Anomaly detection at ATLAS

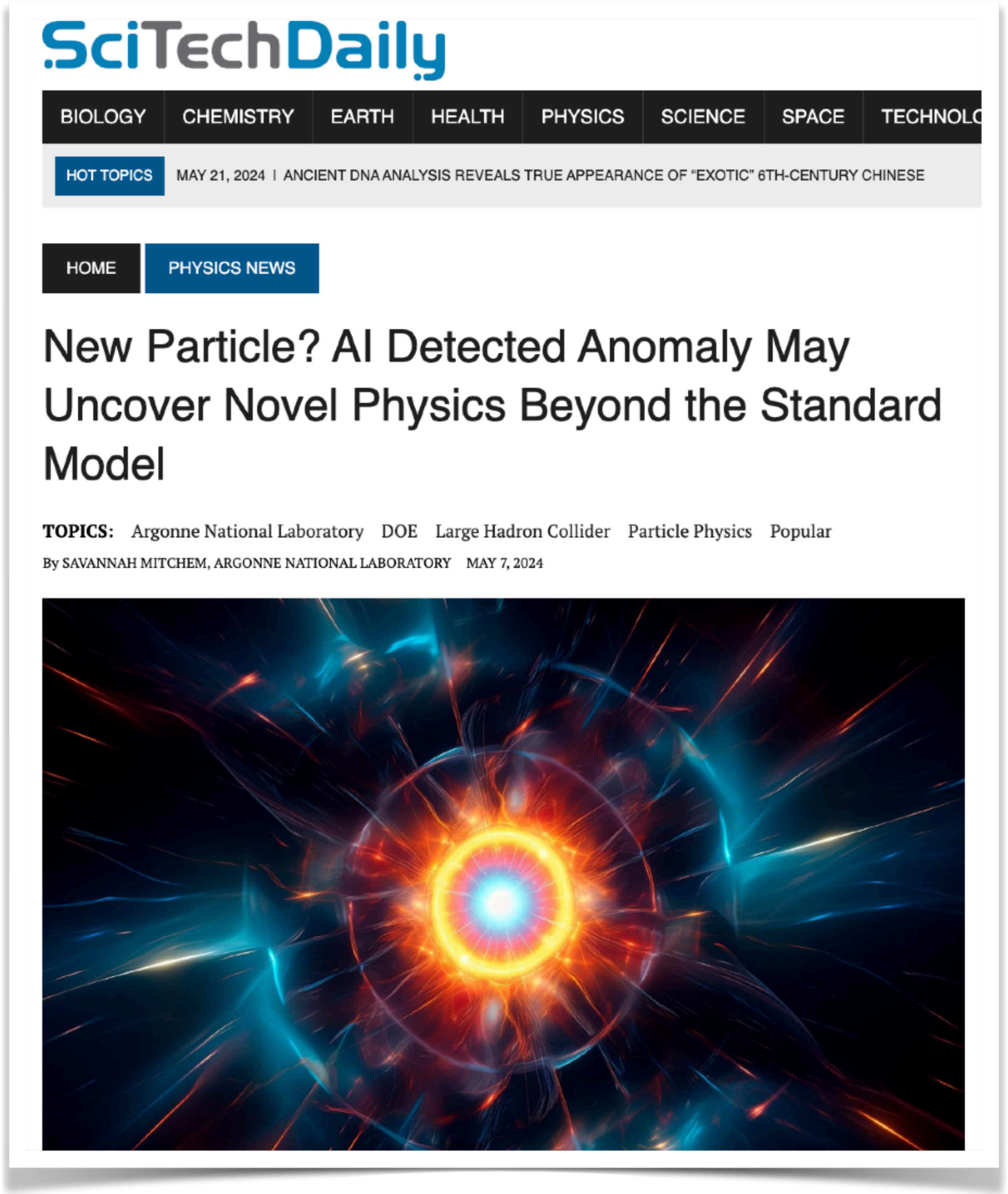
Dag Gillberg, June 5, 2024
LHCP 2024, Boston

ATLAS EXPERIMENT

Carleton University

Overview

- Anomaly detection
  - Growing use in experimental particle physics
  - Generic, model-independent searches
    - No need of signal MC for all models to be probed
- Three ATLAS results covered in this presentation **Analysis 1**
  - Dijet resonance search with CWoLa, *weakly supervised learning*, PRL (2020)
  - Search for $\boldsymbol{H}$ + generic new particle $\boldsymbol{X}$, *unsupervised learning*, PRD (2023)
  - Search for J + generic new particle $\boldsymbol{X}$, *unsupervised learning*, PRL (2024)
    - *Making headlines* **Analysis 3**



SciTechDaily

BIOLOGY | CHEMISTRY | EARTH | HEALTH | PHYSICS | SCIENCE | SPACE | TECHNOLO

HOT TOPICS MAY 21, 2024 | ANCIENT DNA ANALYSIS REVEALS TRUE APPEARANCE OF "EXOTIC" 6TH-CENTURY CHINESE

HOME | PHYSICS NEWS

New Particle? AI Detected Anomaly May Uncover Novel Physics Beyond the Standard Model

