Search for new physics using unsupervised machine learning for anomaly detection in $\sqrt{s} = 13$ TeV pp collisions recorded by the ATLAS detector at the LHC

Elham E Khoda

University of Washington

on behalf of the ATLAS Collaboration

BOOST 2023
August 1, 2023

UNIVERSITY *of* WASHINGTON

a3d3 Accelerated AI Algorithms for Data-Driven Discovery

ATLAS EXPERIMENT

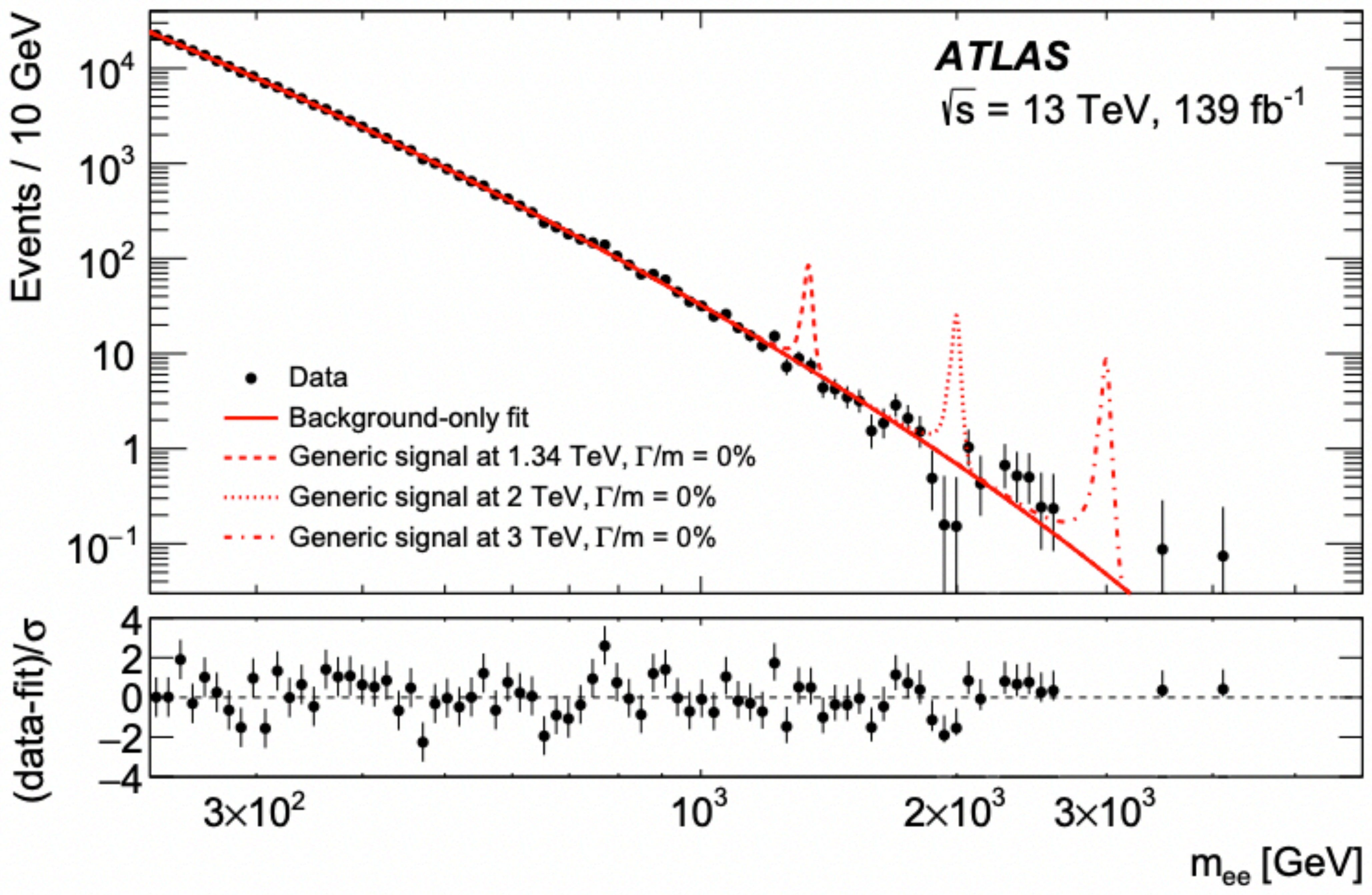
ATLAS Searches: Model Dependent

Two general approaches:

- Model dependent and Model independent Searches

- Choose a well-motivated New Physics Scenario with a signal model
- Maximize the search sensitivity based on signal properties
- Most sensitive approach for a particular new physics scenario
  —> Unlikely to probe a different scenario

ATLAS dilepton Resonance Search
[Phys. Lett. B 796 (2019) 68]

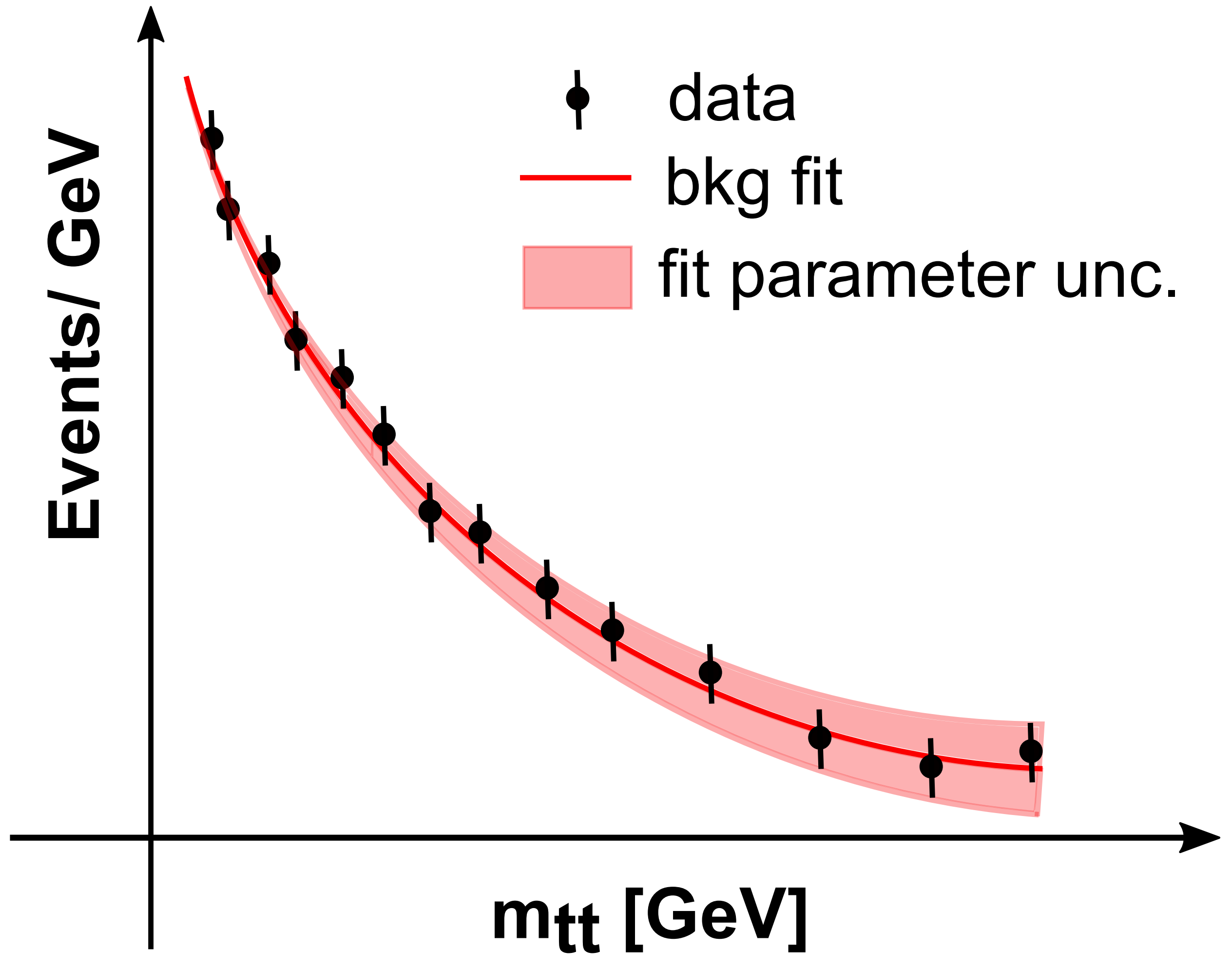


ATLAS Searches: Model Independent

Two general approaches:
- Model dependent and Model independent Searches

- Minimal set of assumptions on the signal properties
- Look for deviations from the background-only hypothesis while searching the data
- Not optimal for a specific signal hypothesis, but likely to be sensitive to wide range of signal models

Estimate the background (SM) and look for deviation in data



This talk will highlight some results based on model-independent anomaly search

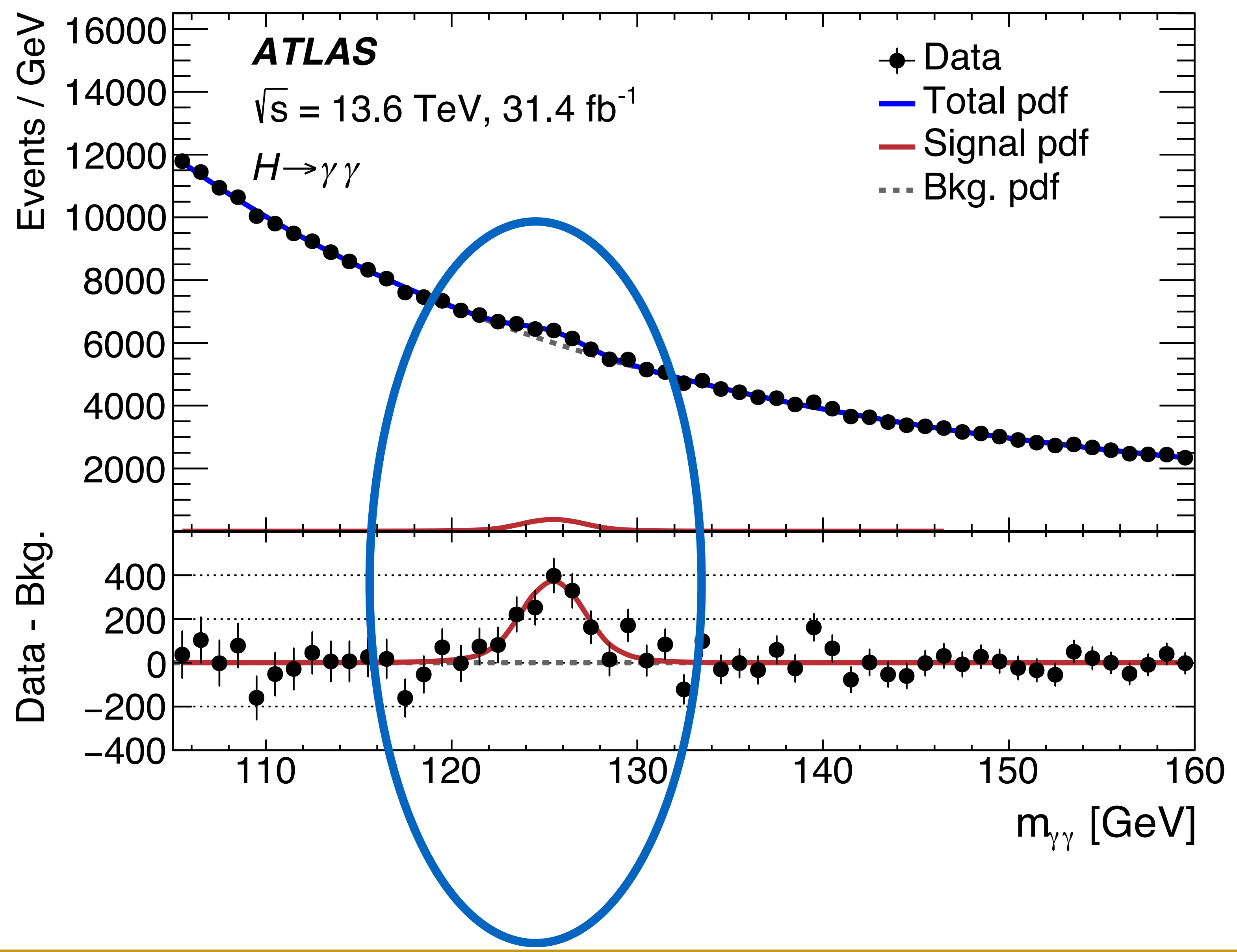
What is an Anomaly?

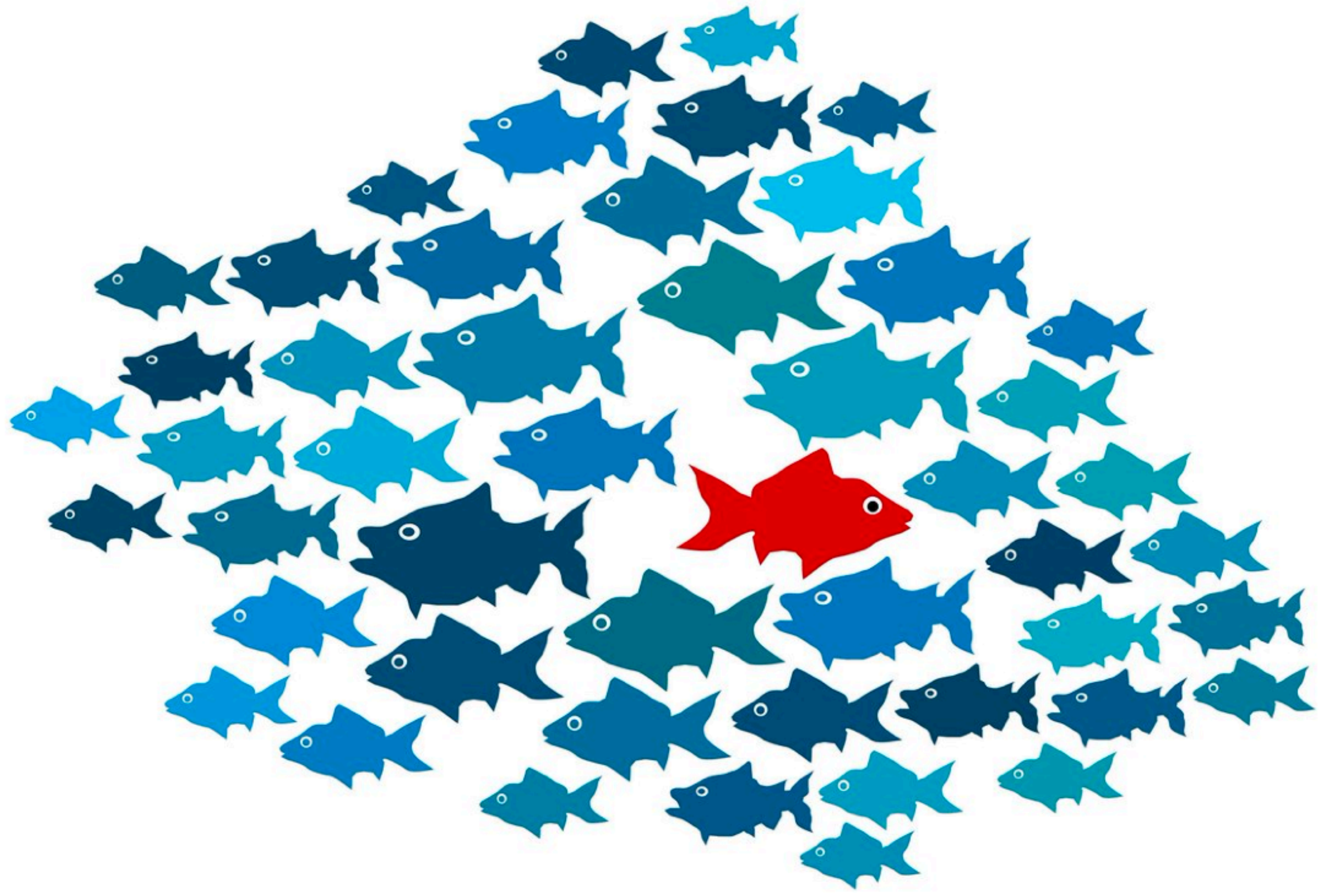
Anomaly detection: Identify data with features that appear inconsistent with those of the majority of the dataset

