

# Anomaly Detection at the ATLAS experiment

Antonio D'Avanzo, on behalf of the ATLAS collaboration

University of Naples Federico II and INFN Naples

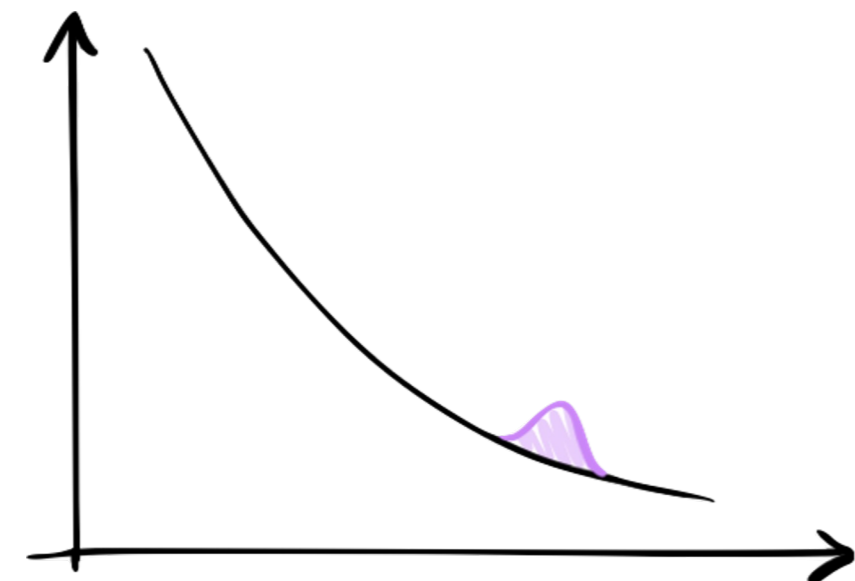
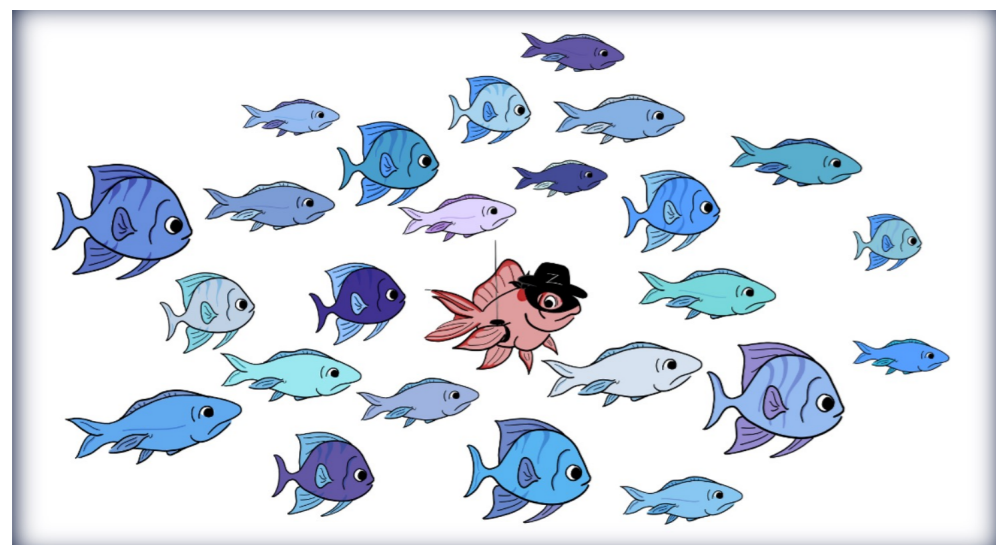
INFN

After discovering the Higgs boson at the LHC [1], no major breakthroughs in Beyond Standard Model physics have occurred. Anomaly Detection, could complement research by broadening the phase space while maintaining sensitivity to potential signals. ATLAS performed **weakly supervised** resonant searches and unsupervised searches, selecting events based on deviations from a background model. Focus is on finding heavy particles decaying into other particles in full Run 2 of LHC data (2015-2018,  $139 \text{ fb}^{-1}$ ), with particular attention to the first unsupervised Machine Learning application in ATLAS: a search with a SM Higgs boson  $H$  and a new particle  $X$  produced in the final state

