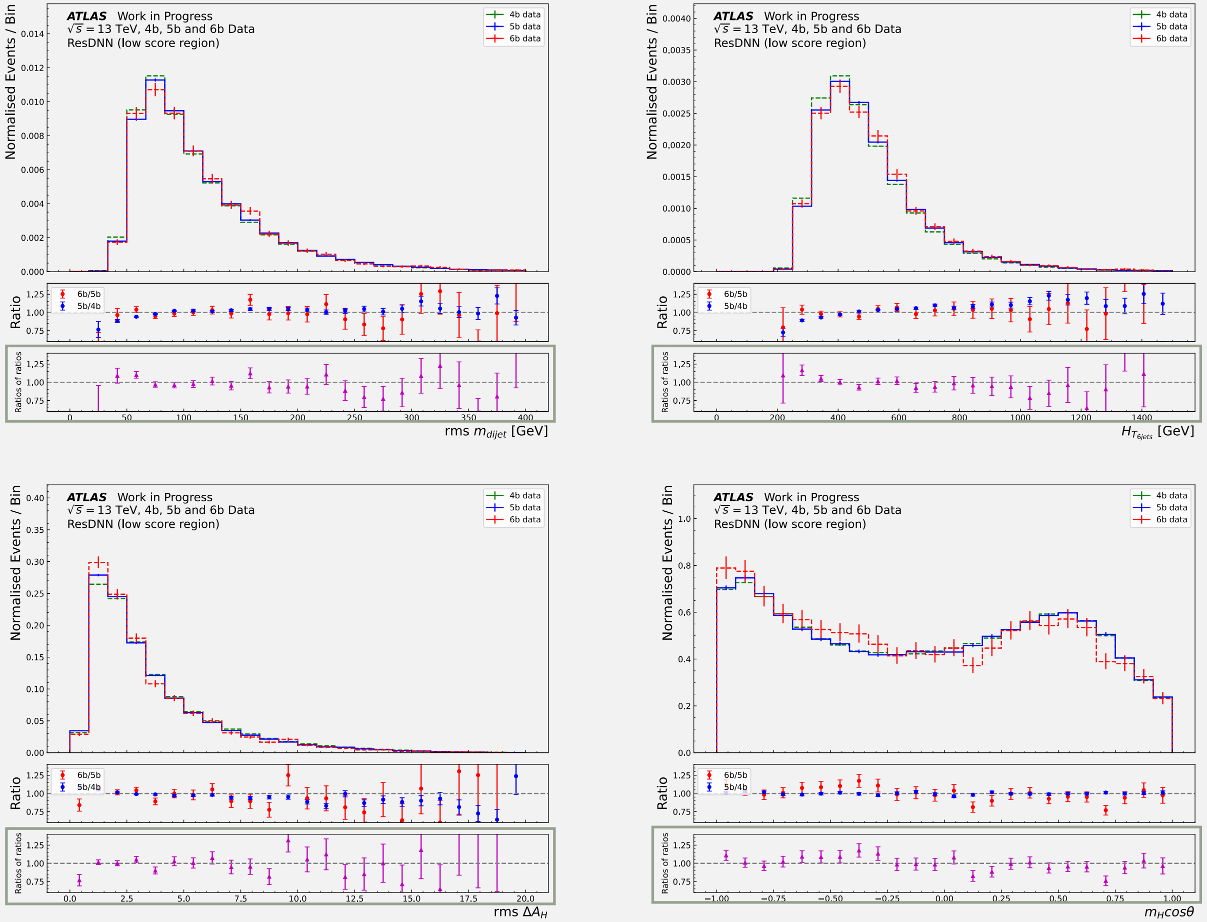
Background estimate in good agreement with 6b data in low-score region