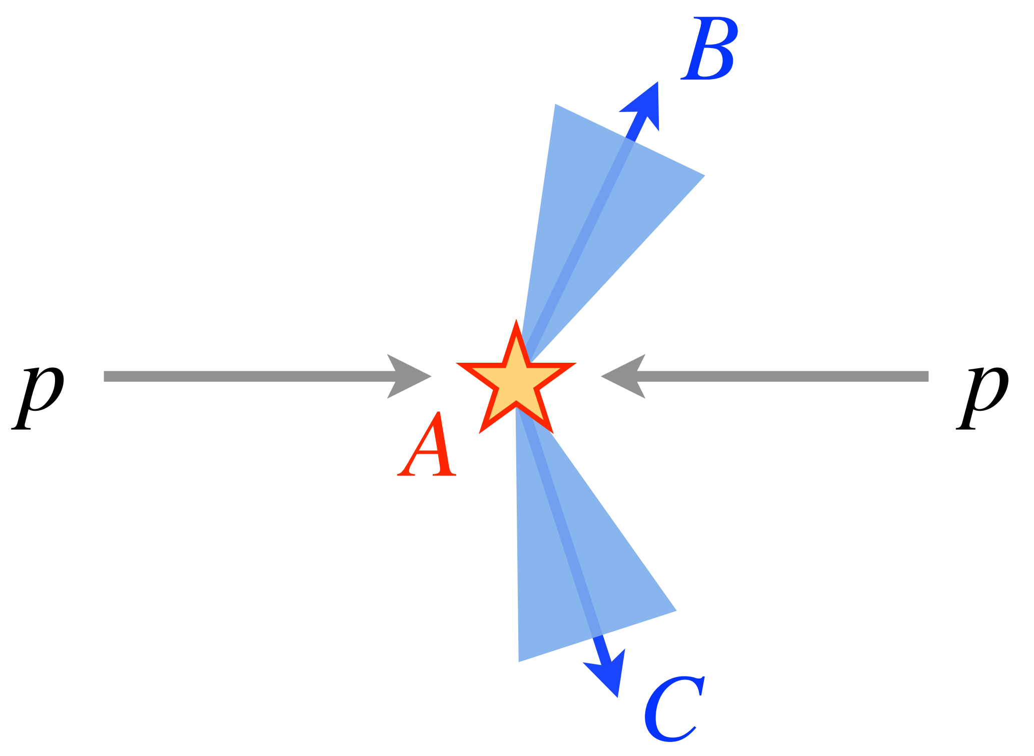
TOPICS: Argonne National Laboratory DOE Large Hadron Collider Particle Physics Popular

By SAVANNAH MITCHEM, ARGONNE NATIONAL LABORATORY MAY 7, 2024

Note: None of these analyses use MC for the main analysis!
(Only for validation of the method)

PRL 125 (2020) 131801

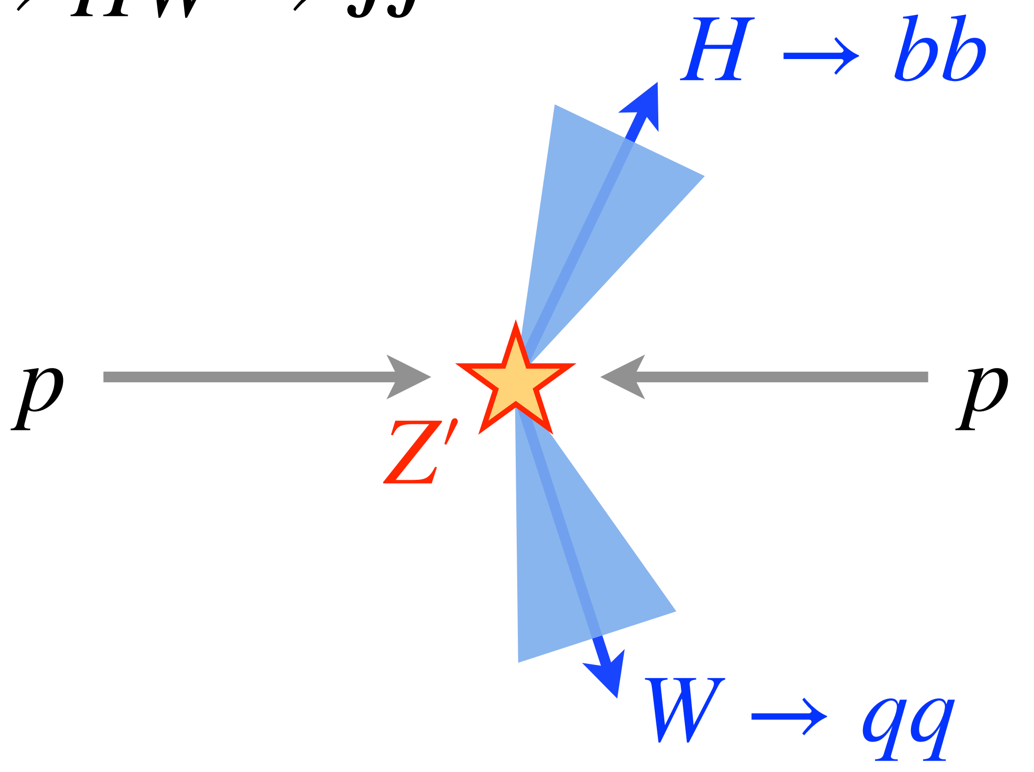
Dijet search with CWoLa: Overview



- Generic signal search for $A \to BC,\ B, C \to$ hadrons
- Final state: two large radius jets (JJ)
- Key parameters are mass of the three objects: m_A, m_B, m_C
    - Probing large parameter space:
      $m_A > 1.9\,\text{TeV}$, $m_B > 30\,\text{GeV}$, $m_C > 30\,\text{GeV}$
- Jets: anti-k_t $R = 1.0$, local hadron calibration
    - Trimming applied
- Dataset: Full Run 2 dataset (2015-2018) $139\,\text{fb}^{-1}$
- Selection
    - Unprescaled large-radius jet triggers
    - Two large radius jets with $p_\text{T} > 200\,\text{GeV}$
    - Leading jet $p_\text{T} > 500\,\text{GeV}$
    - Rapidity separation $\Delta y_{JJ} < 1.2$