## Anomaly Detection Preview

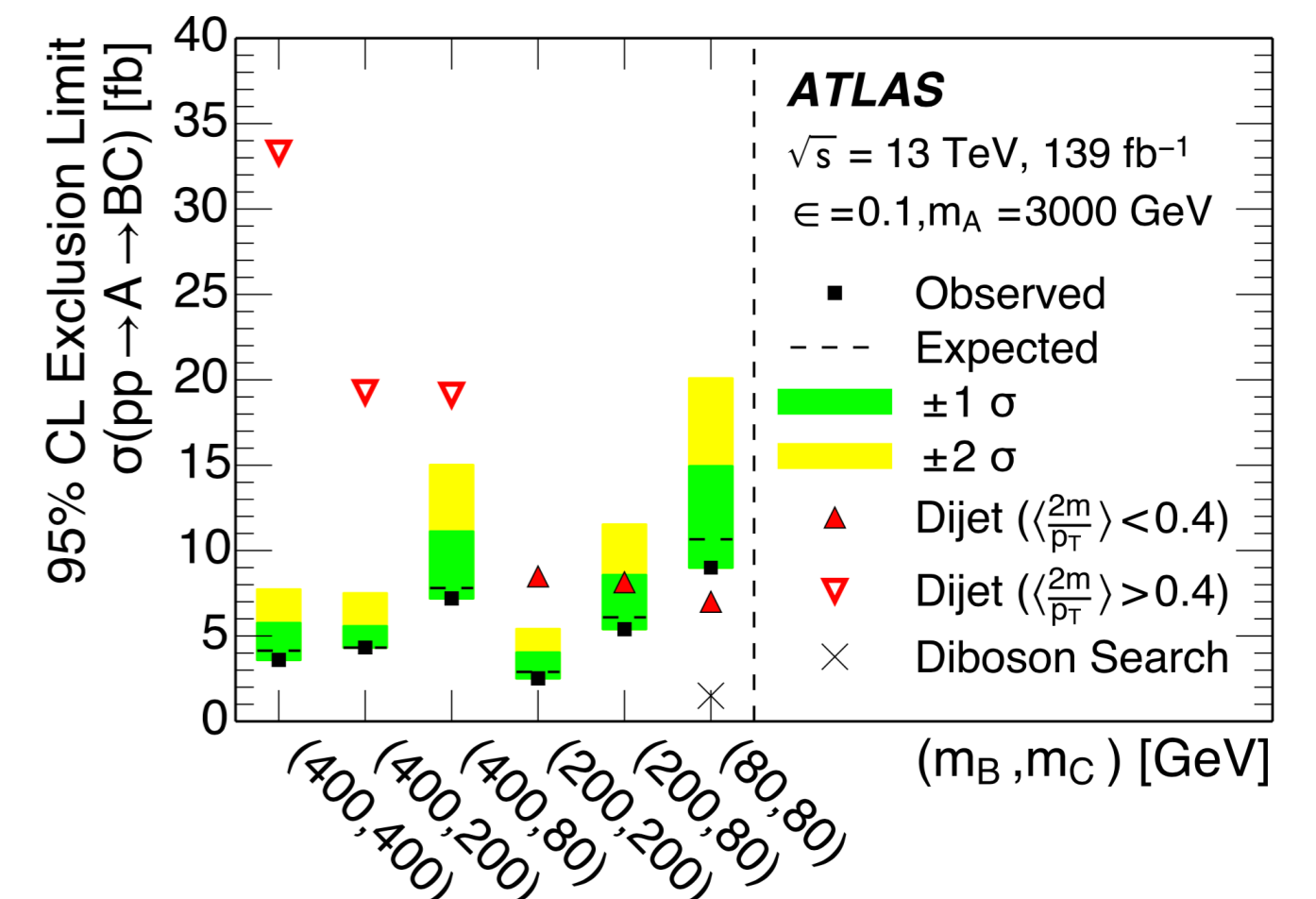
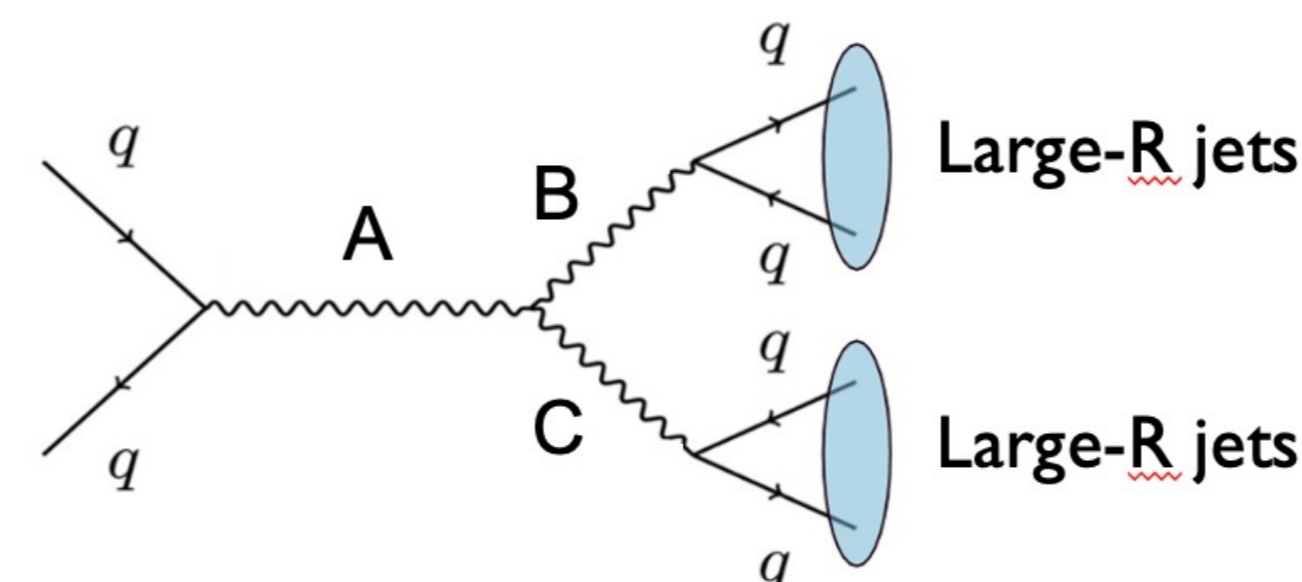
Anomaly Detection (AD) [3] consists in the use of ML algorithms to identify outliers in a set of "standard" objects.