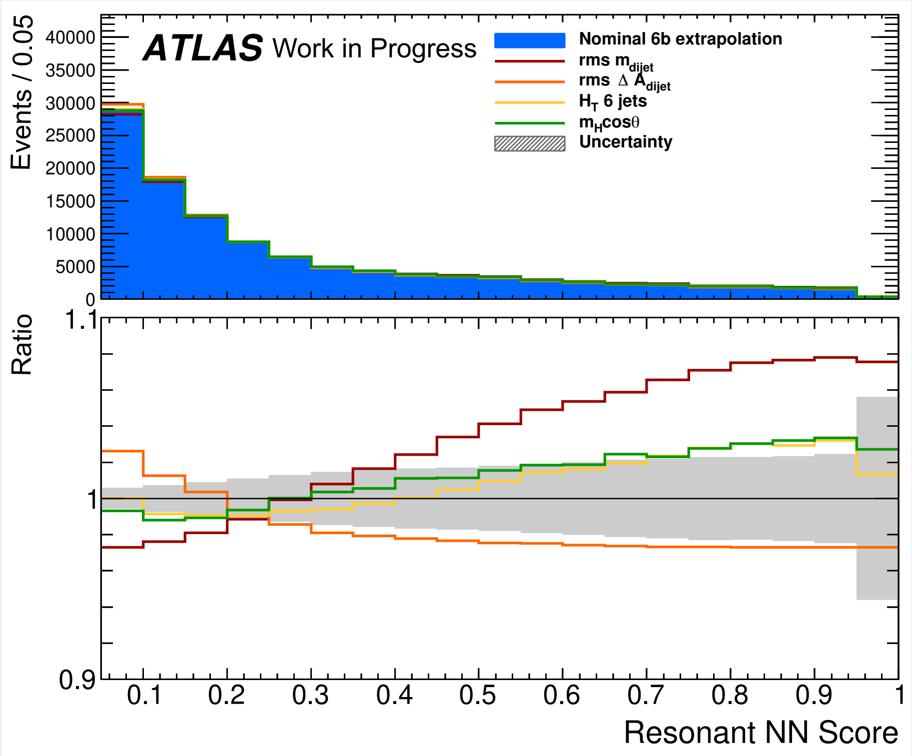




Uncertainties on Background Estimate

- DNN score variation broken down into representative variations from individual input features

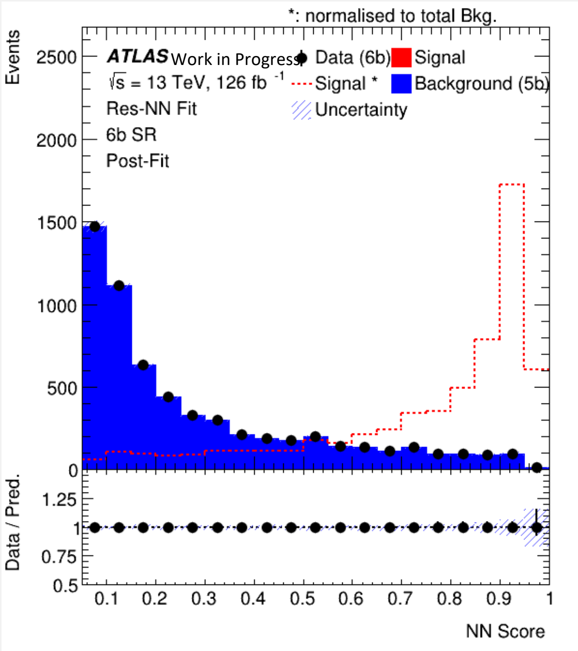




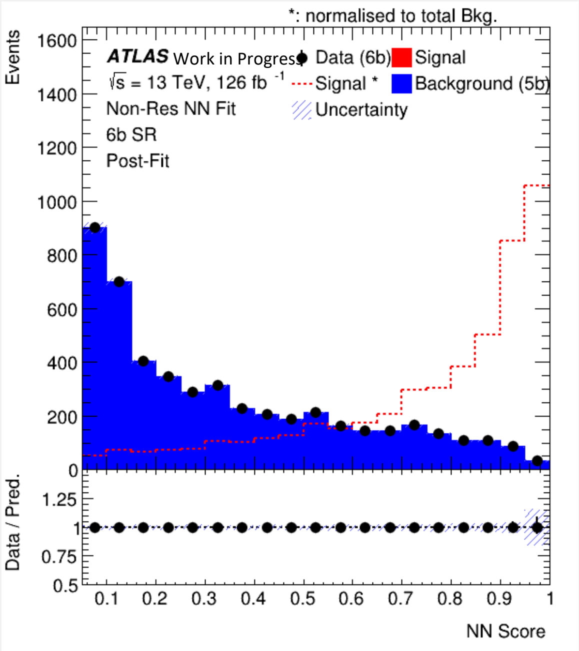
Expected Sensitivity to Signals

- Profile likelihood fits on the DNN output score for each mass hypothesis in the TRSM model and SM HHH signal
- Expected upper limits on cross-sections placed using Asimov fits (data = background)
  - Given a theory prediction of ~ 50 fb
- Uncertainties from detector, trigger, signal modelling and background estimation included