Example of BSM process probed:

$Z' \to HW \to JJ$



$m_B \approx 125\,\text{GeV}$, $m_C \approx 80\,\text{GeV}$

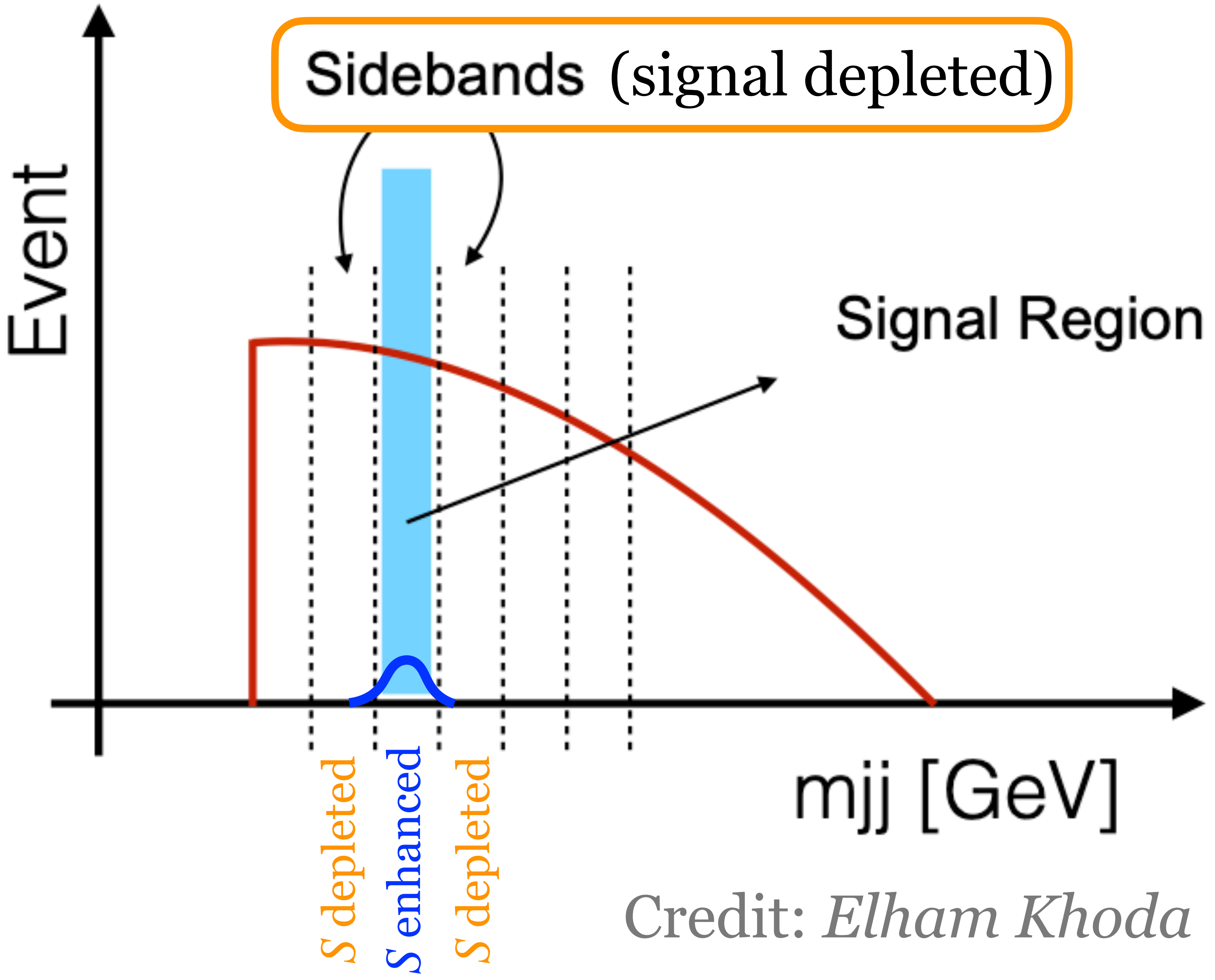
$m_A = m_{Z'}$ probed in [2.3, 6.8] TeV

PRL 125 (2020) 131801

Dijet search with CWoLa: Method

- *Classification Without Labels*
- **Only data** used as input, split in **two parts** based on m_{JJ}
  - SR centred on probed dijet mass $m_{JJ} \approx m_A$
  - Sidebands: nearby m_{JJ} (similar kinematics)
- Assumption: **Mixed samples** *different admixtures of S and B*
- Train DNN to distinguish (classify) SR vs sidebands using the two large radius jet masses as input: m_1 and m_2
  - *Training classifier using impure samples is asymptotically equivalent to training pure samples*
  - Train:Validation:Test split is 3:1:1
- Eight different SRs considered, bins of m_{JJ}
  - A separate analysis is performed for each such bin

CWoLa is ***weakly supervised anomaly detection method***.
Need to define two datasets with different presumed admixtures of *S* and *B*.