HEP → identification of features of detector data inconsistent with the expected background (bumps)



Classification Without Labels (CWoLa) is the first AD search in ATLAS with a weakly supervised machine learning technique

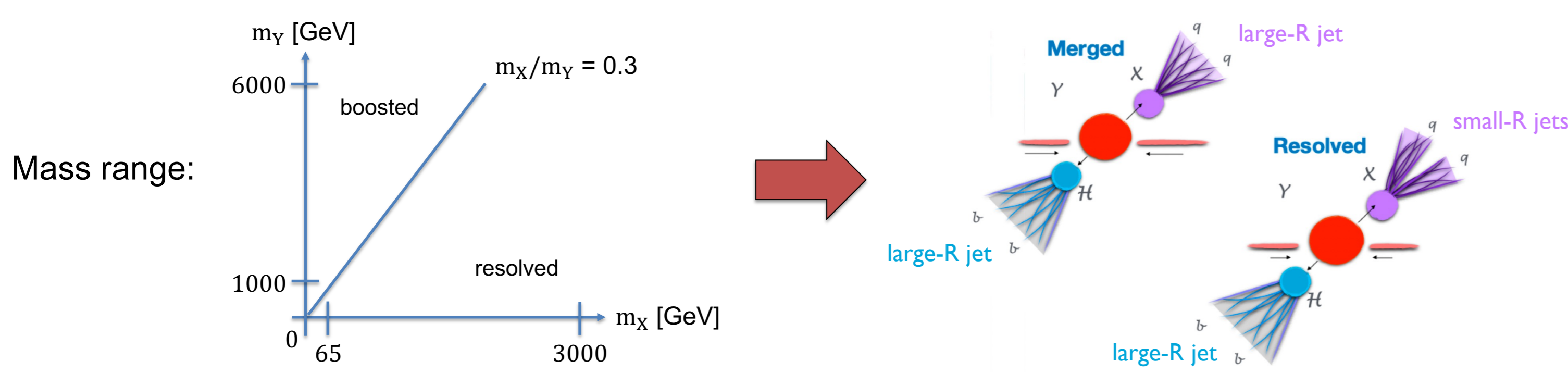
Carried on by ATLAS with a fully unsupervised based search



## Fully unsupervised search for new resonances decaying into a Higgs boson and a new particle X in hadronic final states [4]

### 1. Analysis overview

Search for a heavy-mass resonance  $Y$  decaying in a Higgs boson ( $H \rightarrow b\bar{b}$ ) and a new particle  $X$  in the fully hadronic channel