Resonant TRSM signal
(m_X, m_S) = (500, 300) GeV



SM HHH signal



Expected cross-section 95% CL limits for a few mass hypotheses:

***ATLAS** Work in Progress*

Interpretation	Resonant		Non-resonant	
m_X [GeV]	425	500	500	SM
m_S [GeV]	250	300	225	SM
Exp. upper limit [fb]	128.1	112.7	65.1	61.0

Summary

The first ATLAS search for triple Higgs boson production using Run-2 data

- Targeting a BSM production (TRSM) and SM production
- Using the 6*b* final state
- Dedicated DNNs trained to separate signal and background
- Data-driven method to estimate dominant QCD backgrounds
- Expected upper limits set on the production cross-sections of a range of mass hypotheses
- Next: unblinding and κ_3, κ_4 scans

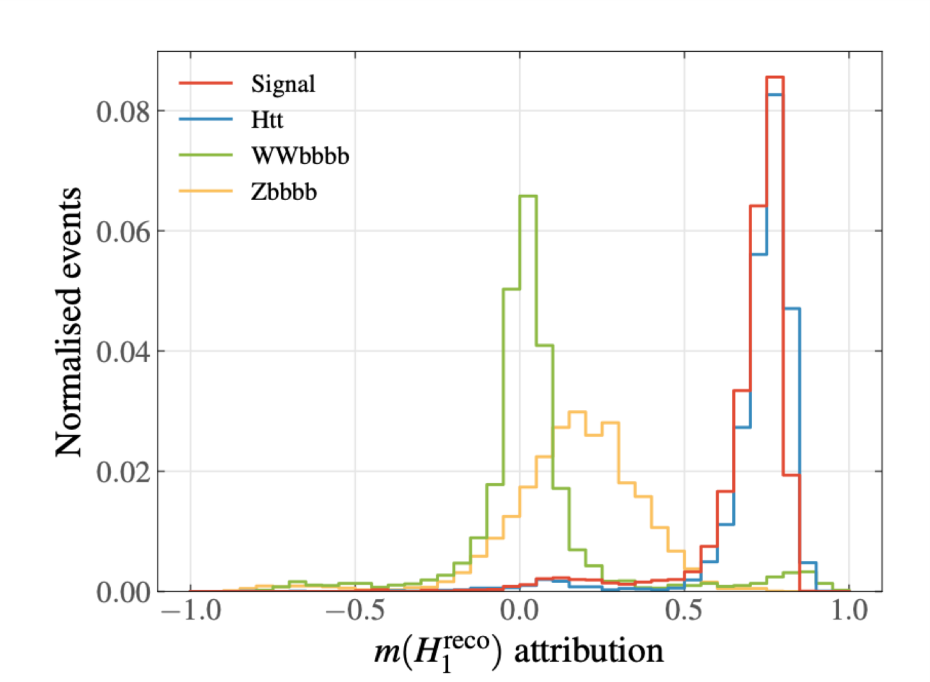
Backup

Other final states of HHH

- $4b2\tau$ is competitive
  - $BR(H \to b\bar{b}) = 0.584$
    - Background: QCD multijet production, *ttH*
  - $BR(H \to \tau^+\tau^-) = 6.627\times10^{-2}$
    - Backgrounds: *WWbbbb, Zbbbb, ttH, ttZ*