Credit: *Elham Khoda*

m_{JJ} boundaries for the 8 m_{JJ} regions [TeV]
[1.90, 2.28, 2.74, 3.28, 3.94, 4.73, 5.68, 6.81, 8.17]

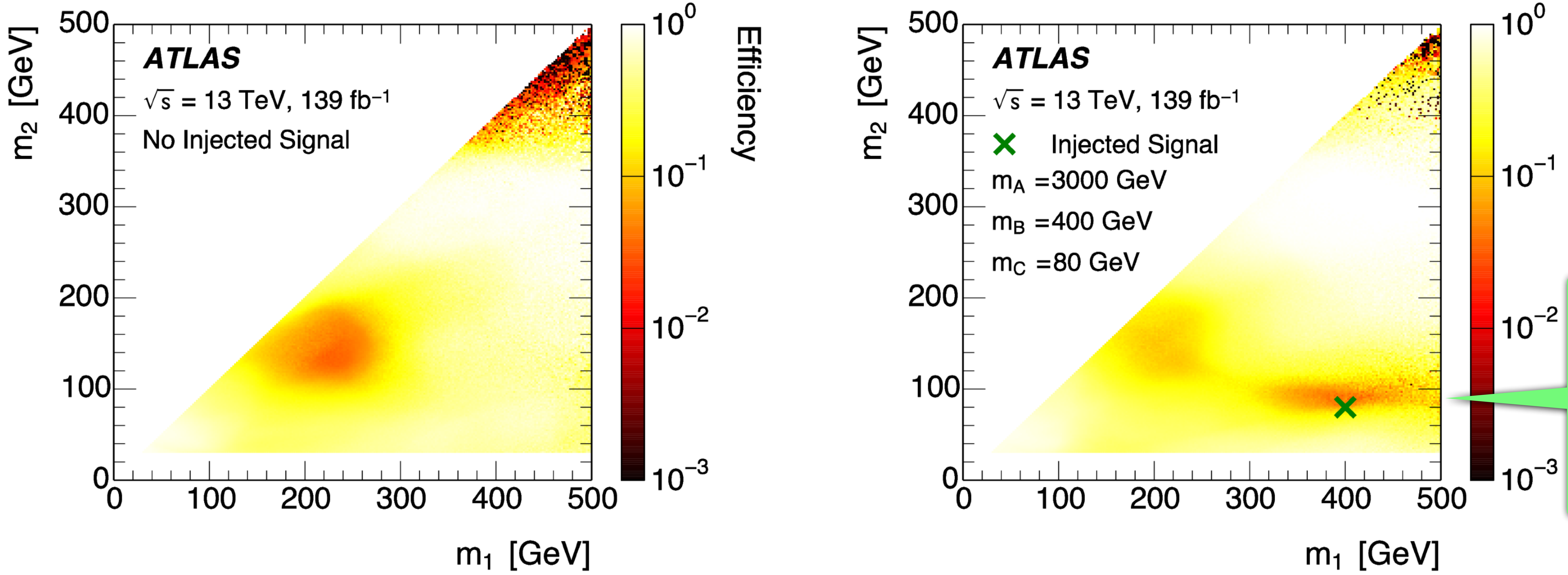
Bin widths > detector resolution.
A narrow m_A resonance stay within a bin.

PRL 125 (2020) 131801

Dijet search with CWoLa: Demonstration

- Two selections are performed on the DNN classification output
  - Medium, $\varepsilon = 0.1$: Keeping the 10% most anomalous (SR-like) events
  - Tight, $\varepsilon = 0.01$: Keeping the 1% most anomalous events
- Can scan the sample to see if these events cluster at any given (m_{J1}, m_{J2}) region

Mapped efficiency in SR defined by $m_{JJ} \in [2.74, 3.24]$ GeV



Test of method using MC input. CWoLa DNN correctly identifies anomaly around injected signal.