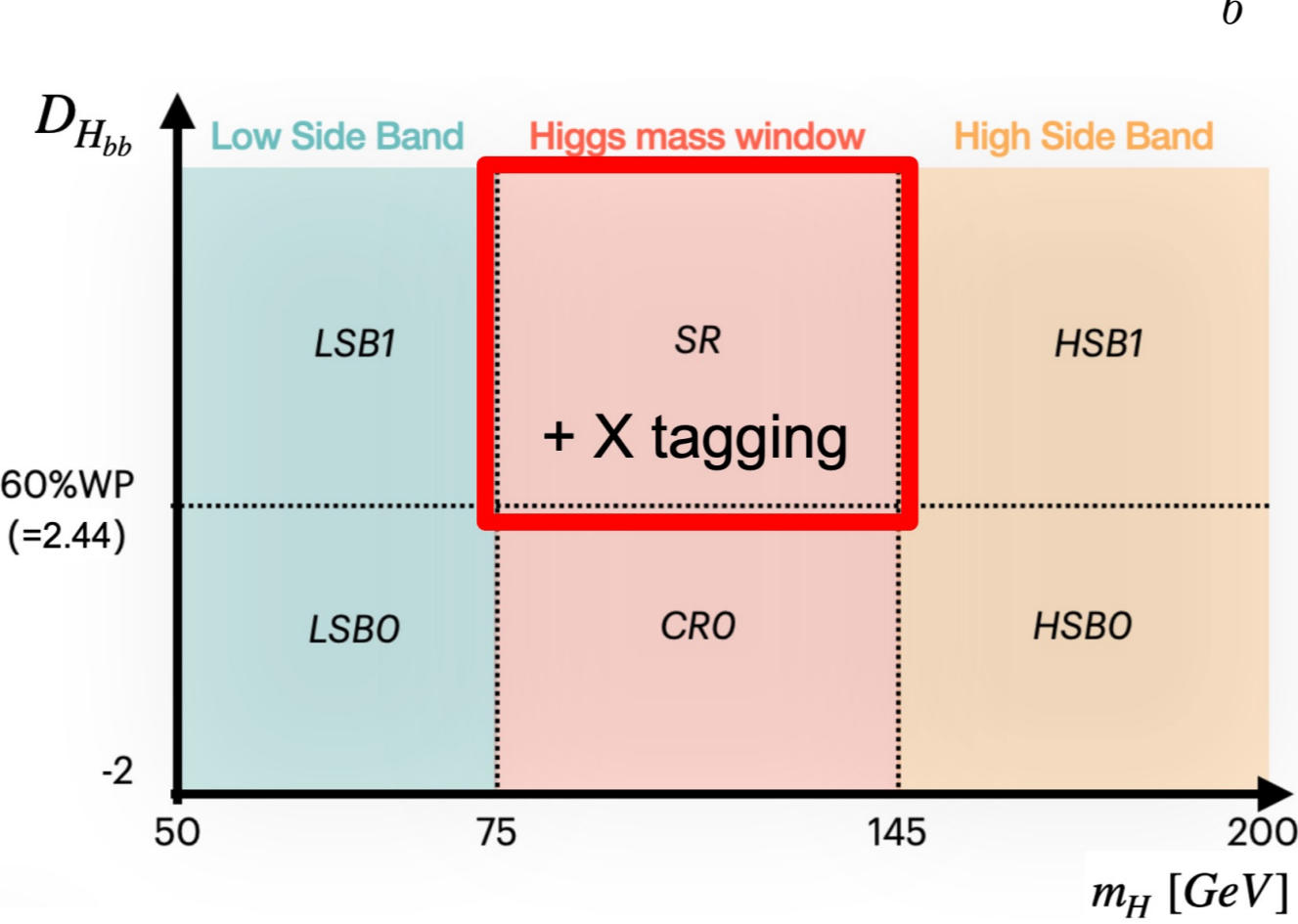
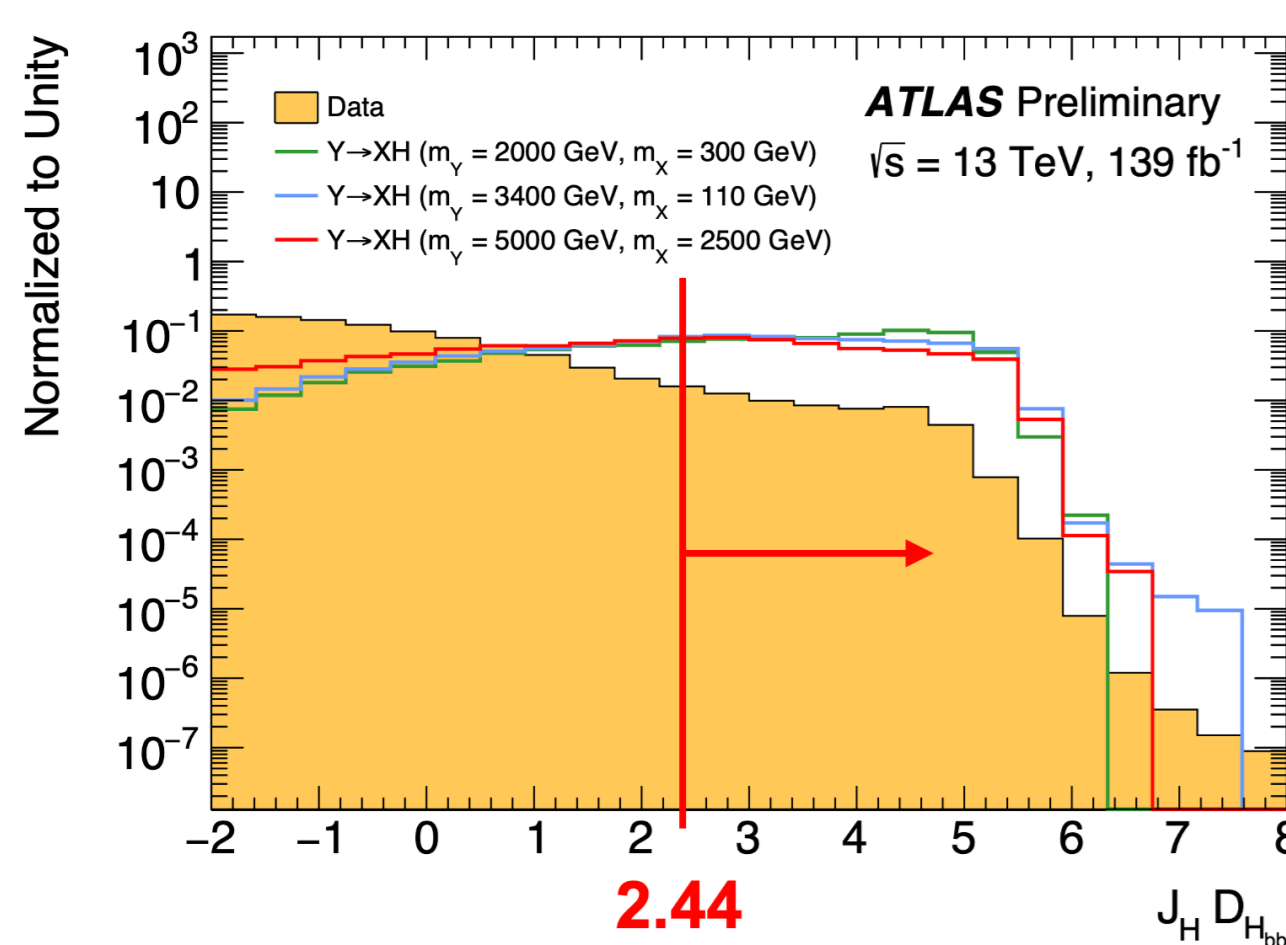
2 approaches to X tagging:

- Model dependent:  $X \rightarrow q\bar{q}$  boosted ( $m_X/m_Y < 0.3$ ) or resolved ( $m_X/m_Y > 0.3$ )
- Model independent: anomalous hadronic decay reconstructed in a **large-R jet** (!)