It is often associated to outlier detection

For HEP application: A New Physics signal will show up as an ensemble of events rather than a single outlier

arXiv:2306.11379





Usual approaches:

- *Unsupervised** = no labels
- *Weakly-supervised** = noisy labels
- Semi-supervised = partial labels

* in this talk

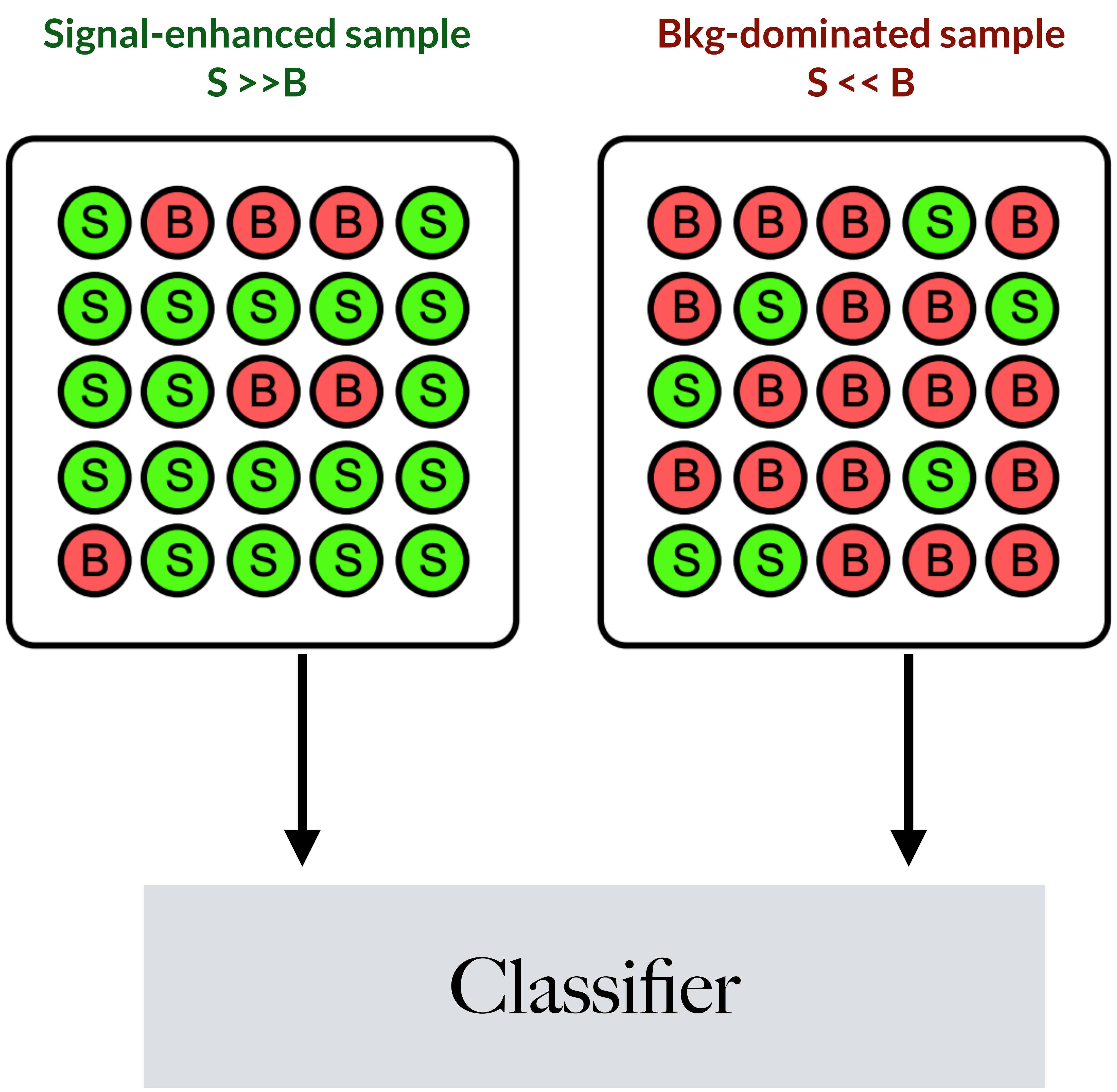
Weakly Supervised Method

Weakly supervised = noisy labels

Let's assume we have:
Two mixed samples of events with no label information

CWoLa (Classification Without Labels) method shows:
It is possible to train a classifier to distinguish red from green

Training on impure samples (different admixture of S and B) is asymptotically equivalent of training on pure samples



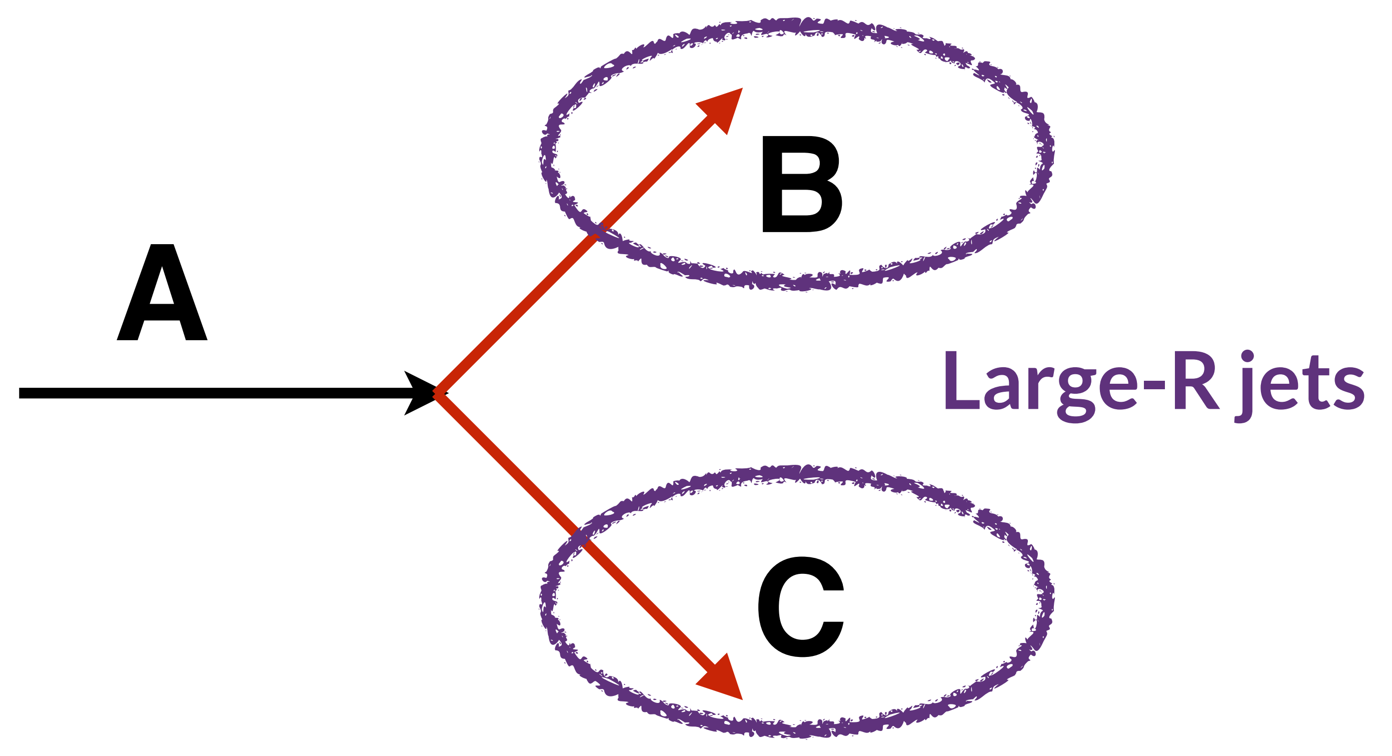
JHEP 10 (2017) 174

Dijet Resonance Search with CWoLa

Phys. Rev. Lett. 125 (2020) 131801

Generic A → B C search

where A, B and C can be BSM particles



CWoLa *method requires* ***two samples*** *that are admixtures of S and B with different fractions of each*

Mixed Samples: Signal Regions and Sidebands

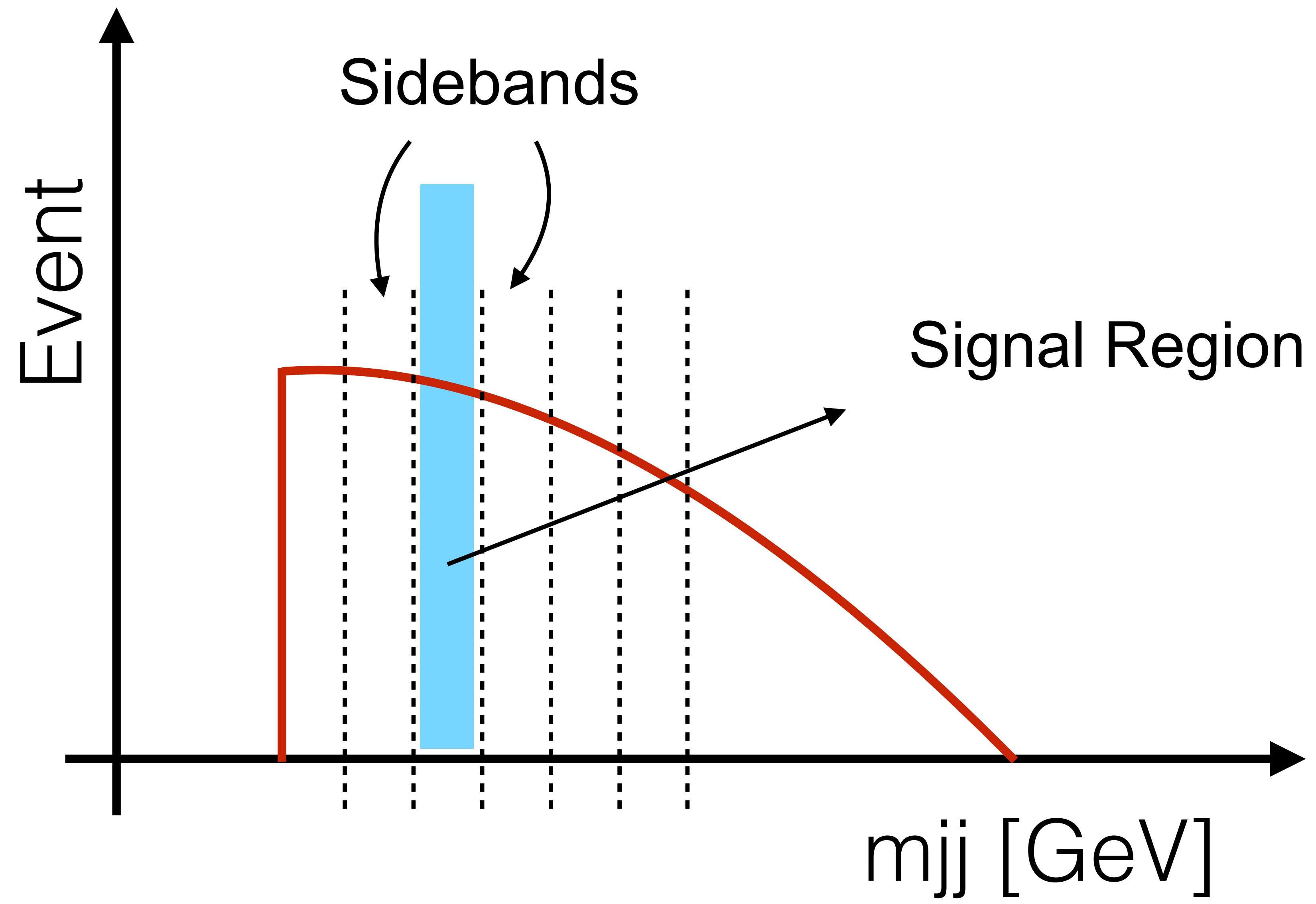
- Signal regions in m_{JJ} with width 20% × m_{JJ}

➡ *Corresponds to the detector resolution for a narrow resonance*

- The signal regions are labeled 0–7 and have boundaries: [1.90, 2.28, 2.74, 3.28, 3.94, 4.73, 5.68, 6.81, 8.17] TeV.

- **Classifier training:** signal vs two sidebands

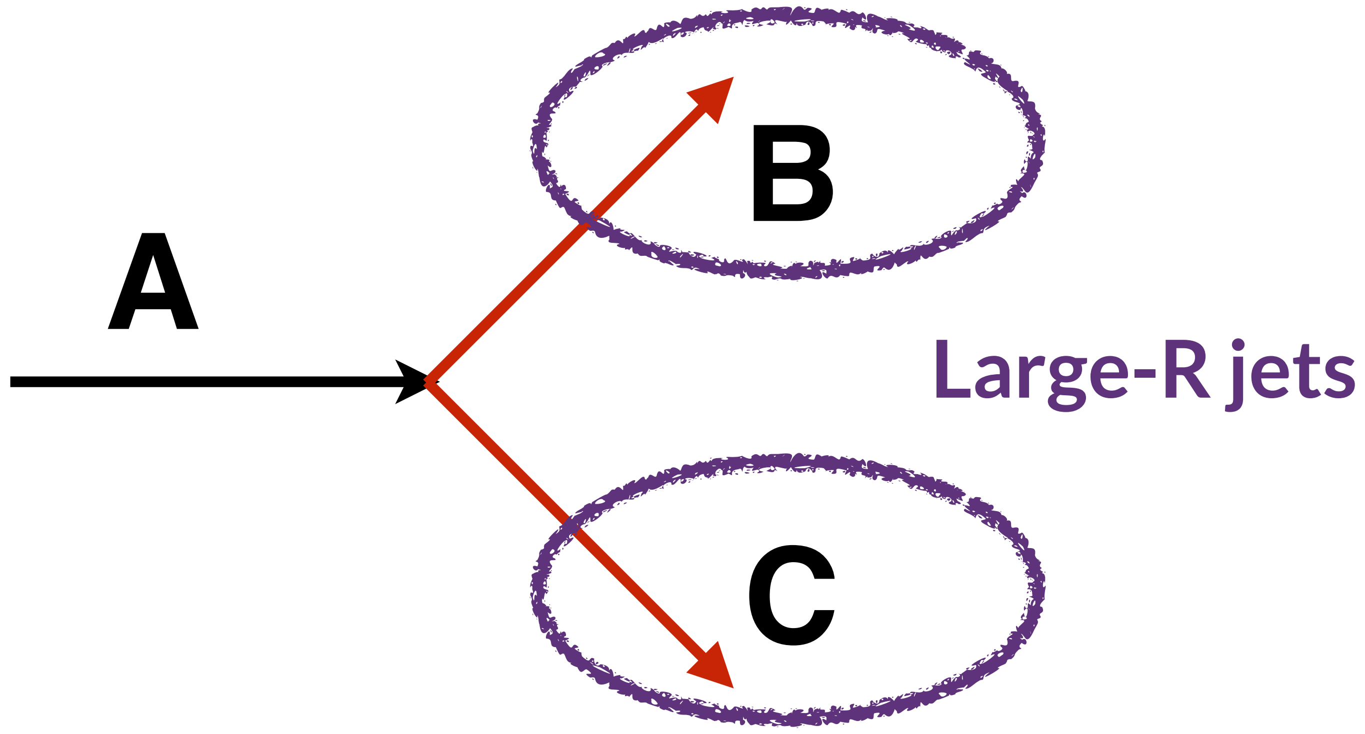
- **Signal Regions:** bins 1-6



Analysis Pre-selection

Phys. Rev. Lett. 125 (2020) 131801

Trigger: Lowest unpre-scaled large-R jet trigger (offline p_T>500 GeV)
- Trimmed large-R jets



Dataset: Full Run 2 (2015-2018) (139 fb^{-1})

Selections:
1. **Two jets** with p_T>200 GeV
   I. leading jet p_T>500 GeV

2. Jet mass > 30 GeV and m < 500 GeV (limit learning differences in tails of distribution)
3. |y1-y2| < 1.2 (rapidity difference)

Mass range: 1.1 - ~8 TeV (fit range: 1.8 – 8.2 TeV)

Mass bins: Assume narrow width, bins set by dijet mass resolution (20% of mJJ)