PRL 125 (2020) 131801

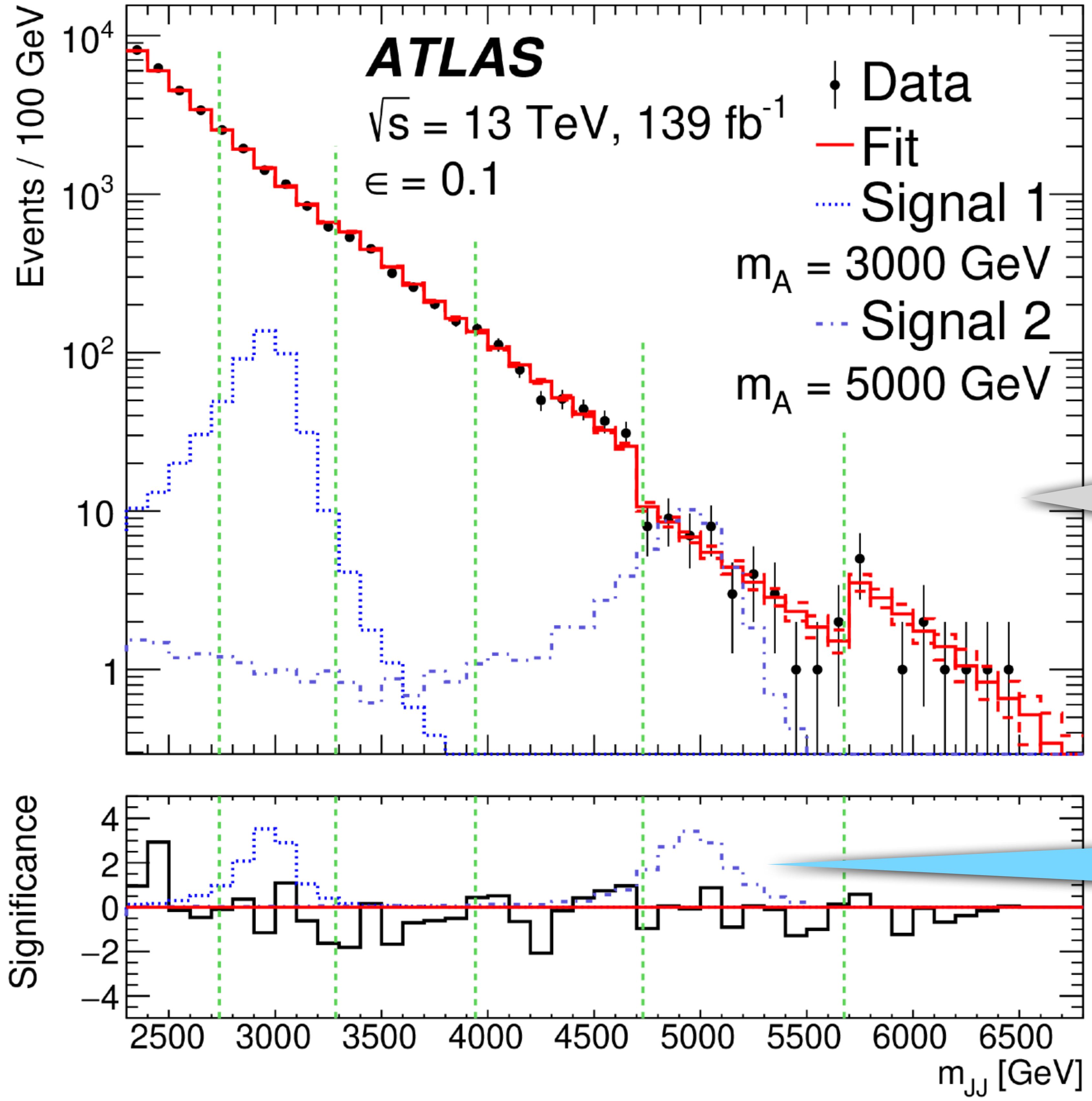
Dijet search with CWoLa: Results

- After applying the anomaly selection (e.g. $\epsilon > 0.1$), the m_{JJ} spectrum is fit using a parametric function
  - If any signal is present, S/B will be greatly enhanced, and hence signal sensitive increased
- 3 separate parametric functions are used, and a bump hunt is performed

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

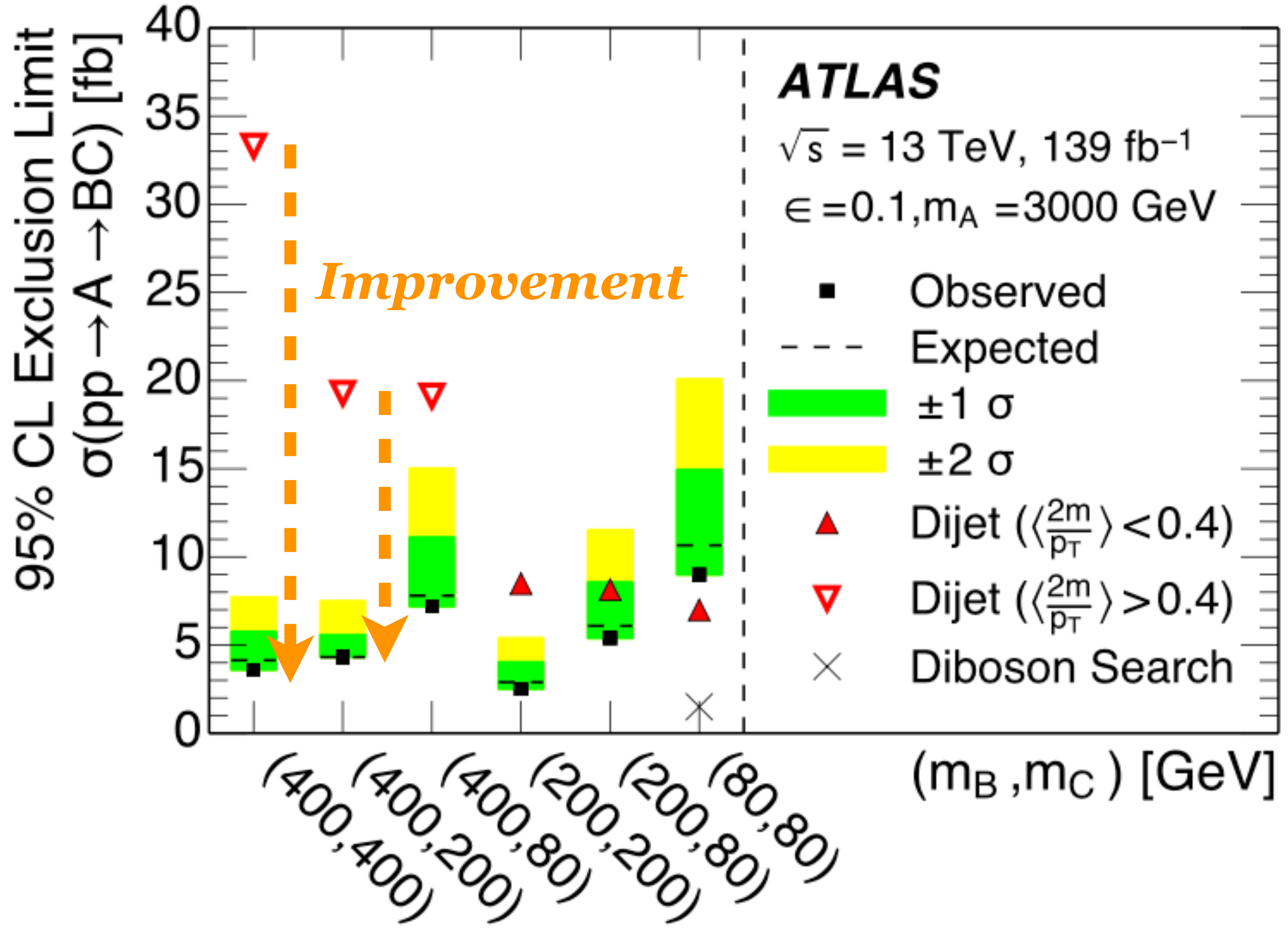
$$\frac{dn}{dx} = p_1(1-x)^{p_2-\xi_1 p_3}x^{-p_3+(p_4-\xi_2 p_3-\xi_3 p_2)\log(x)} \quad (2)$$