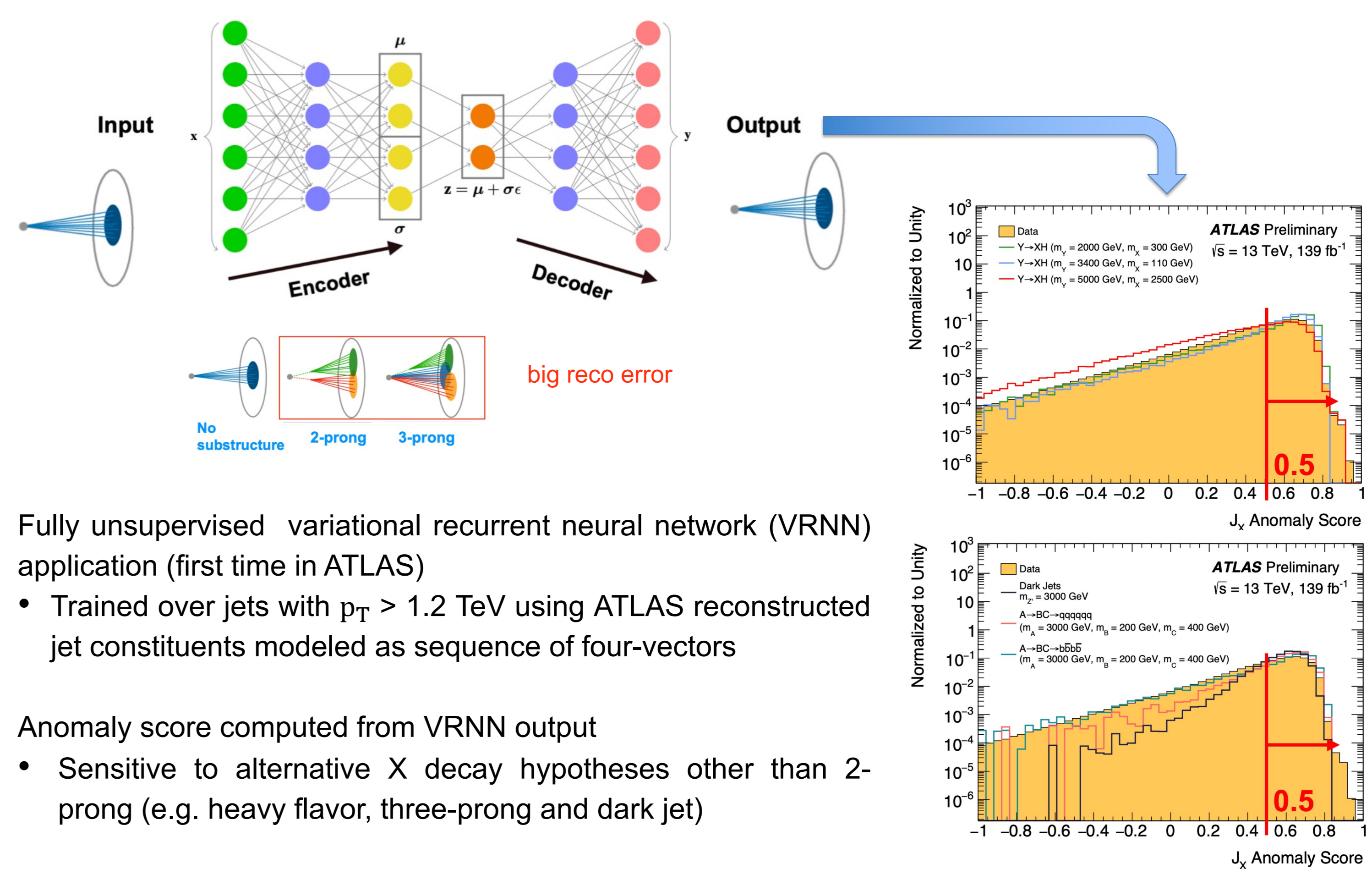
X and H candidates identification performed on the two most-leading jets.

Ambiguity resolved by DNN  $H \rightarrow b\bar{b}$  tagger[5],  $D_{Hbb}$  score computed

- H candidate chosen based on highest score



### 2. X tagging with unsupervised Anomaly Detection



Fully unsupervised variational recurrent neural network (VRNN) application (first time in ATLAS)

- Trained over jets with  $p_T > 1.2 \text{ TeV}$  using ATLAS reconstructed jet constituents modeled as sequence of four-vectors

Anomaly score computed from VRNN output

- Sensitive to alternative X decay hypotheses other than 2-prong (e.g. heavy flavor, three-prong and dark jet)