Neural Network Performance

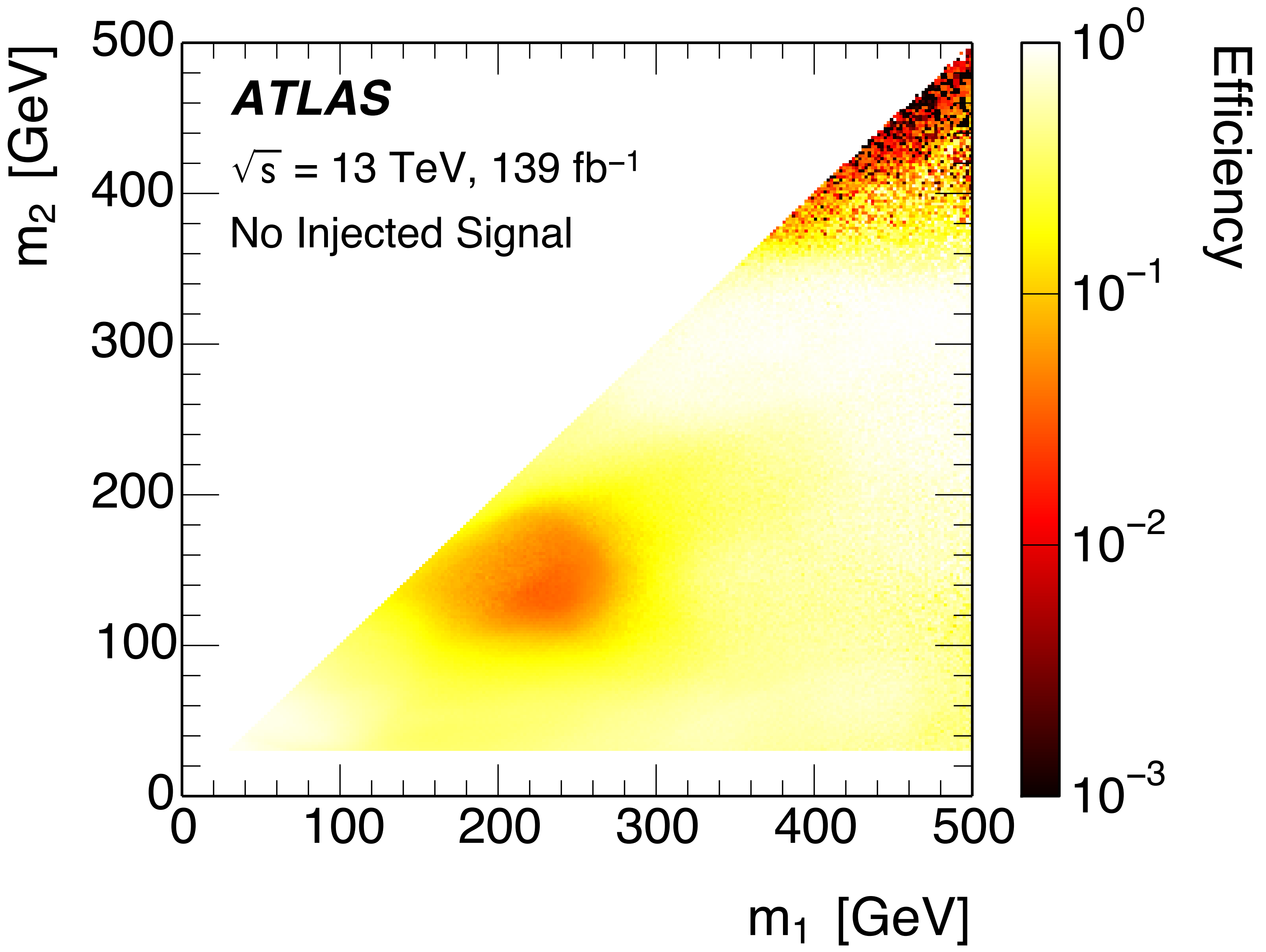
Phys. Rev. Lett. 125 (2020) 131801

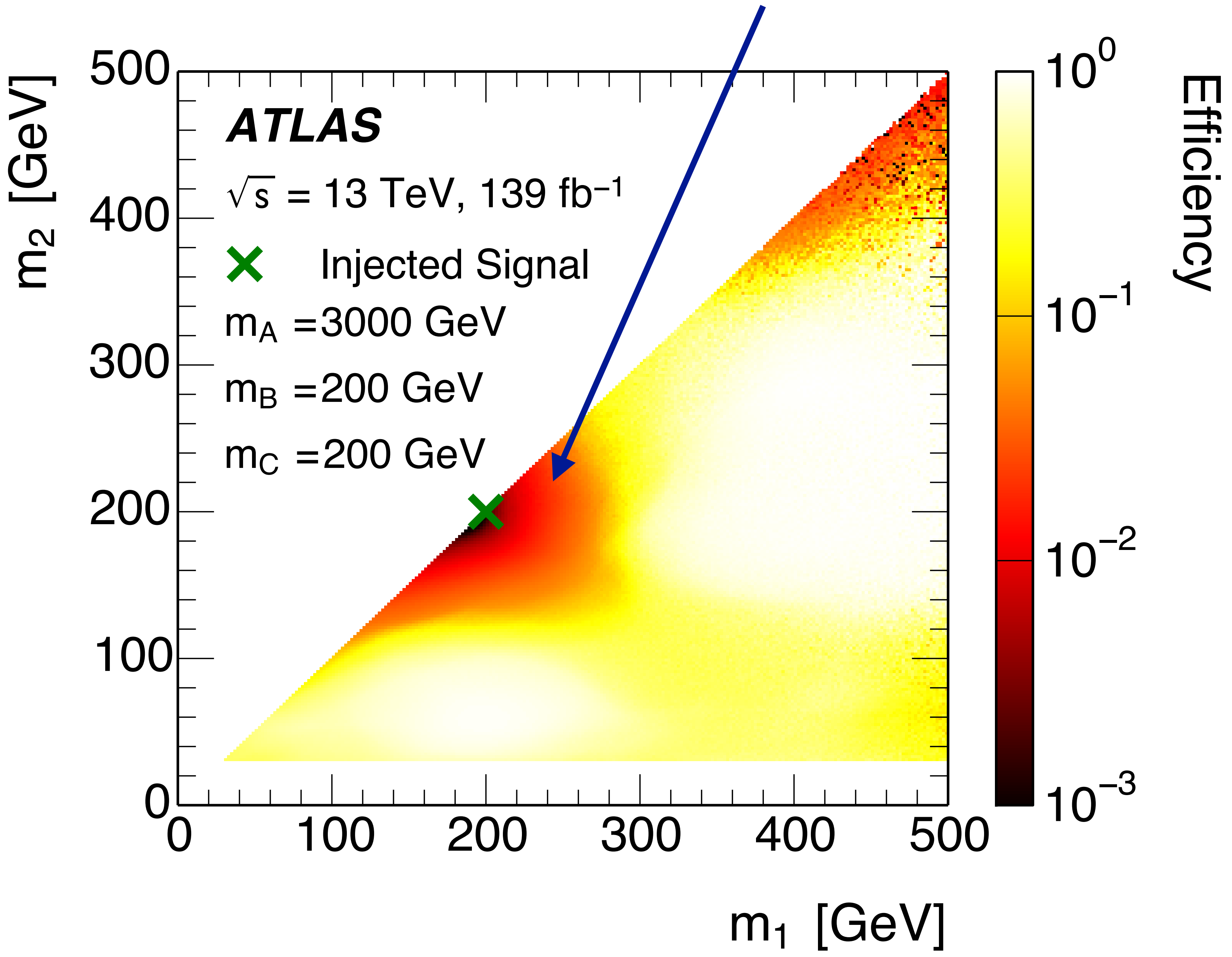
- DNN is trained two jet masses: m1 and m2
- Thresholds applied on NN outputs = Efficiency (ϵ)

Two different selections:

1. $\epsilon = 0.01$ —> *keeps 1% most signal-region-like events*
2. $\epsilon = 0.1$ —> *keeps 10% most signal-region-like events*

The NN score is able to catch the signal if it is present in data





Background fitting

Phys. Rev. Lett. 125 (2020) 131801

Apply classifier score cuts: *enhance signal sensitivity*

Background Estimation: *functional fits*

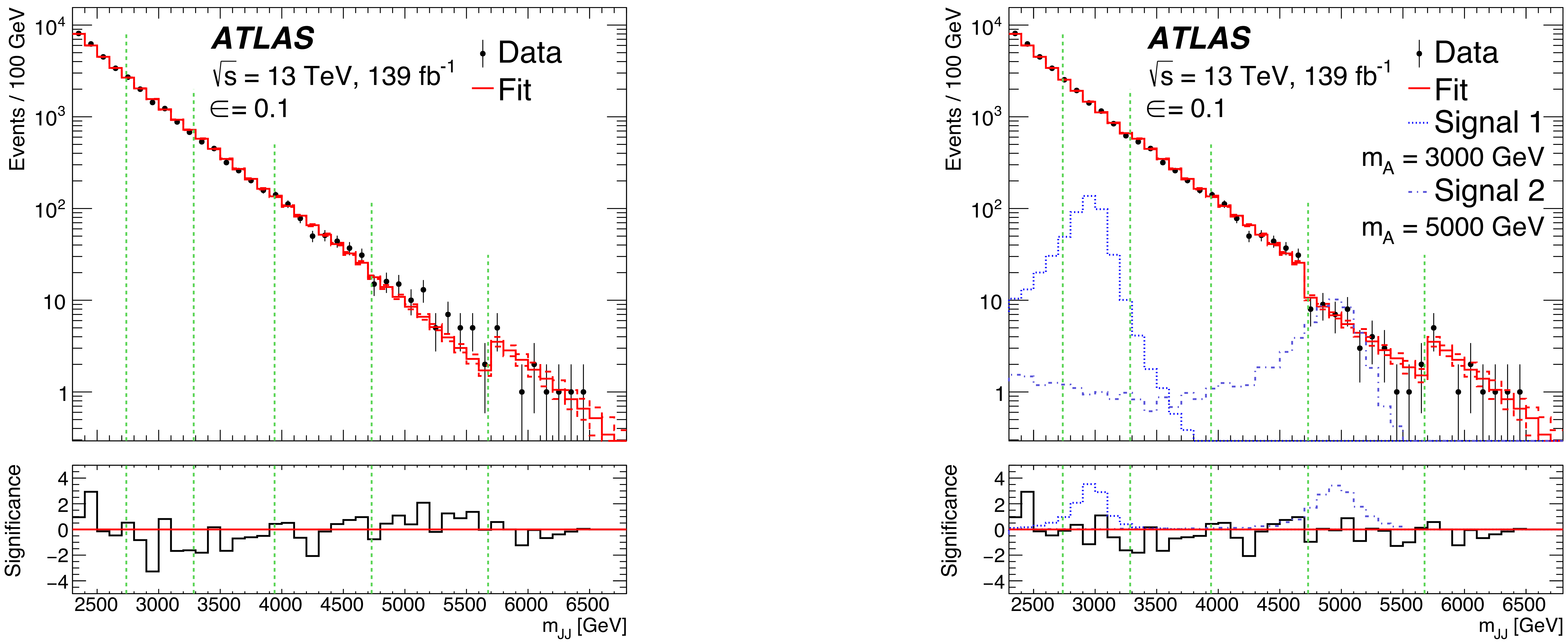
m_{JJ} distributions are fit with a parametric functions in each Signal Region

- **3 Fit functions are used:**

$$\frac{dn}{dx} = p_1(1-x)^{p_2-\xi p_3} x^{-p_3} \quad (1)$$

$$\frac{dn}{dx} = p_1(1-x)^{p_2-\xi_1 p_3} x^{-p_3+(p_4-\xi_2 p_3-\xi_3 p_2)\log(x)} \quad (2)$$

$$\frac{dn}{dx} = p_1 x^{p_2-\xi_1 p_3} e^{-p_3 x+(p_4-\xi_2 p_3-\xi_3 p_2)x^2} \quad (3)$$



Upper Limits on Signal Cross-Section

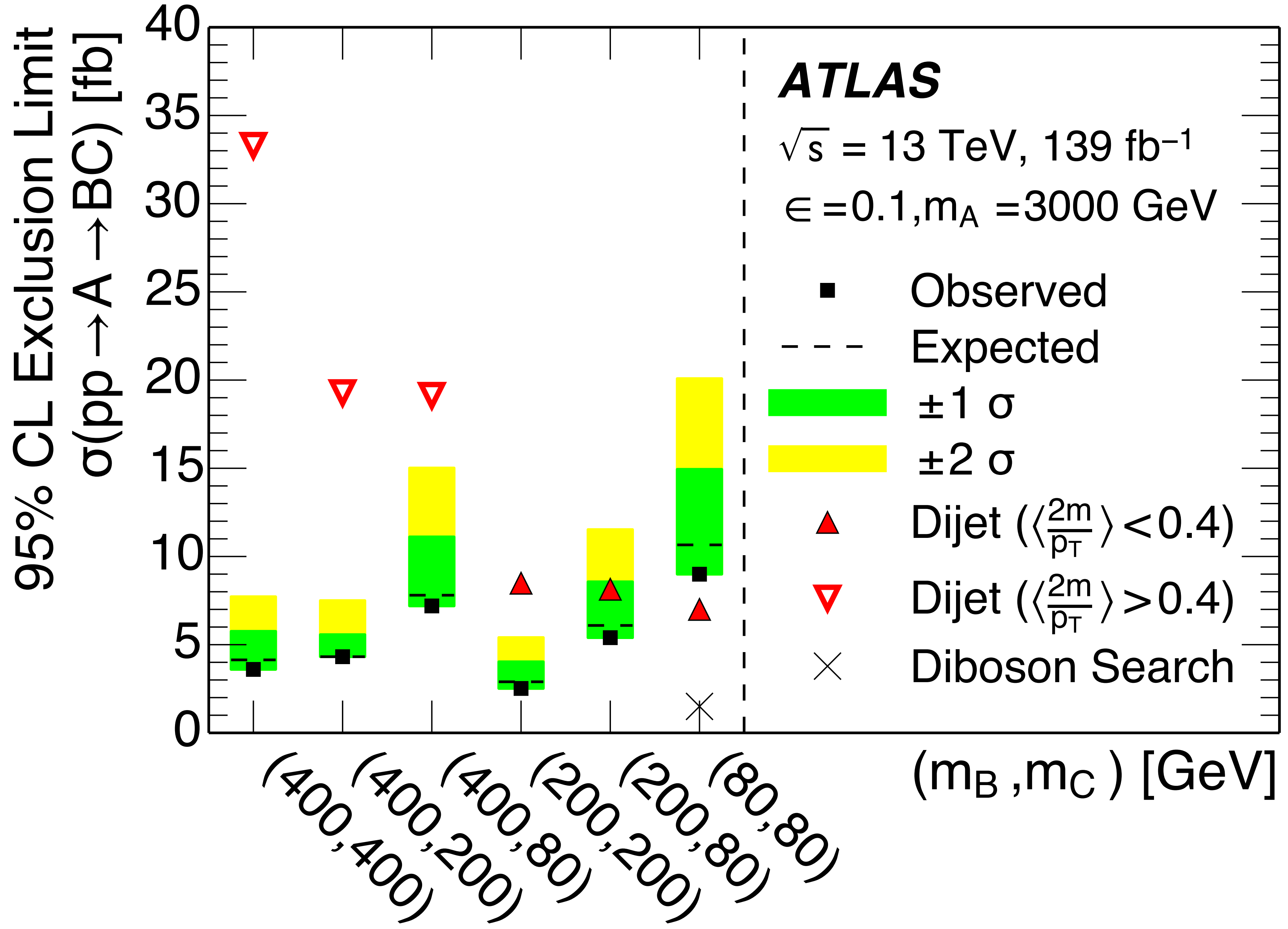
Phys. Rev. Lett. 125 (2020) 131801

The NN is most powerful when the local signal-to-background ratio is high
Weaker limits on some signal models

Comparison with ATLAS diboson and dijet resonance search:

- Dedicated diboson searches show greater sensitivity when mB, mC ~ mW, mZ
- But no sensitivity away from this point