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



Results from six SRs. Each SR use own NN and hence cut efficiency. Its fit extends beyond its m_{JJ} range. This plot is hence 'stitched': shows key range of six SRs.

Expected results from BSM signals with $m_A = 3$ TeV and $m_A = 5$ TeV with cross section just below previous upper limits.

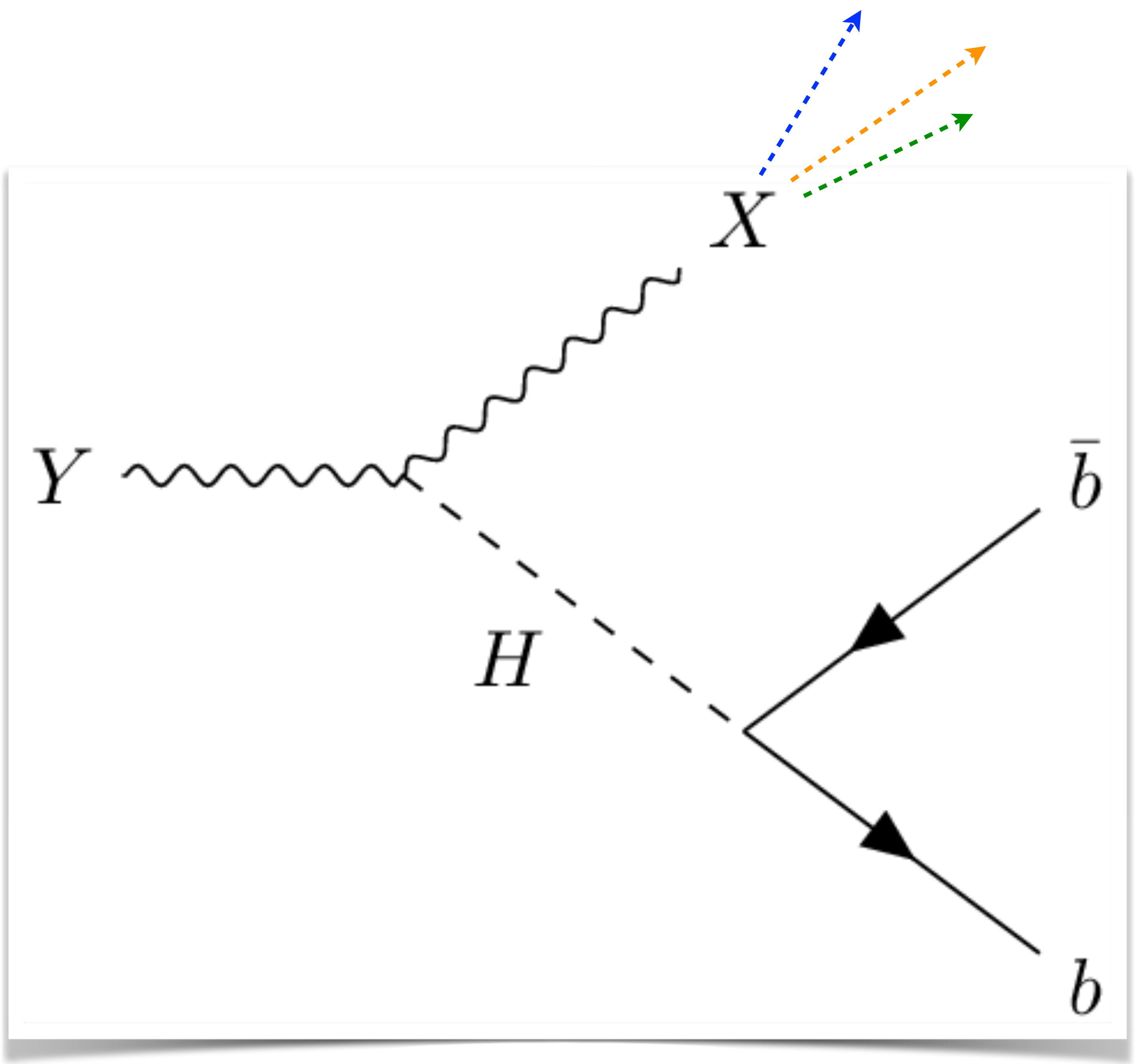


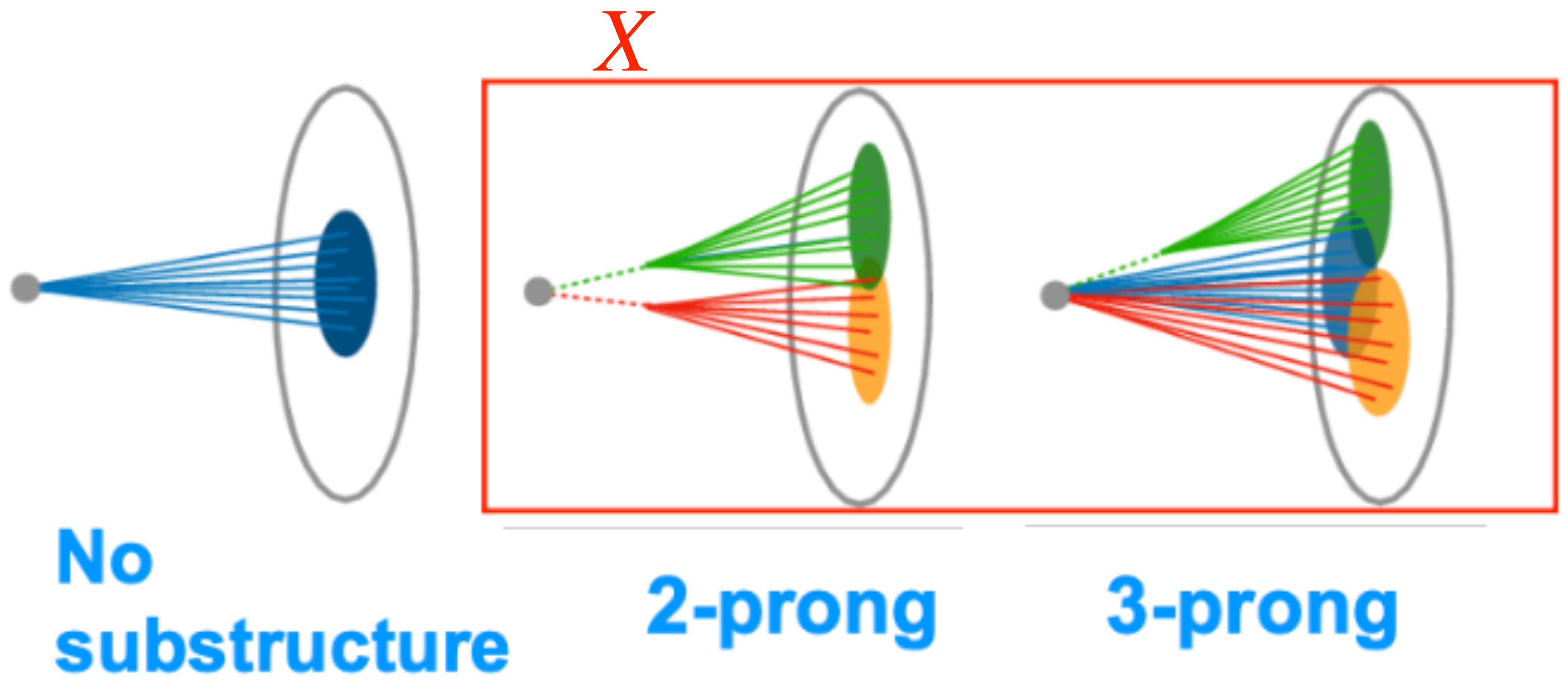
Example limits for different (m_A, m_B, m_C) points. Significant improvement wrt previous analysis!

PRD 108 (2023) 052009

Higgs + X Search: Overview

- Search for heavy resonance Y decaying to HX, with $H \rightarrow bb$, $X \rightarrow$ hadrons
  - Extensive experience identifying $H \rightarrow bb$ with high efficiency
- Heavy resonance Y probed in range $m_Y \in [1,6]$ TeV
- Analysis sensitive to wide arrange of possible decays of **particle** X
  - Two-prong decay ($X \rightarrow qq$) used for benchmark
  - Many other decay topologies checked: three-prong, displaced vertices (heavy-flavour), **dark jets** (patterns of missing and visible energy)
  - Sensitive to wide range of mass: $m_X \in [\sim50, \sim3000]$ GeV
- Signal grid: (m_Y, m_X)
- Run 2 dataset collected using large-radius jet triggers
  - $p_{\mathrm{T},J} > 500$ GeV, $m_{JJ} > 1.3$ TeV
  - $H \rightarrow bb$ identified using NN-based tagging+$m_{bb} \in (75,145)$ GeV