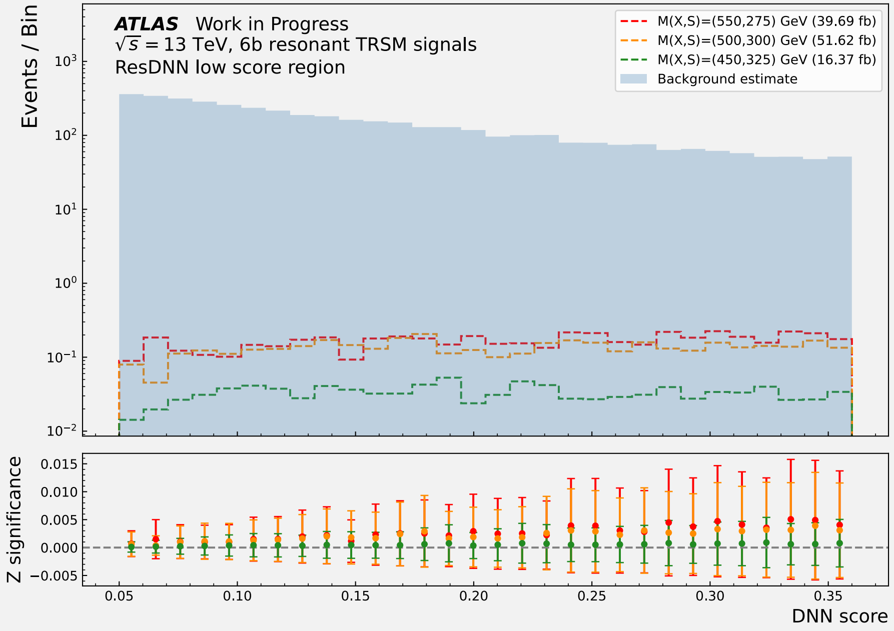
arXiv:2312.04646

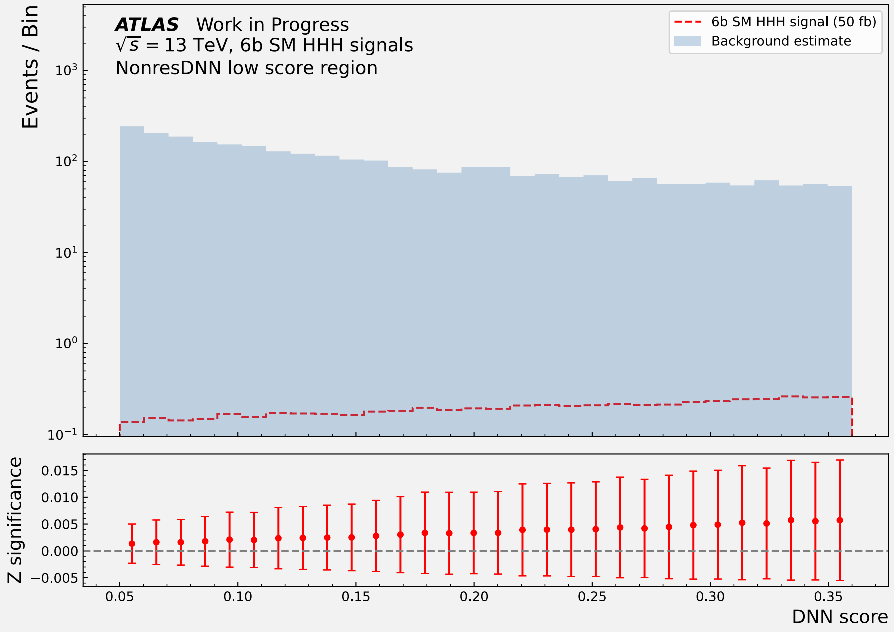
	σ(gen.)(fb)	σ(sel.)(fb)	σ(NN)(fb)
$tt(H \to \tau\tau)$	3.8	0.17	0.011
$WWbbbb$	31	4.6	8.1×10^{-3}
$tt(H \to bb)$	3.5	0.89	3.8×10^{-3}
$Zbbbb$	4.3	0.45	3.3×10^{-4}
$tt(Z \to bb)$	0.77	0.15	3.1×10^{-4}
$tt(Z \to \tau\tau)$	4.7	0.080	2.2×10^{-4}
$tttt$	0.38	0.091	2.1×10^{-4}