→ the semi-supervised method is sensitive everywhere



$Y \rightarrow X\,H$ Search: Overview and Setup

arXiv:2306.03637

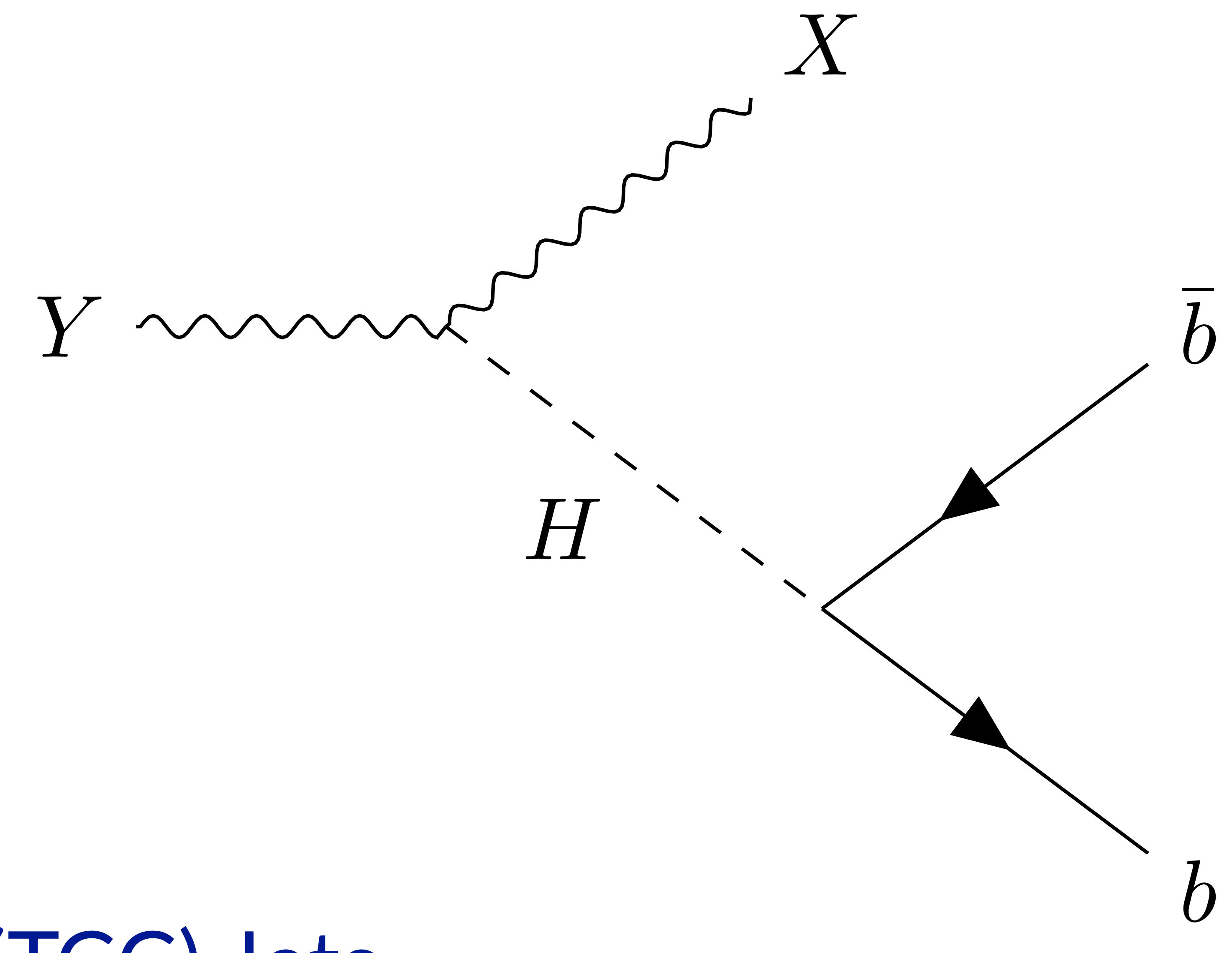
Search for two new particles:

a heavy resonance Y (~ 1TeV) decaying to a SM Higgs (H→bb) and another BSM particle X

- Analysis is sensitive to X masses spanning several orders of magnitude $\mathcal{O}(10)$ GeV to $\mathcal{O}(1)$ TeV

Three signal regions (SR):

1. ***Anomaly SR*** $\rightarrow$ $\mathcal{O}(\sim 100)$ **GeV to** $\mathcal{O}(10)$ **TeV**
2. Two-prong regions (X $\rightarrow qq$): *merged* and *resolved*



- High (~1-6TeV) Y mass resulting in X and H **boosted**
  - Y Reconstructed with two large-R jets →Track CaloCluster (TCC) Jets

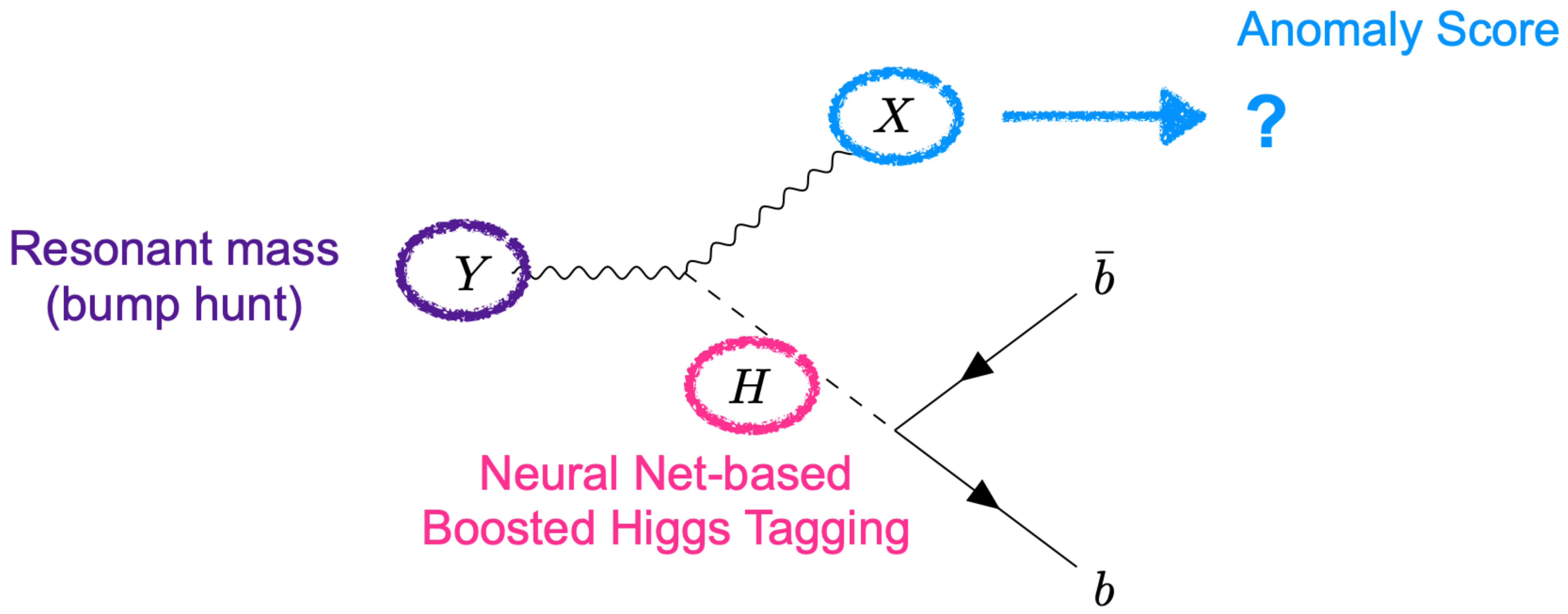
Anomaly Detection in Y → X H Search

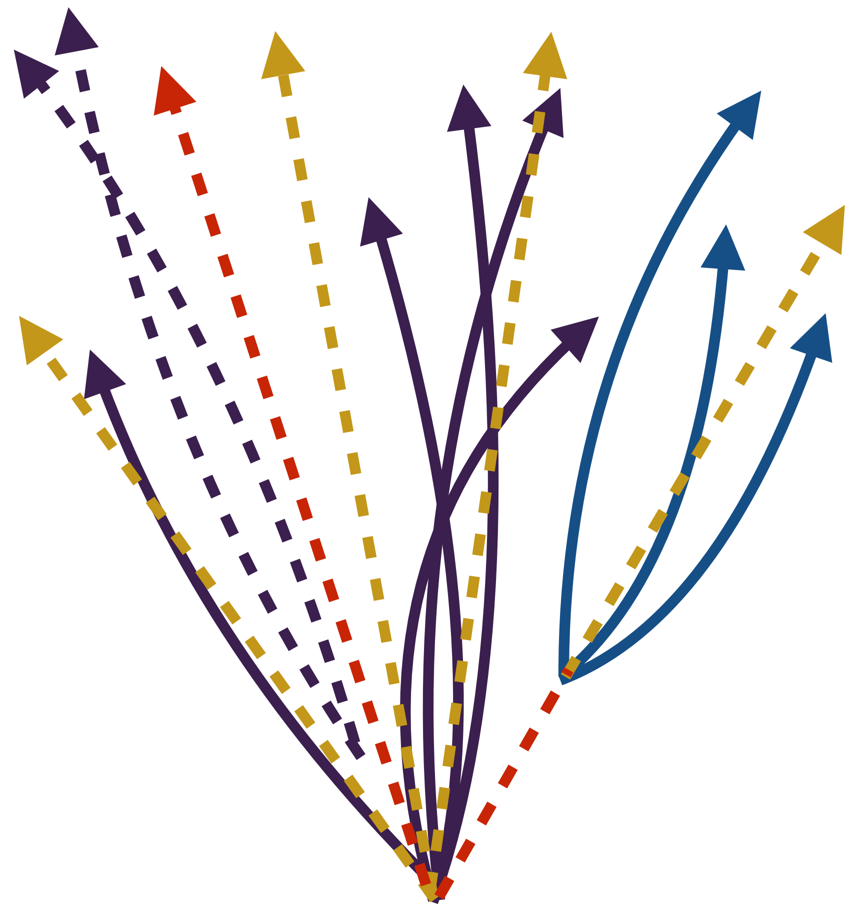
arXiv:2306.03637

- **Unsupervised technique:** Per-jet anomaly score to perform model-independent tagging of anomalous X candidate jets

- Neural net-based tagging of boosted H→ bb topology

- DNN-based (density ratio estimation) reweighting procedure for data-driven background estimation





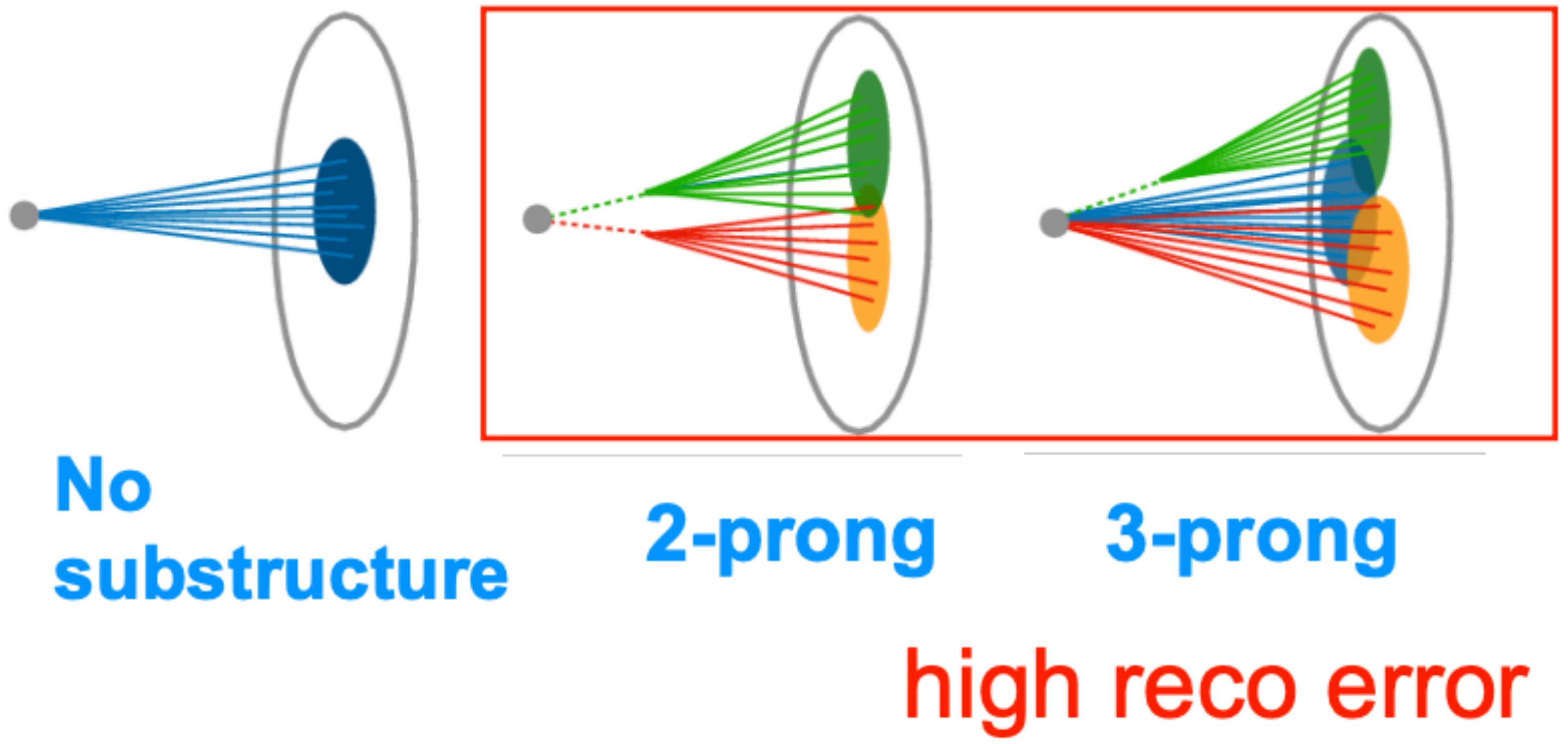
Anomalous Jet Tagging

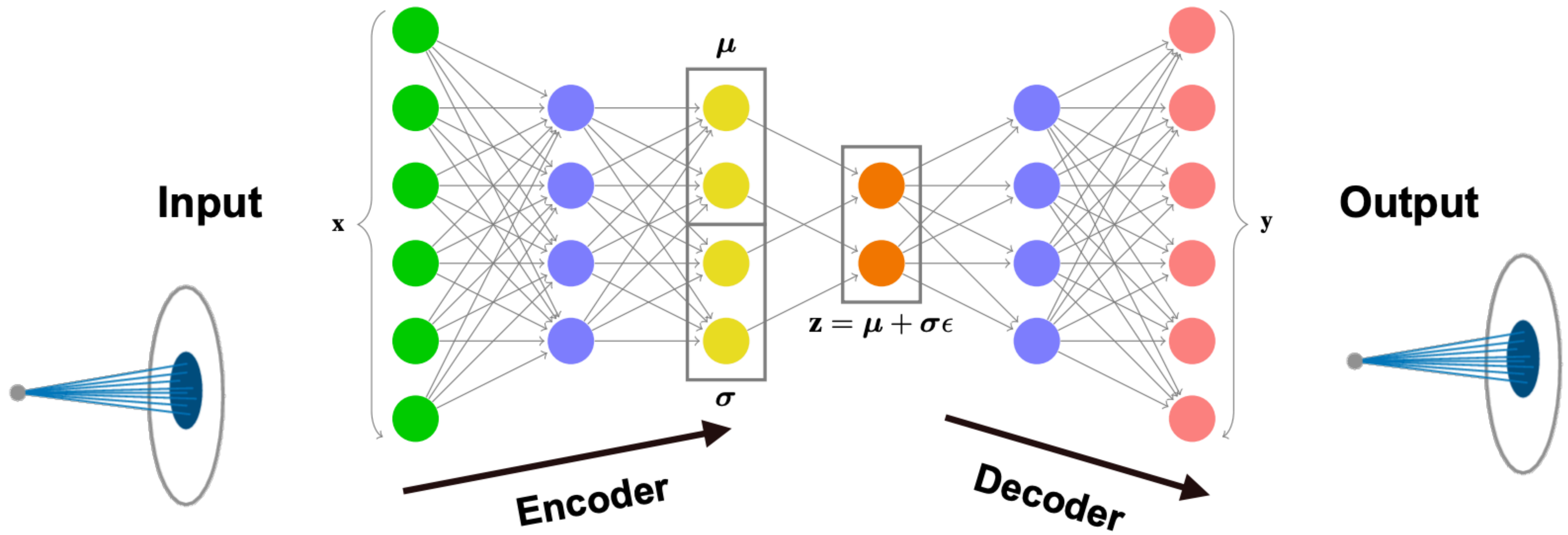
Jet-level anomaly score given by a **variational recurrent neural network (VRNN)**

arXiv:2306.03637

- Unsupervised training over jets in data modeled as sequence of kt-ordered constituent 4-vectors: no signal model!