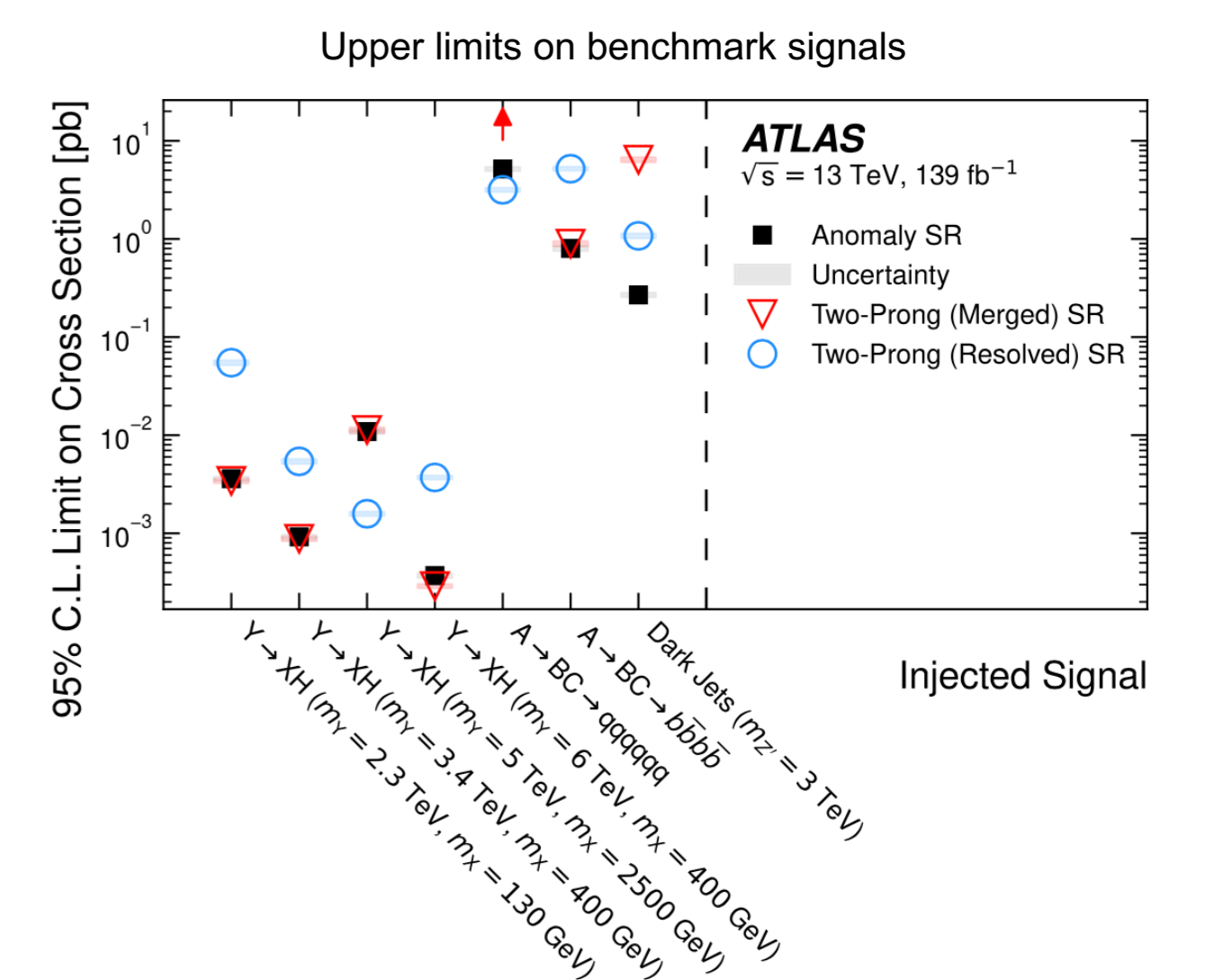
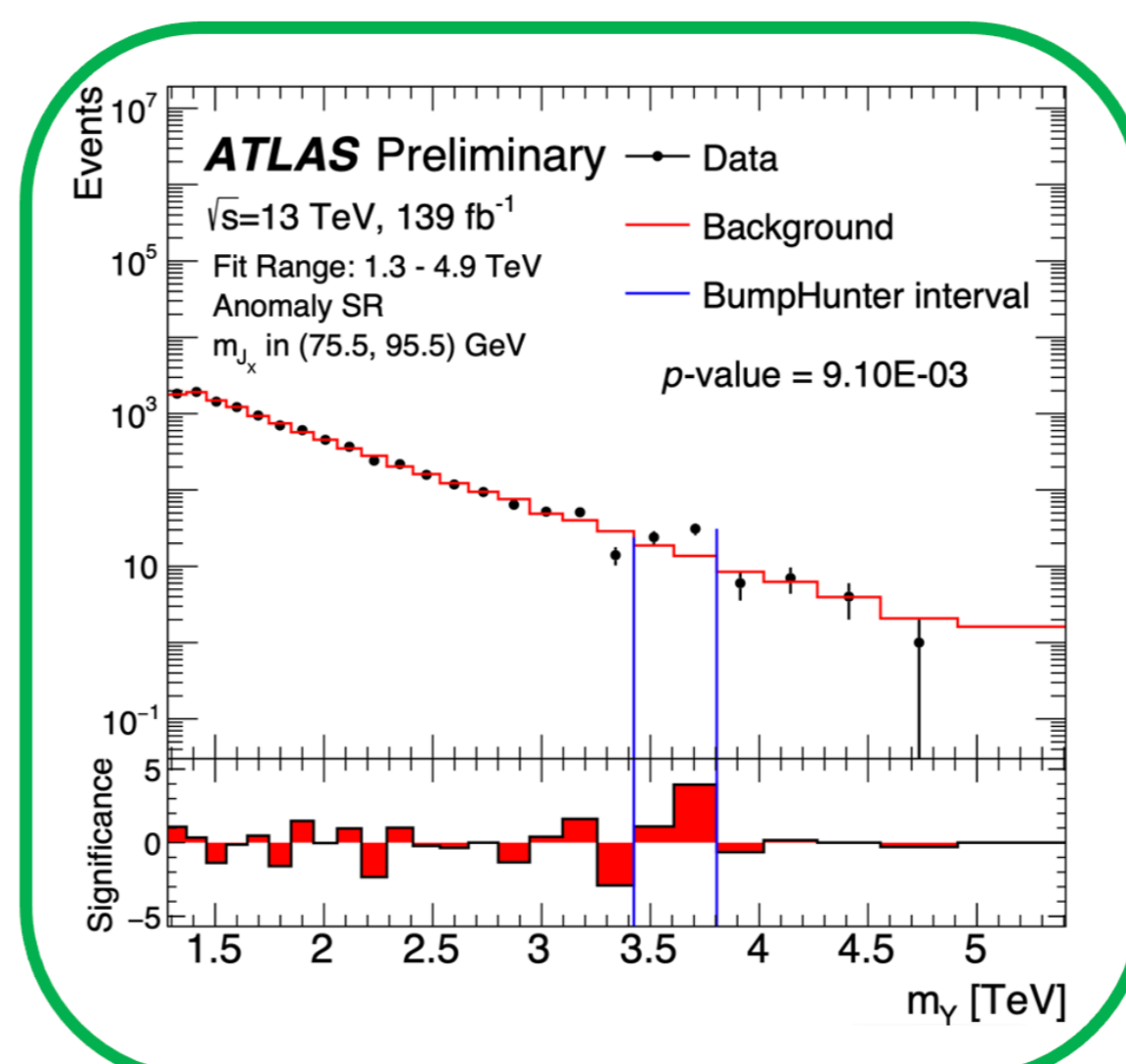