6b Signal Contamination Check in the DNN Low-score Region

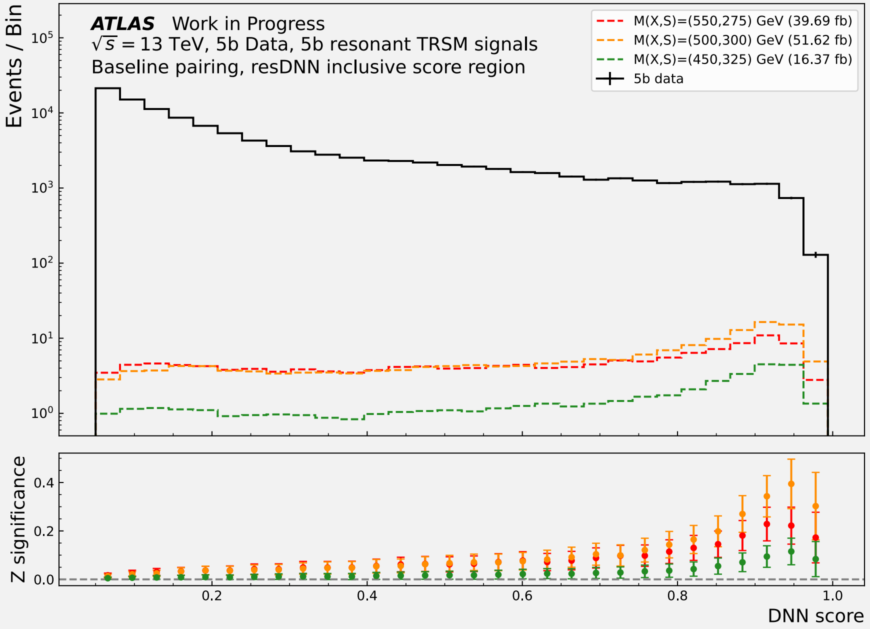
- Low-score region: $0.05 < \text{score} < 0.36$ for both resDNN and nonresDNN
- $\sim 10\%$ of $6b$ events from the combined TRSM signal / SM signal in this region of resDNN / nonresDNN