- Define anomaly score (AS) per jet as a function of VRNN loss





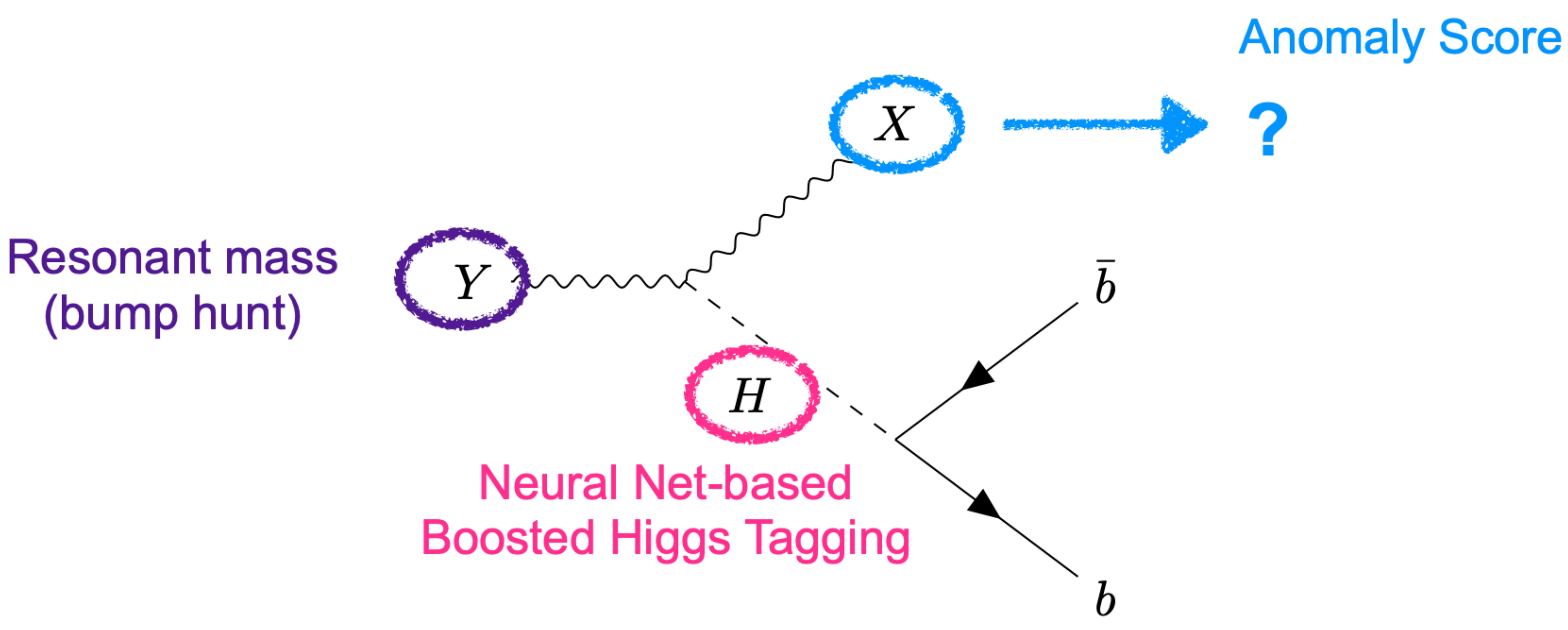
arxiv:1506.02216

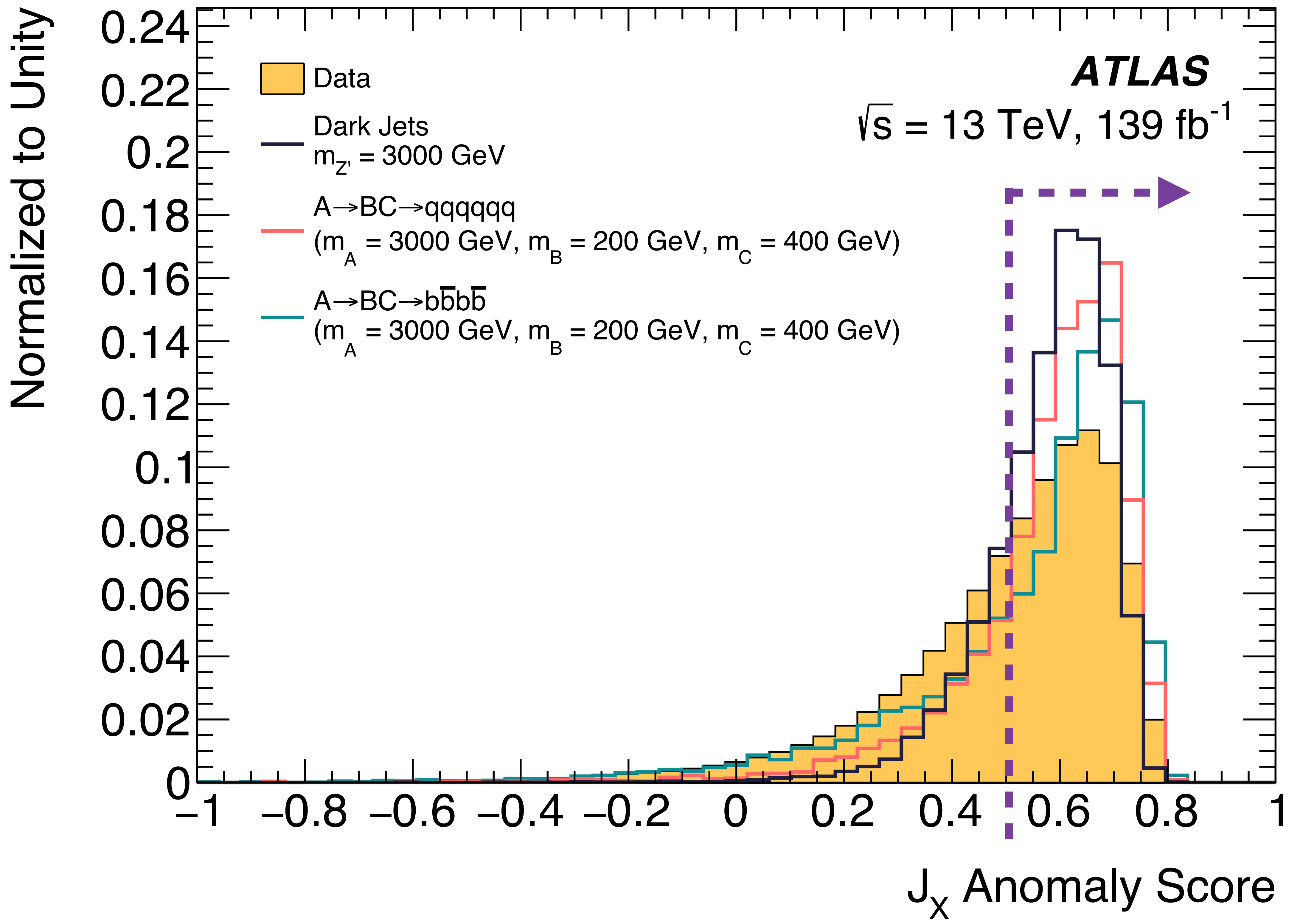
$Y \rightarrow X\ H$ Search: Anomalous X Tagging

arXiv:2306.03637

- Trained using jets from full Run-2 dataset

- Test model-independence by studying AS discrimination performance on 4 jet topologies: 2-prong, 3-prong, heavy flavor (displaced vertices), and dark jets (pattern of missing and visible energy)

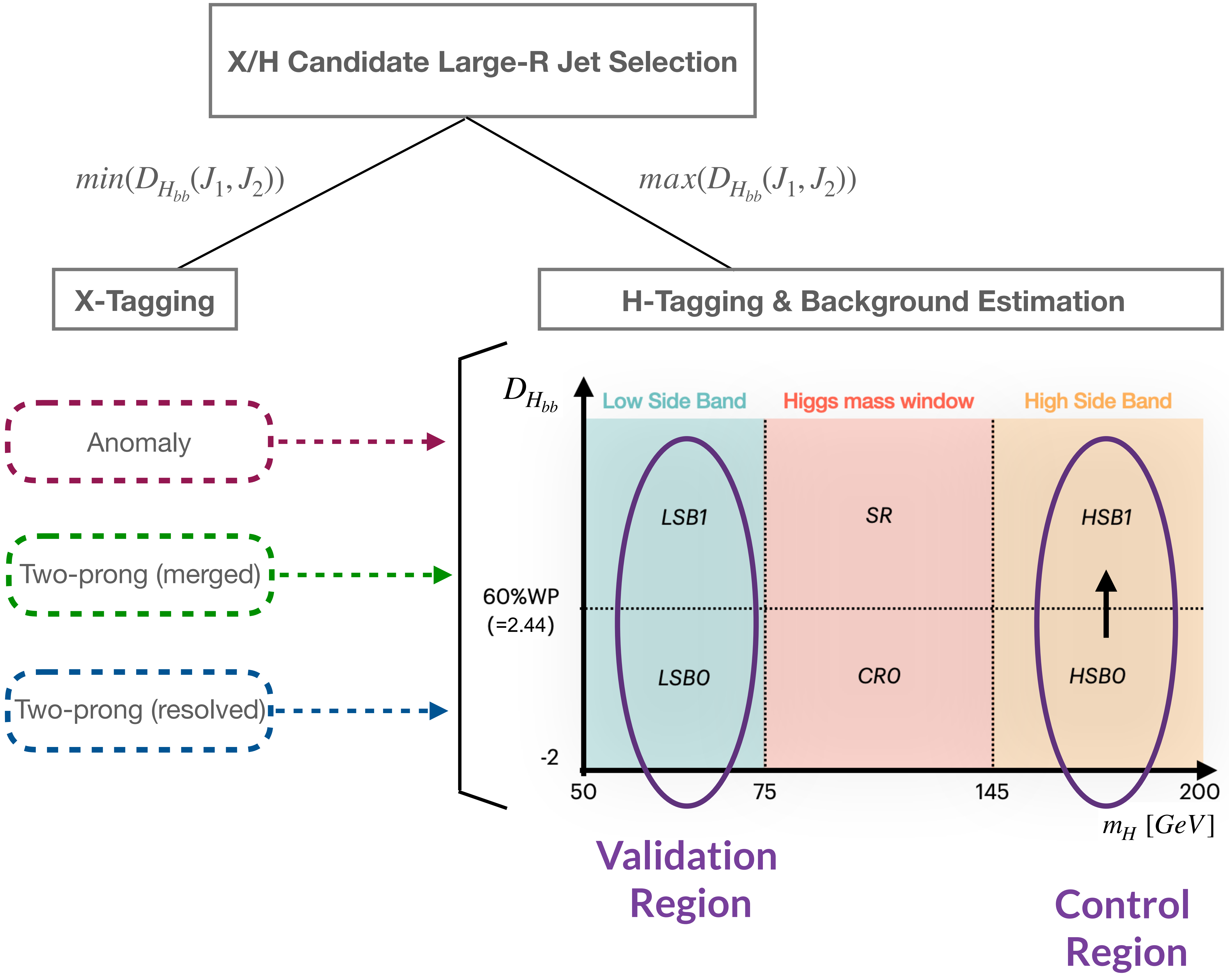




$Y \to X\,H$ Search: Analysis Flow

arXiv:2306.03637

1. **Large-R jet trigger:** J1(pT) > 500 GeV and mJJ > 1.3 TeV
2. **Ambiguity resolution:** jet with highest D_{Hbb} score is Higgs candidate
3. **X-tagging:** AS of X candidate > 0.5 (model independent search) or 2-prong regions (model dependent search)
4. **Higgs tagging:** DHbb of H candidate > 2.44



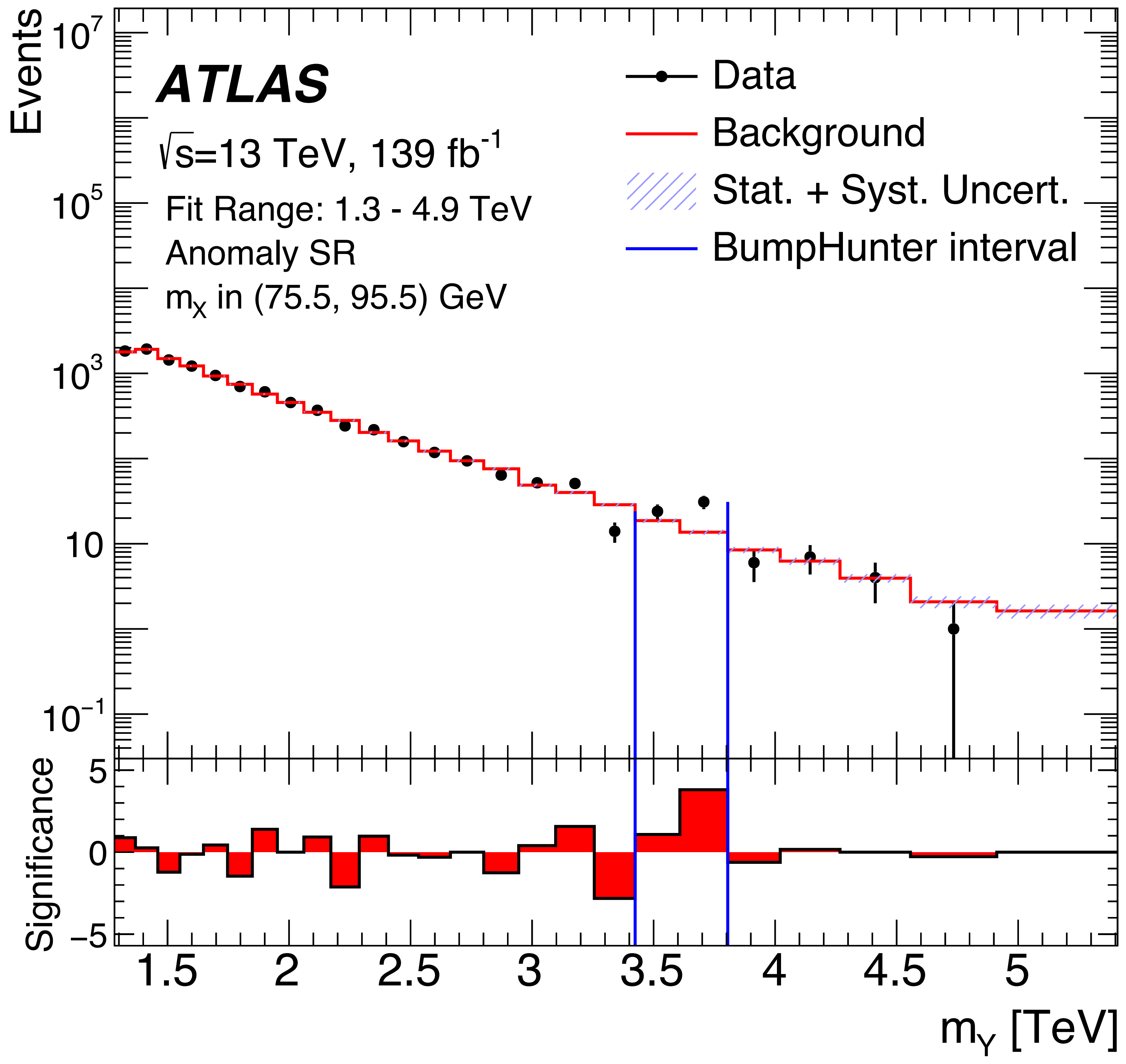
- **SR selection:** 75 < mH < 145 GeV
- **Background estimation:** DNN-derived reweights for untagged high sideband (HSB0→HSB1)
- **Validation:** low sideband (LSB)

Y → X H Search: Search Results

arXiv:2306.03637

- Calculated p-values across all m_Y and m_X bins in the anomaly signal region

- The lowest observed p-value corresponds to the bin with $mY \in [3608, 3805]$ GeV and $mX \in [75.5, 95.5]$ GeV
  - Compatibility p -value of < 0.001
  - 1.4σ global significance in BumpHunter