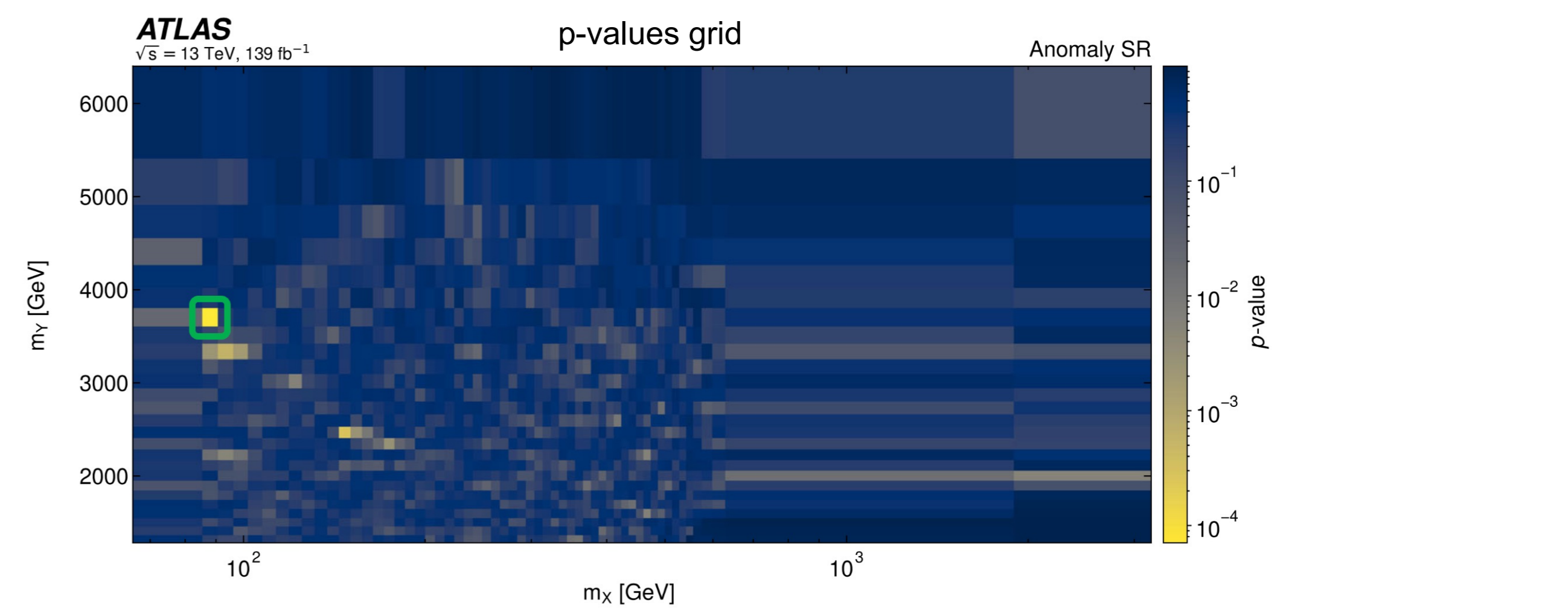
### 4. Results