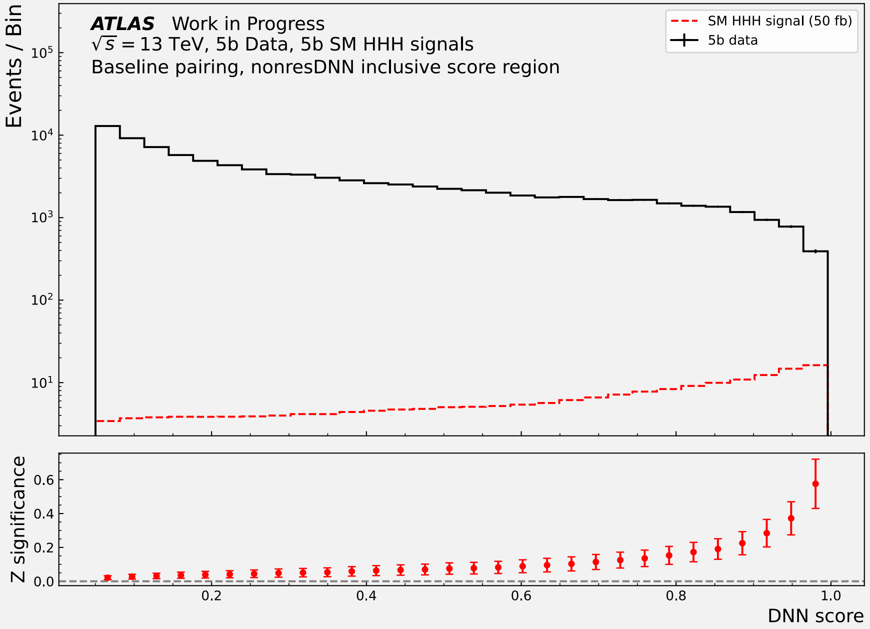




5b Signal Contamination

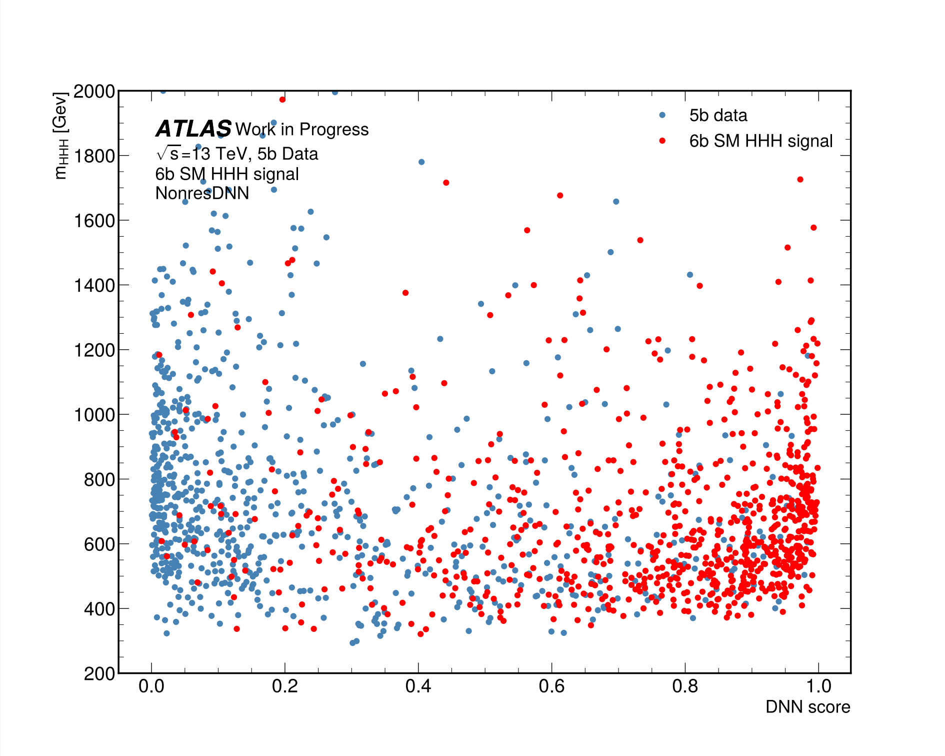
- In the score > 0.05 region, used in statistical fitting
- Small but non-negligible contamination from 5*b* data, especially in high-score
- Taken into account in the statistical fitting by correlating the parameter of interest (signal strength) in the 5*b* and 6*b* regions

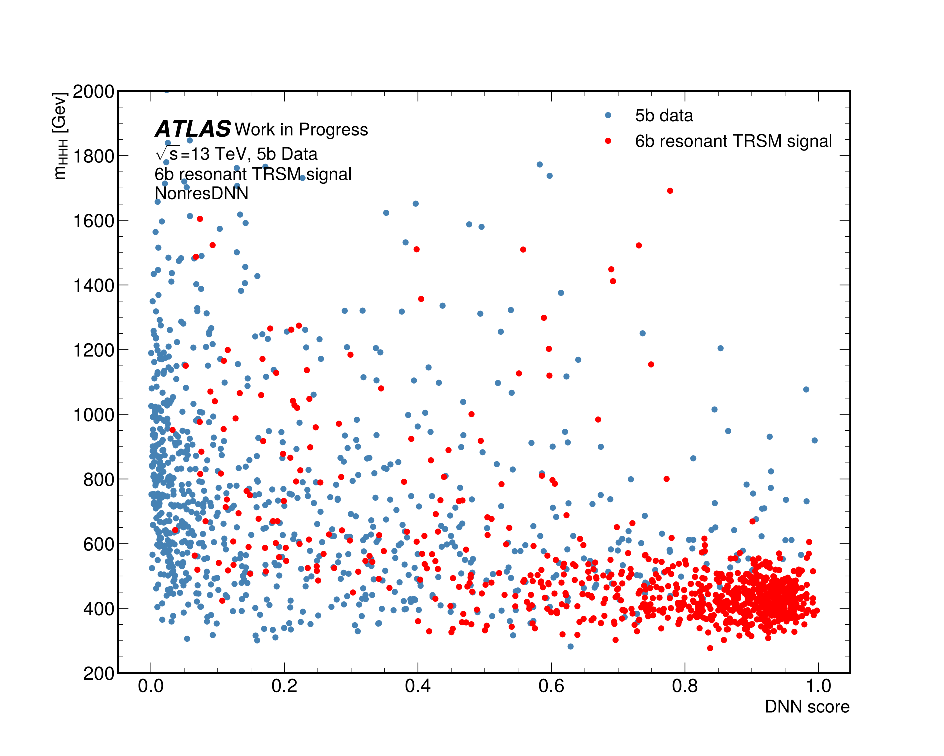




m_{HHH} dependence of the DNNs

- ResDNN is more dependent on the m_{HHH} of the events than nonresDNN
- SM HHH events have a m_{HHH} distribution that is more similar to 5*b* data background





m_H dependence of the DNNs

- The values m_H's are dependent on the jet pairing efficiency
- The resDNN and nonresDNN's dependence on m_H shows how sensitive they are to the correctness of jet pairing
- The DNN scores are very correlated with the m_H's
  - The closer to the (120, 115, 110) GeV point, the higher the DNN scores