Limits from Signal Injection

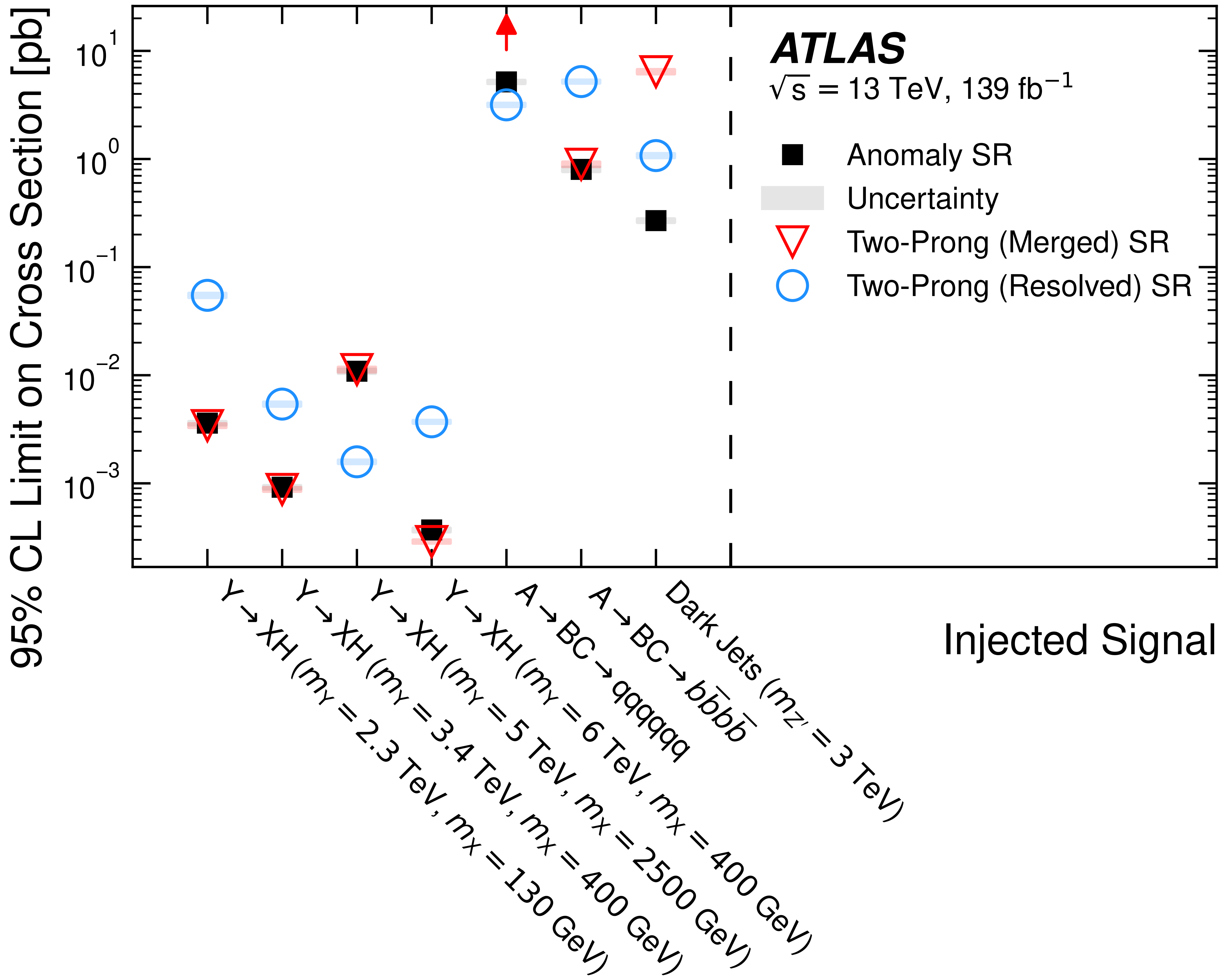
arXiv:2306.03637

Upper limits obtained by injecting benchmark signals

When X particle is highly boosted the upper limits are approximately the same across the merged and anomaly SRs

Signal models with alternative jet substructure has higher sensitivity in the Anomaly SR

Improvements from Anomaly SR:
Dark jet limits are ~10 times better with respect to 2-prong tagging approach

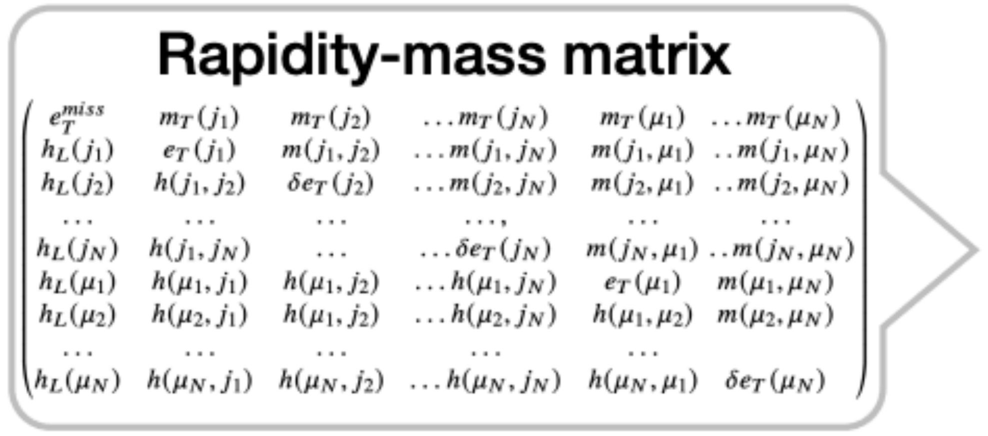


➡ First application of fully unsupervised machine learning to an ATLAS analysis

Anomaly detection in Jet + X final state

- Signal regions are defined based on Autoencoder score
- 9 mass spectrum is analyzed in each SR

arxiv:2307.01612



Trigger (one lepton) and pre-selection ($p_T^l > 60$ GeV, $p_T^{jet} > 30$ GeV)

↓

Reconstruct Rapidity Mass Matrix for each event

↓

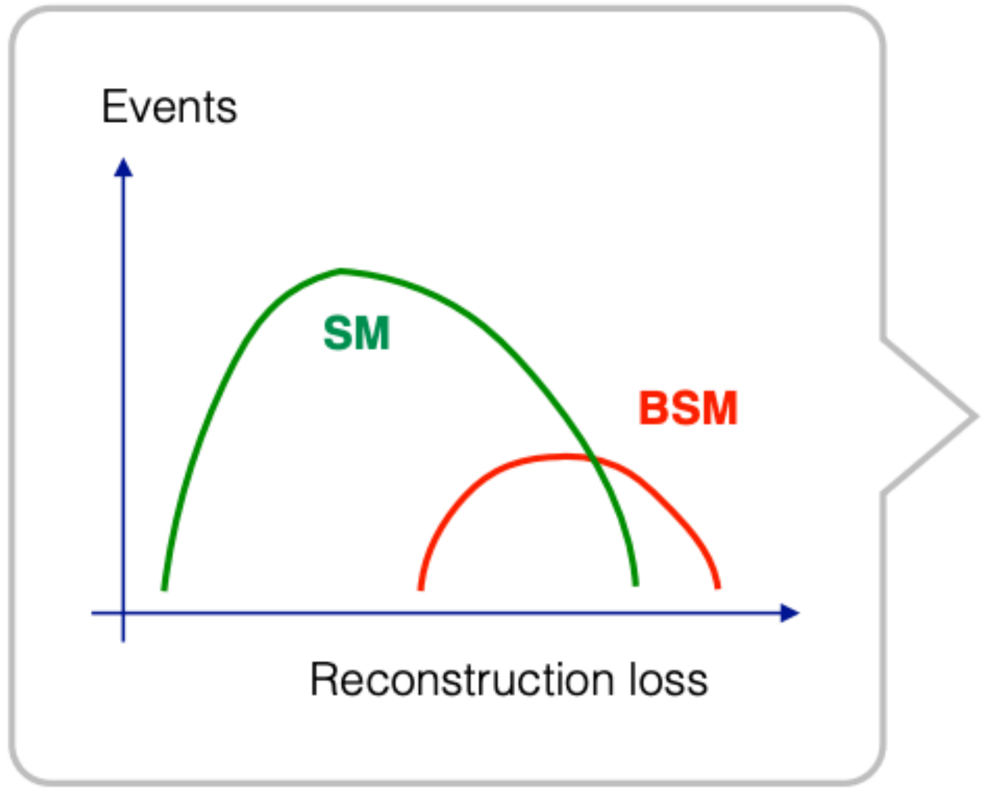
Train autoencoder using 1% ATLAS Run2 data

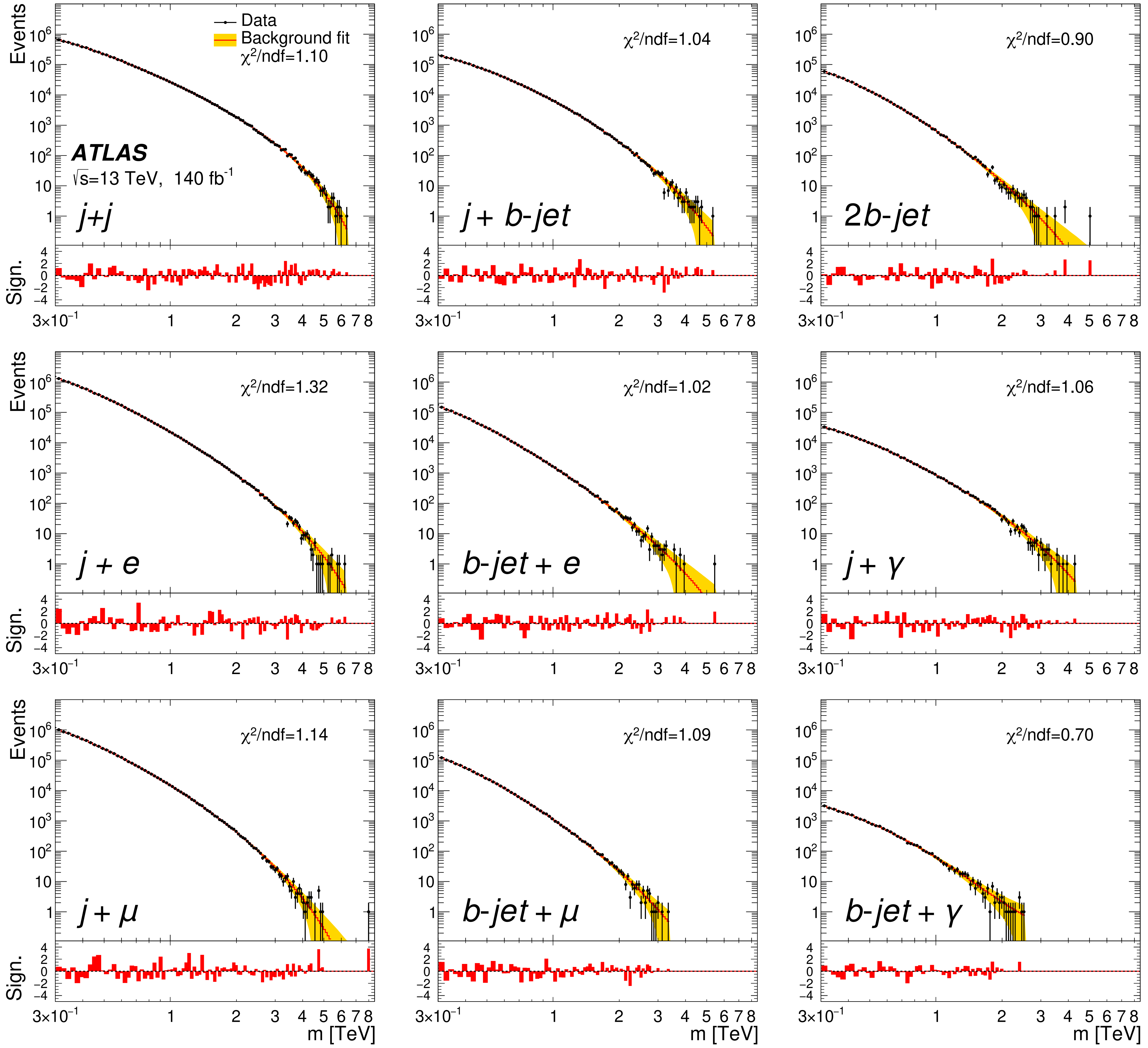
↓

Define signal region using reconstruction loss from autoencoder

↓

Fit invariant mass spectrum, statistical analysis, look for bumps



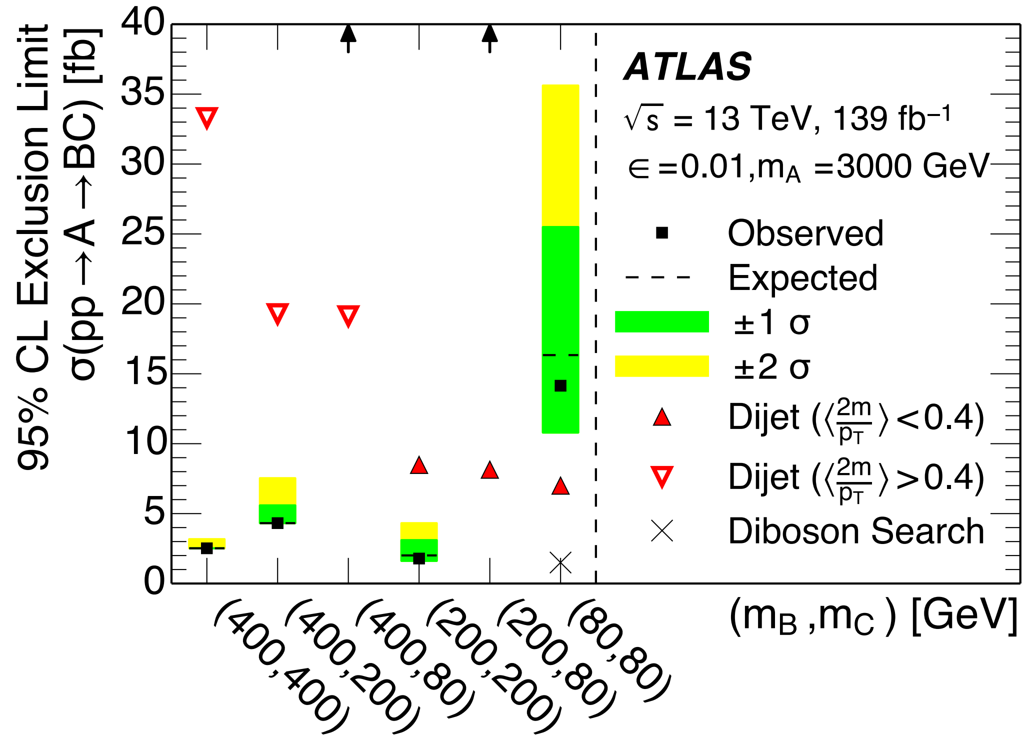


You will hear more from Wasikul's talk

Summary and Outlook

Successful application of semi-supervised and unsupervised Anomaly detection methods