Fit performed on final state invariant mass distribution  $m_{JJ}$  in SR of data, repeated several times in overlapping bins of the X candidate mass

- Many signal regions defined for each  $(m_X, m_Y)$  bin

Calculated stat-only p-values to test compatibility with background only hypothesis using BumpHunter [7]

Max deviation:  $\sim 3\sigma$  local significance, becoming  $1.43\sigma$  global significance

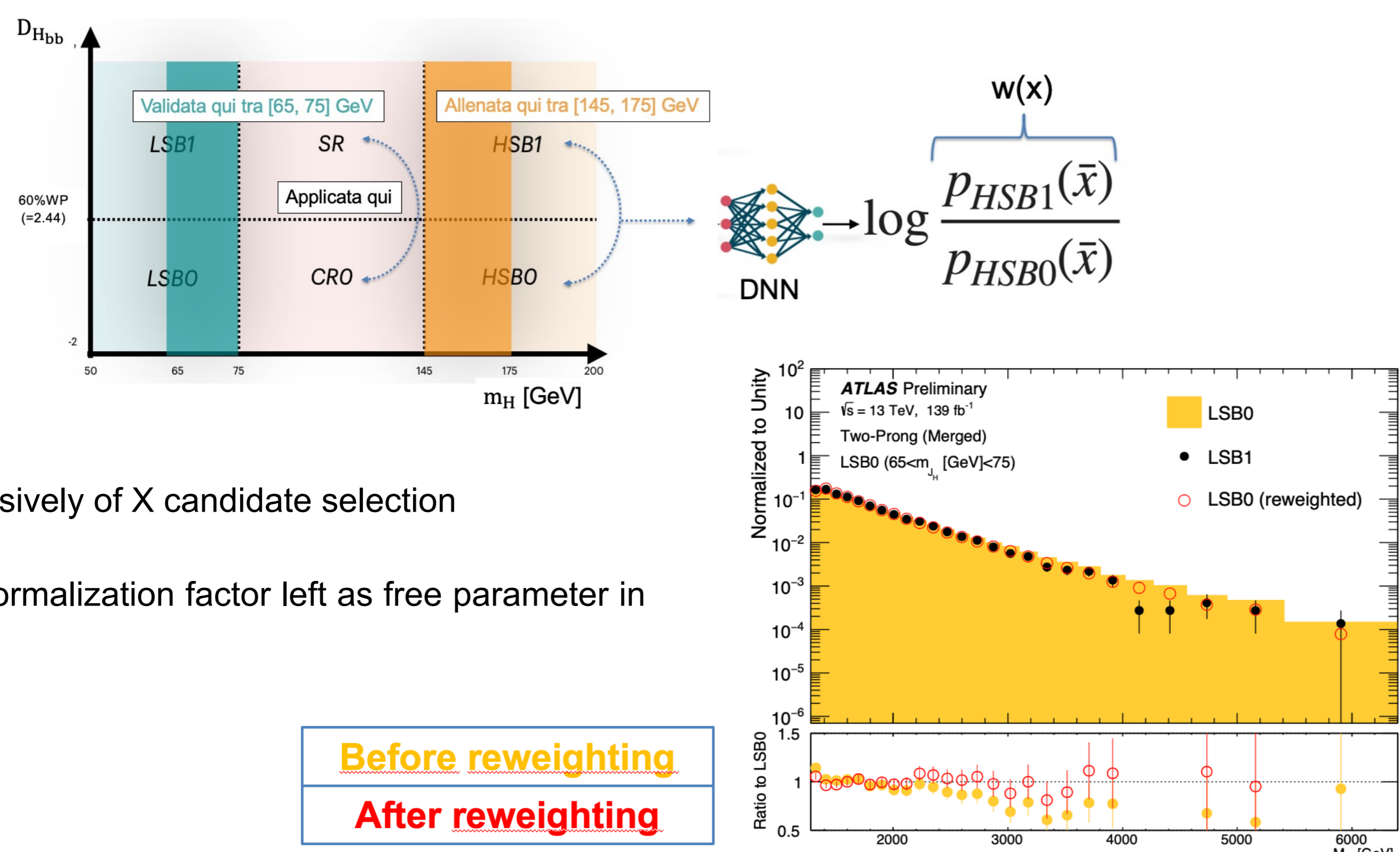


### 3. Background estimation

The background is mainly QCD dijet events ( $\sim 97\%$ )

Signal Region contribution is fully data-driven estimated with Machine Learning technique [6]

- Based on events reweighting by a function  $w(x)$  from control region CR0 to SR



Validated inclusively of X candidate selection

Background normalization factor left as free parameter in final fit

Before reweighting  
After reweighting



Bibliography



1. [j.physletb.2012.08.020](https://arxiv.org/abs/2012.08.020)

2. [Phys.Rev.Lett.125.131801](https://arxiv.org/abs/125.131801)

3. [j.revip.2024.100091](https://arxiv.org/abs/2024.100091)

4. [Phys.Rev.D108.052009](https://arxiv.org/abs/Phys.Rev.D108.052009)

5. [ATL-PHYS-PUB-2020-019](https://arxiv.org/abs/ATL-PHYS-PUB-2020-019)

6. [JournalofMachineLearningResearch10\(2009\)1391-1445](https://arxiv.org/abs/JournalofMachineLearningResearch10(2009)1391-1445)

7. [arXiv:1101.0390](https://arxiv.org/abs/1101.0390)