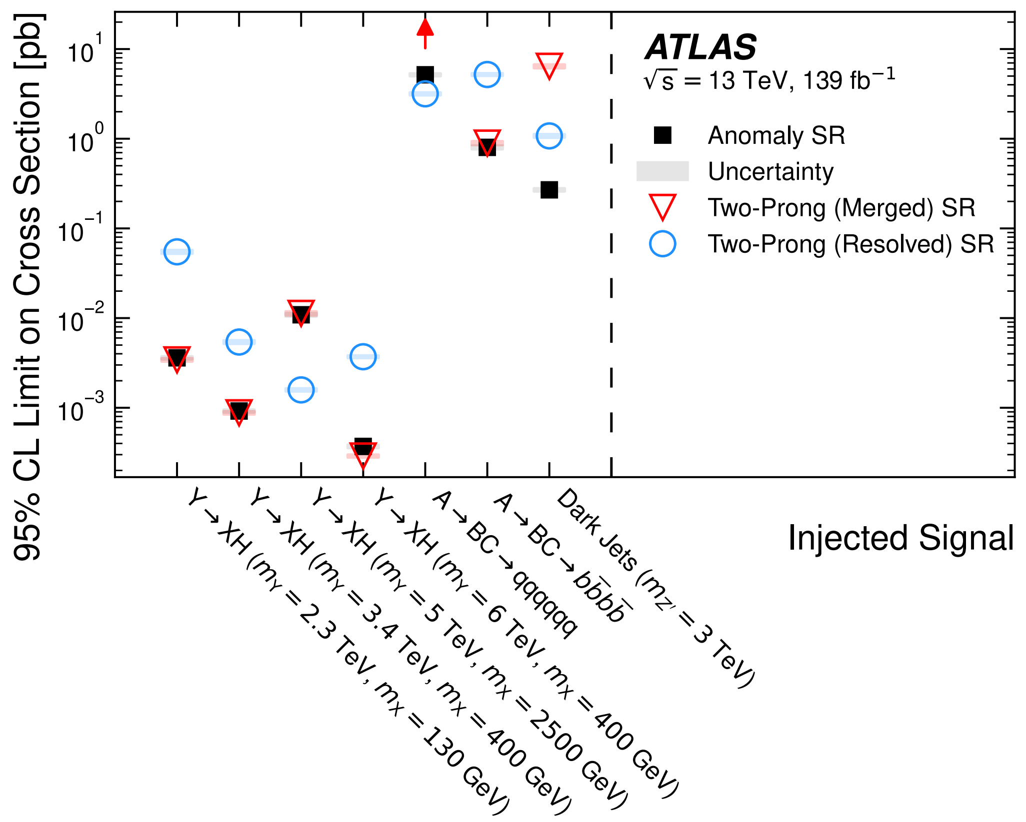
Three published results so far on Full Run-2 data

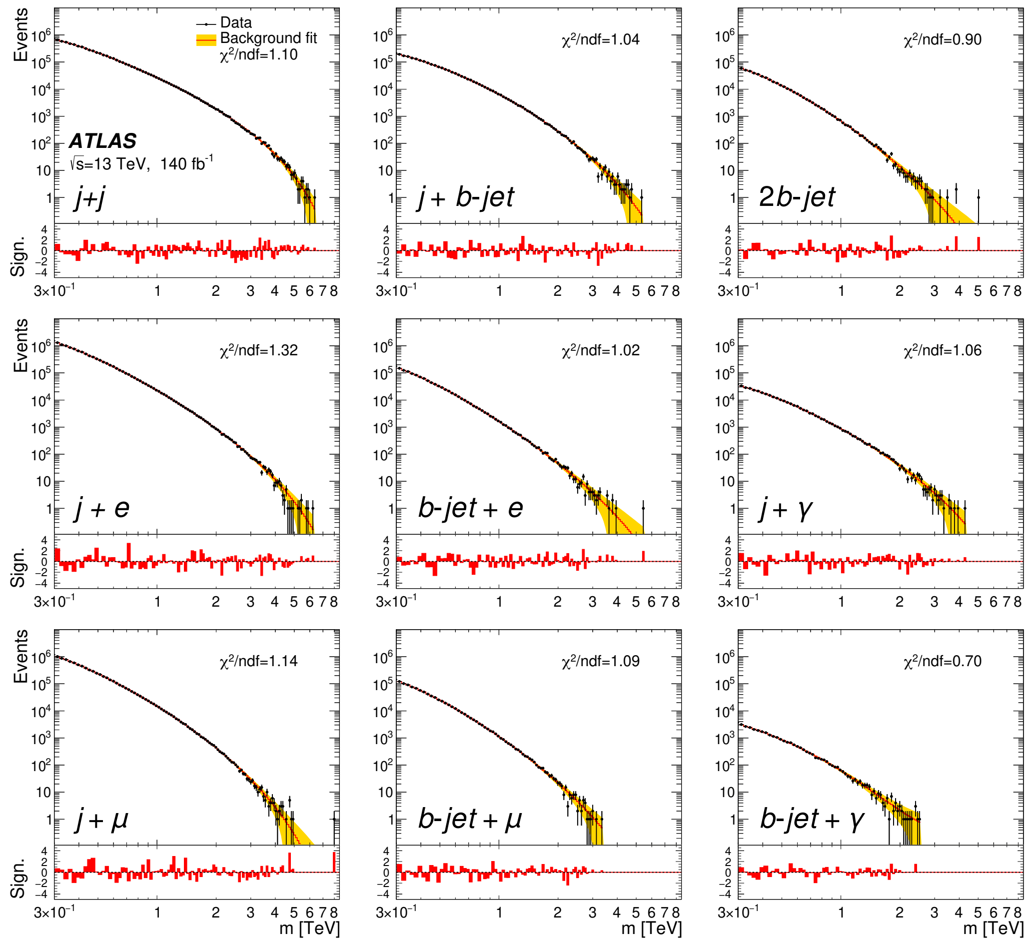


Anomaly Detection searches are ramping up in ATLAS

➡ Efficient way of doing model independent search across different final states

➡ **Stay Tuned!** Many more searches are coming up

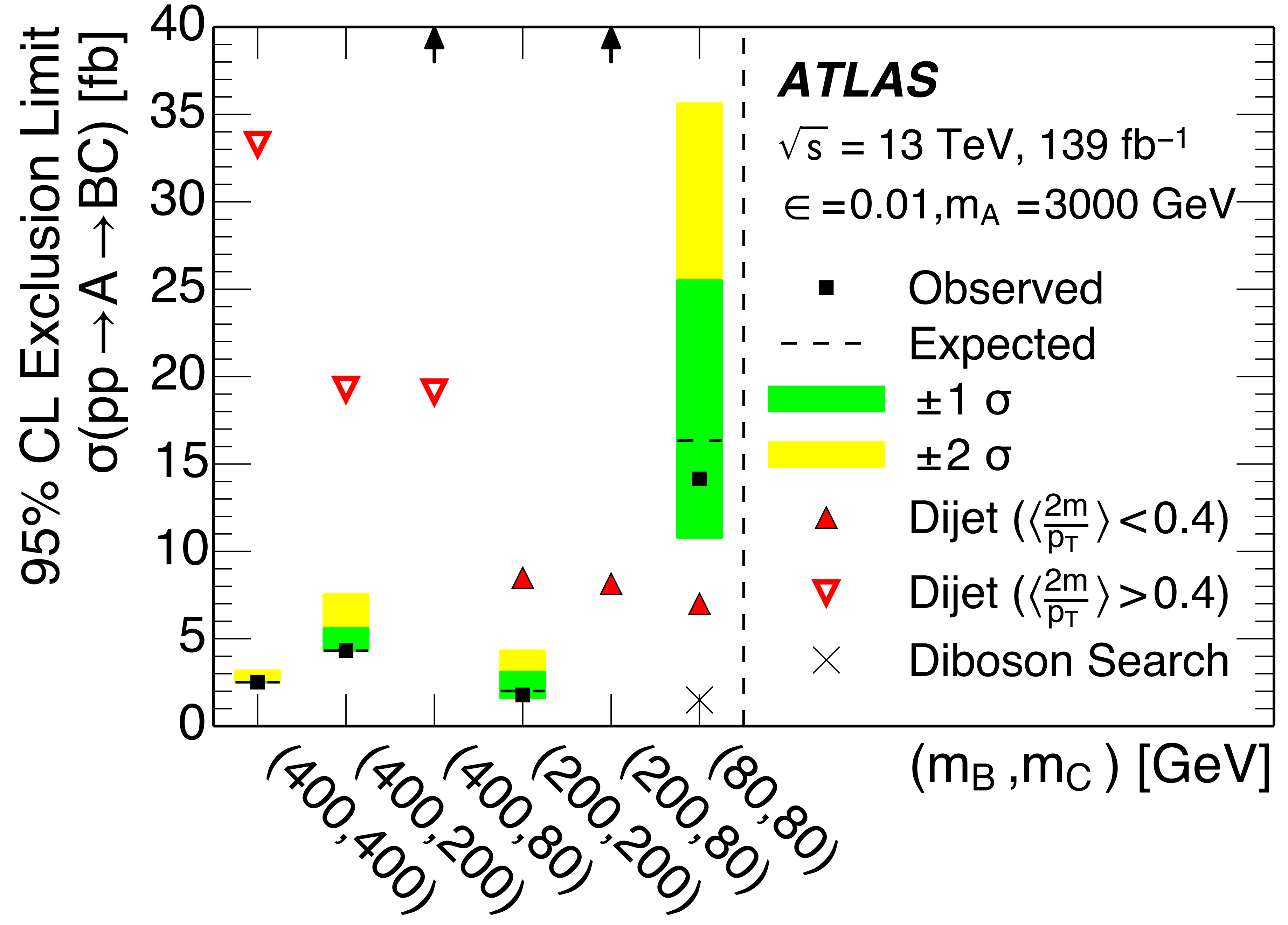




Thank You!

Extra Slides

Dijet CWoLa: Limit



Y -> XH: Event Selection

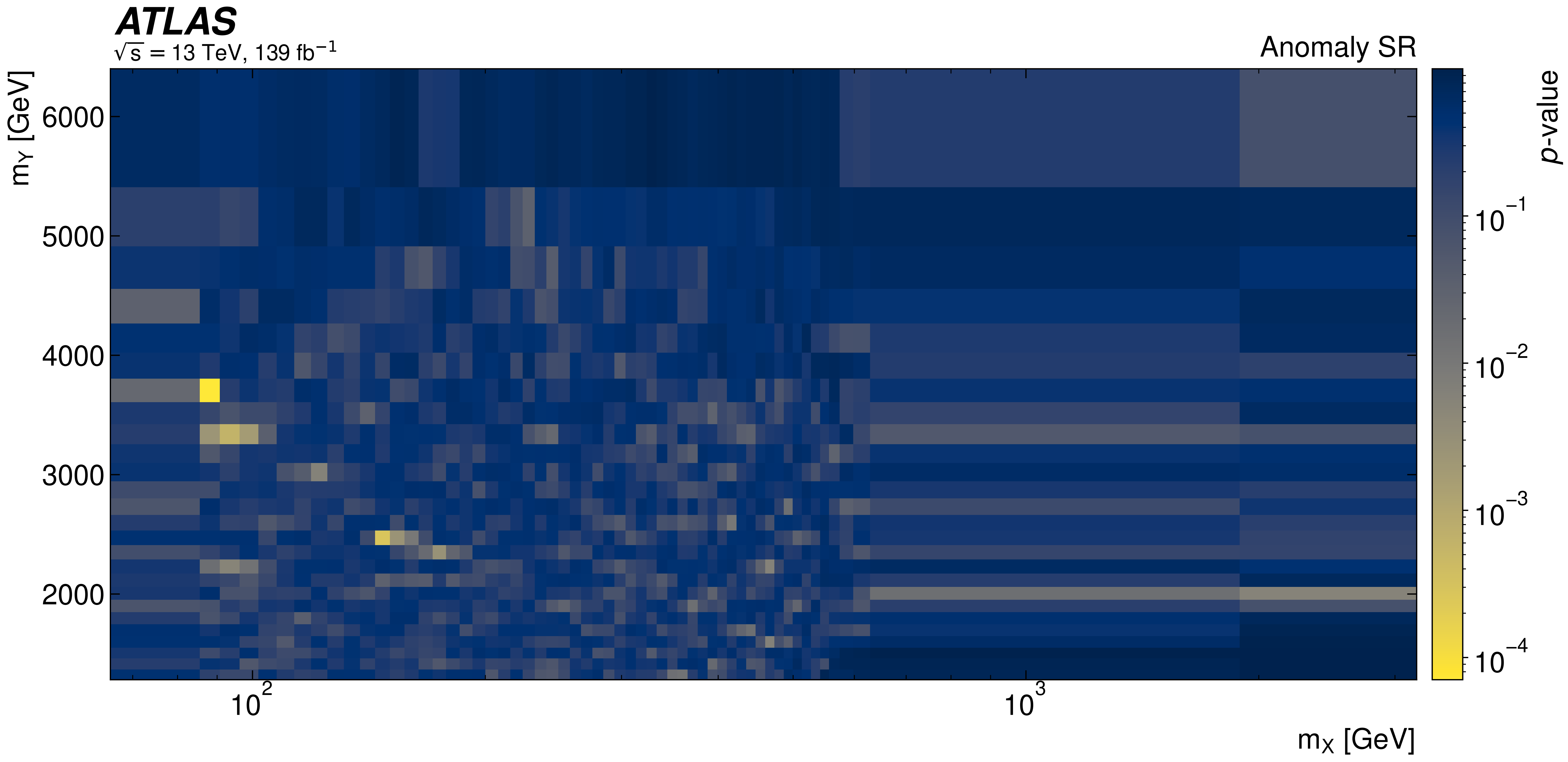
Parameter	Preselection requirements				
m_{JJ} [GeV]	> 1300				
$p_{\mathrm{T}}(J_1)$ [GeV]	> 500				
m_J [GeV]	$m_{J_1} > 50 \parallel m_{J_2} > 50$				
$D_{H_{bb}}$	> -2				
	Signal regions				
	Merged		Resolved		Anomaly
m_H [GeV]	(75, 145)				
$D_{H_{bb}}$	> 2.44				
D_2^{trk}	< 1.2		> 1.2		-
$\lvert\Delta y_{j_1,j_2}\rvert$	-		< 2.5		-
$p_{\mathrm{T}}^{\mathrm{bal}}$	-		< 0.8		-
Anomaly Score (S_{A})	-		-		> 0.5
	Background estimation regions				
	CR0	HSB0	HSB1	LSB0	LSB1
m_H [GeV]	(75, 145)	(145, 200)		(65, 75)	
$D_{H_{bb}}$	< 2.44	< 2.44	> 2.44	< 2.44	> 2.44

Hbb Score

$$D_{H_{bb}} = \ln \frac{P_{\text{Higgs}}}{f_{\text{top}} \cdot P_{\text{top}} + (1 - f_{\text{top}}) \cdot P_{\text{multijet}}}.$$

Y -> XH: 2D p-value

The distribution of observed p-values across all mY and mX bins in the anomaly signal region, comparing data with the background estimates generated by a background-only fit, displayed in the two-dimensional (mX, mY) grid. The p-value calculation is performed at the center of each mX bin, and all statistical and background systematic uncertainties are considered. The lowest observed p-value corresponds to the bin with mY within [3608, 3805] GeV and mX within [75.5, 95.5] GeV.



$Y \rightarrow X\,H$ Results

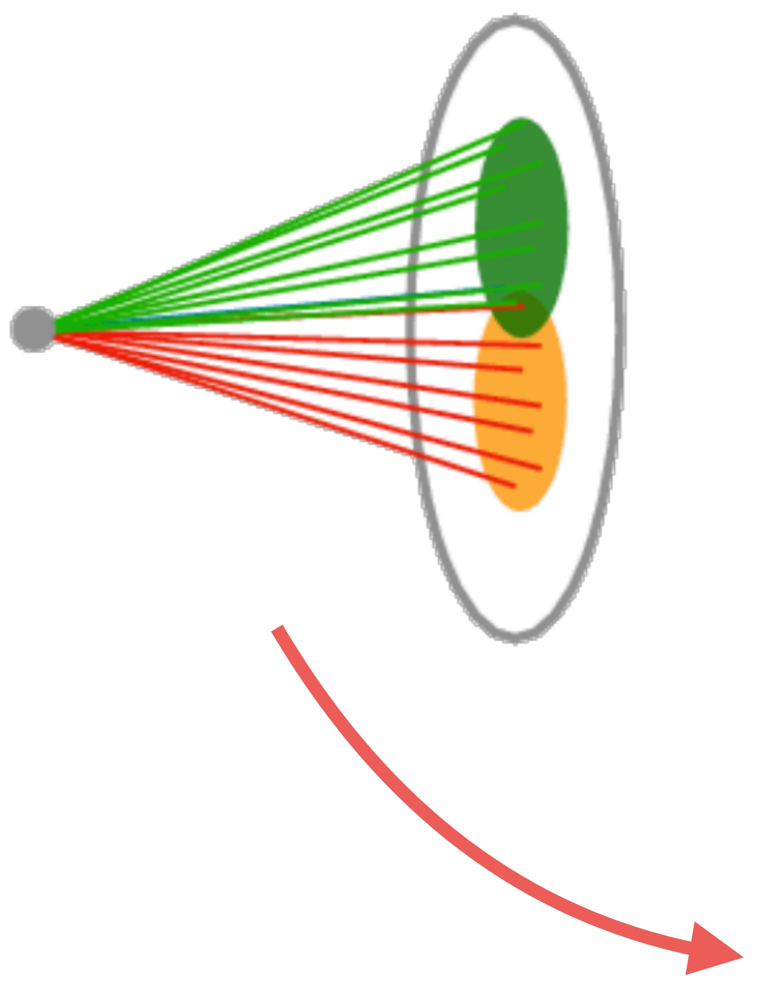
arXiv:2306.03637

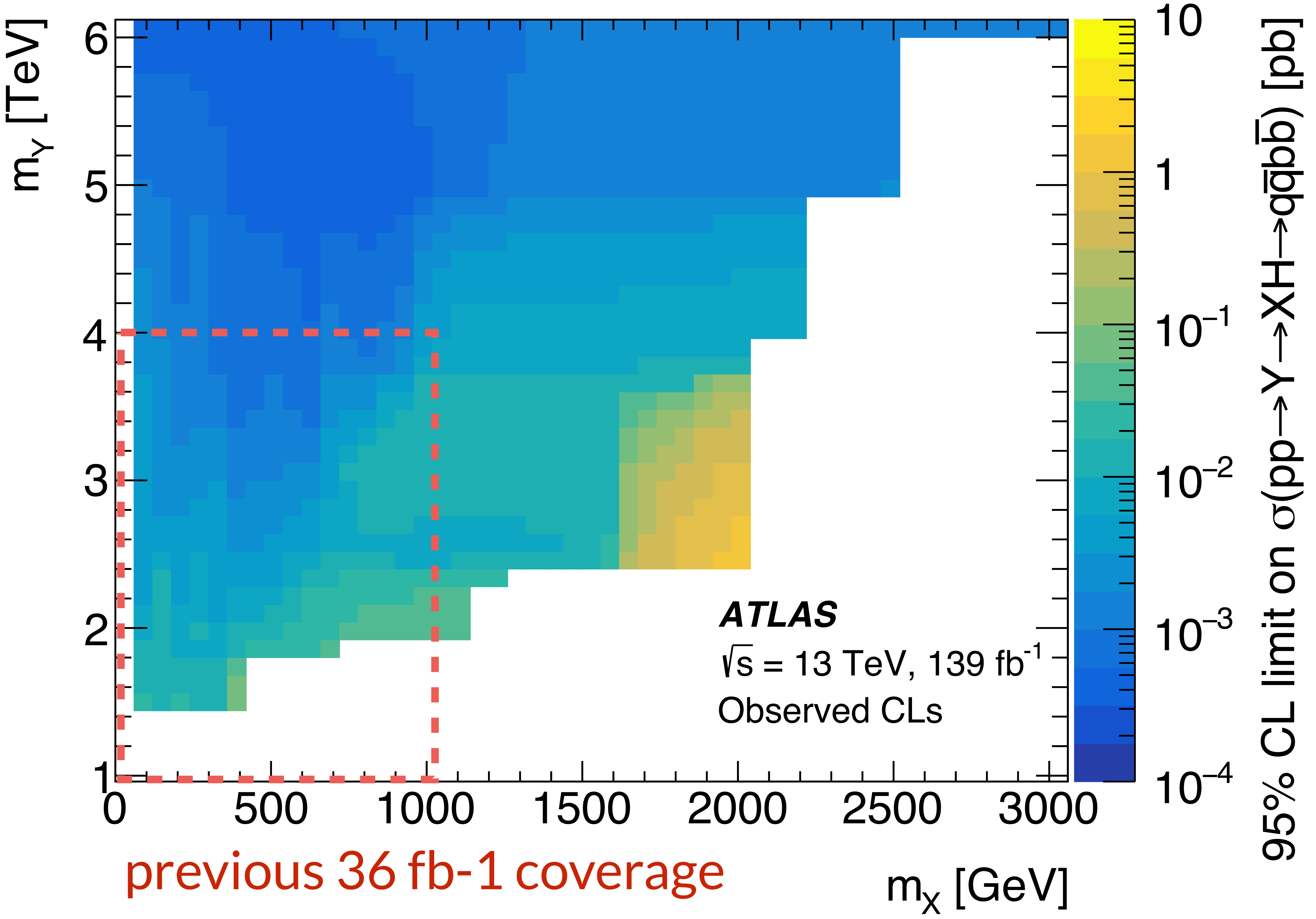
Fit invariant mass of X and H for excesses in overlapping windows of m_X

Results:

no significant deviations in anomaly region across mX bins

Interpret in nominal X→qq, sensitive up to 6 TeV resonance mass



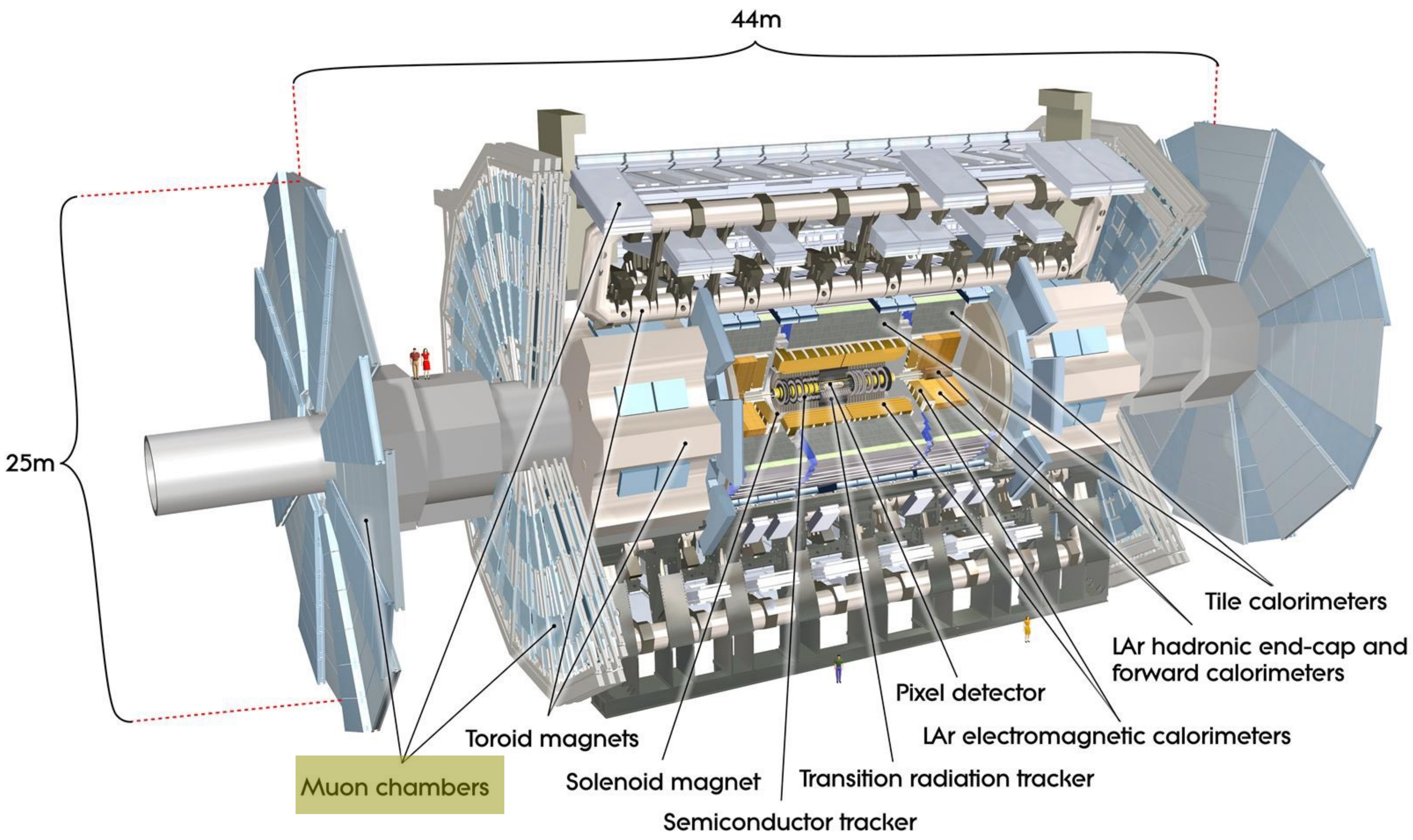


The ATLAS Experiment

General purpose detector

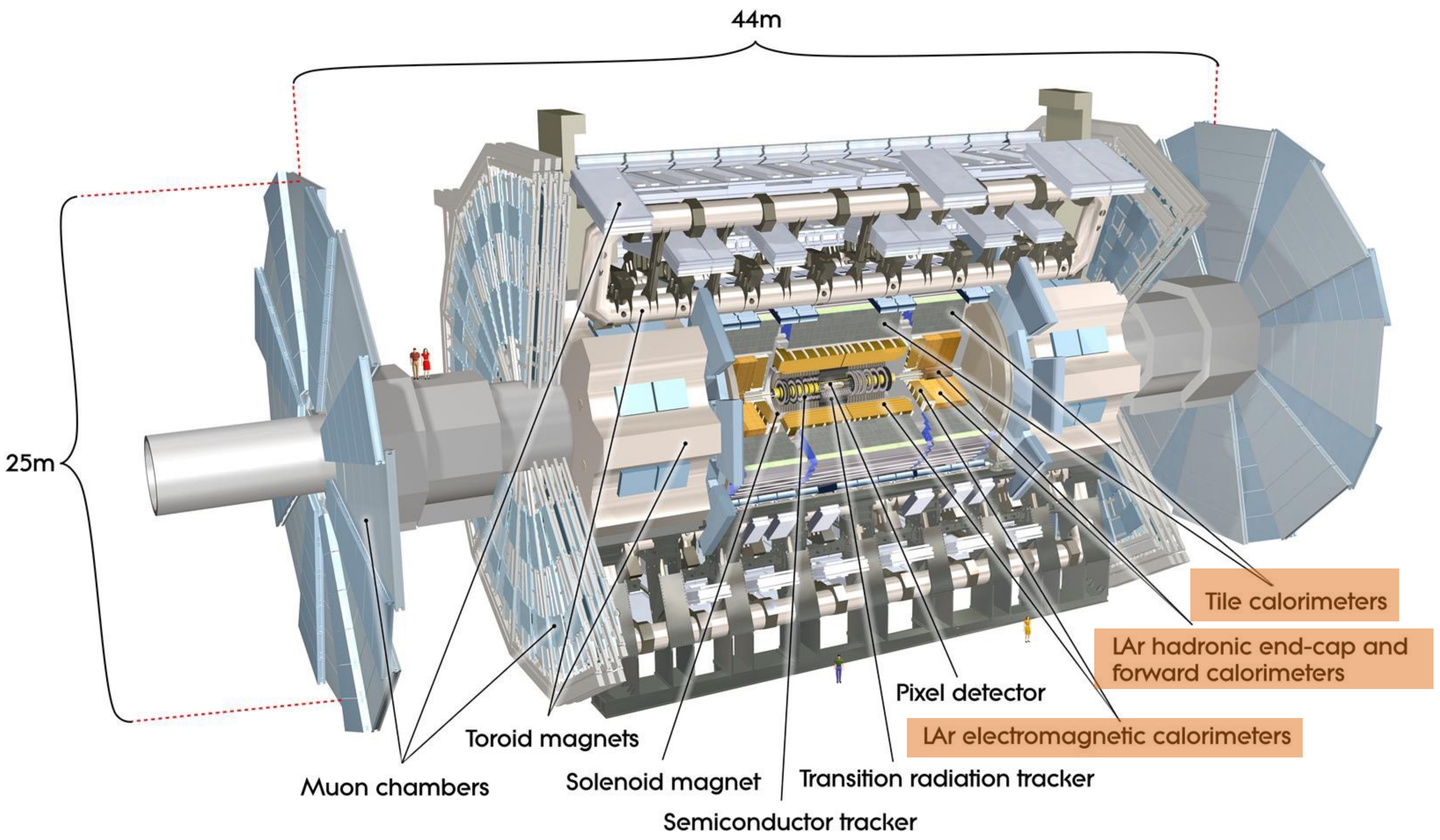
Toroidal Magnet: 0.5 T

Muon Spectrometer:
Four different detector technology



The ATLAS Experiment

General purpose detector



Calorimeter:

Electromagnetic (Liquid Argon), Hadronic (Liquid Argon (endcap) & Tile (barrel))

Solenoid Magnet: 2.0 T

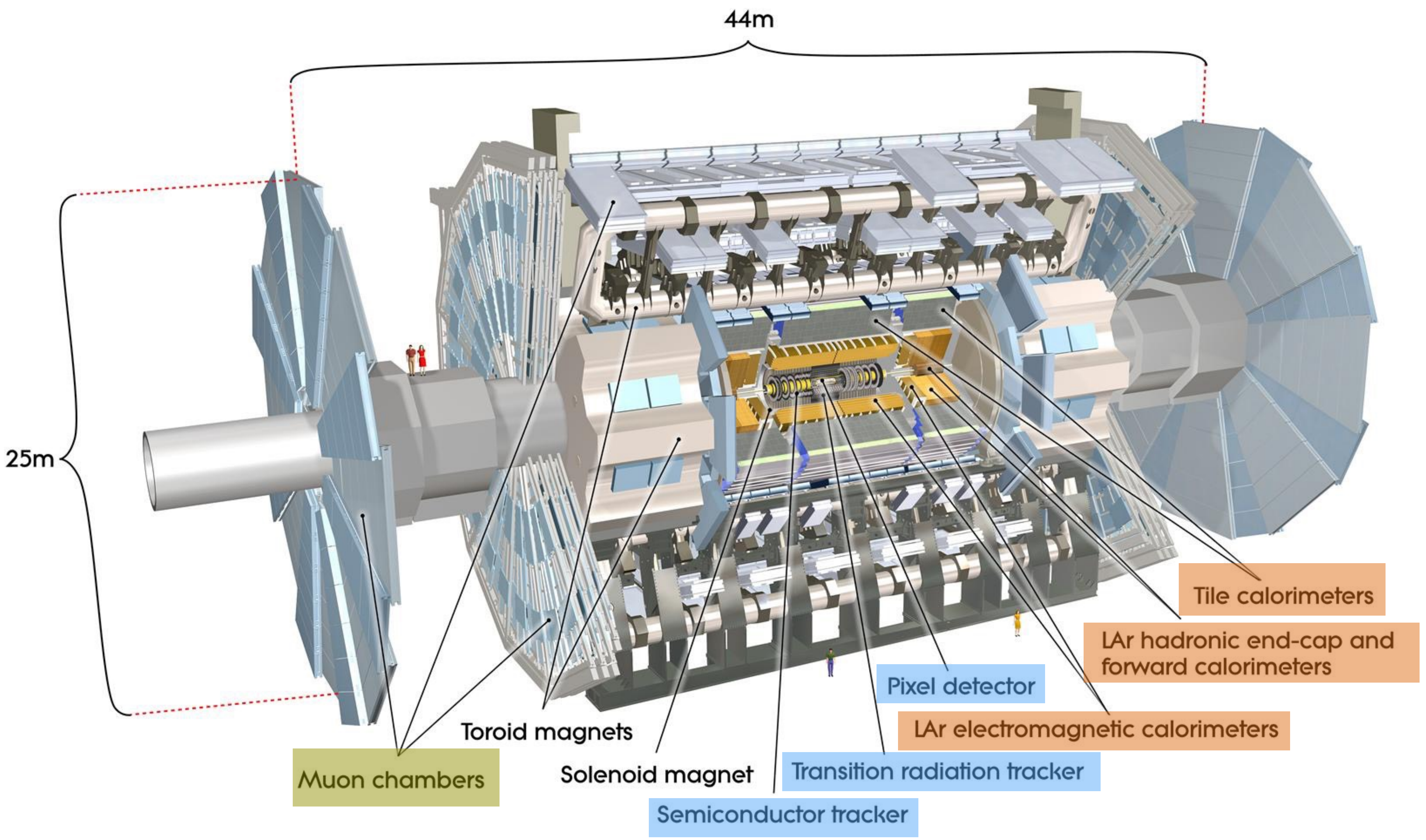
The ATLAS Experiment

General purpose detector

Muon Spectrometer:
Four different detector technology

Calorimeter:
Electromagnetic (Liquid Argon), Hadronic (Liquid Argon (endcap) & Tile (barrel))

Solenoid Magnet: 2.0 T



Inner Detector:
Three different detector technology

1. Silicon Pixel
2. Silicon Strip
3. Straw Tubes: Transition Radiation Tracker (TRT)