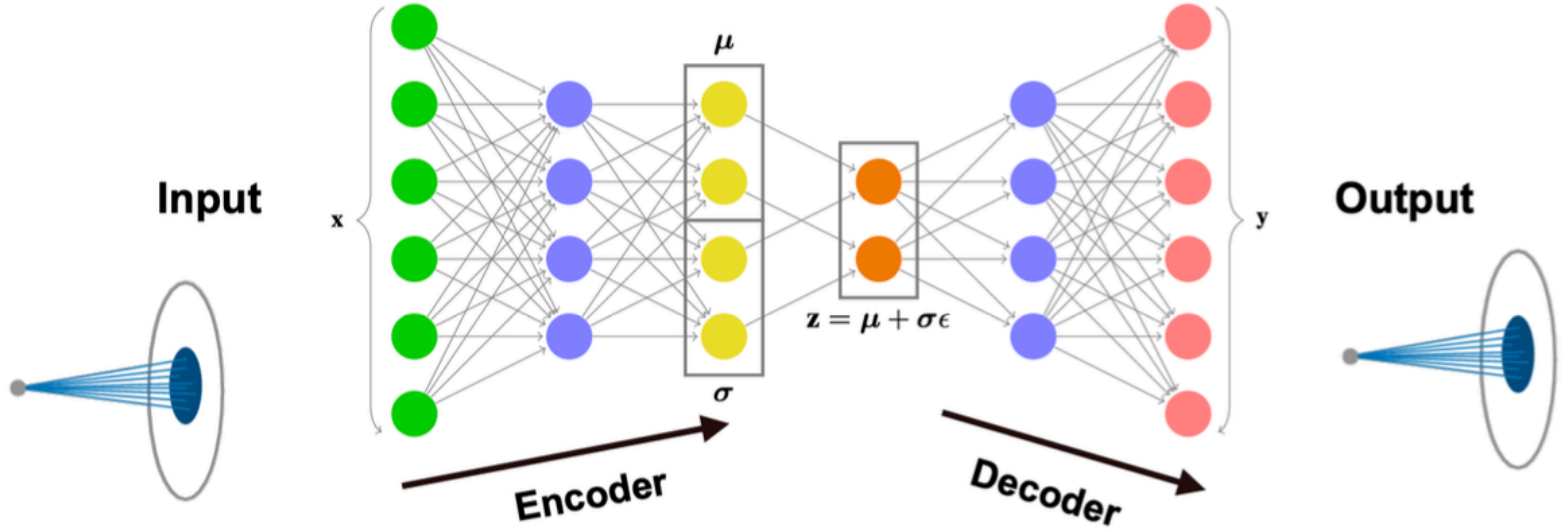
PRD 108 (2023) 052009

Higgs + X Search: Anomalous jet tagging

- Variational recurrent neural networks (VRNN) are used to identify the anomalous X candidate
  - VRNN trained using all jet constituent 4-momenta as input
  - Assigns a per-jet anomaly score (AS)
  - Unsupervised training
  - AS defined from the VRNN loss function:

Loss Output Input Kullback-Leibler Divergence

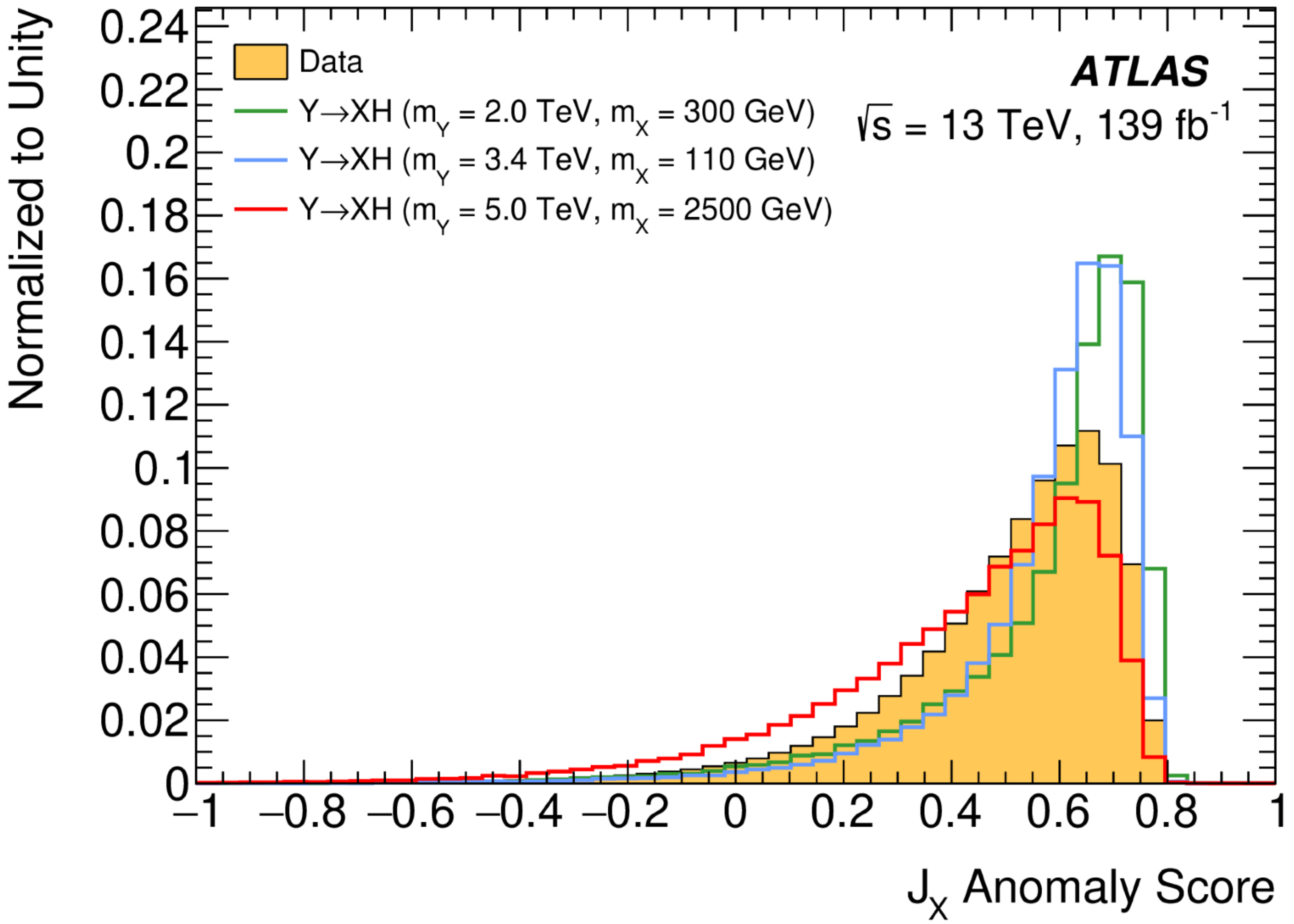
$$\mathcal{L}(t) = |\mathbf{y}(t) - \mathbf{x}(t)|^2 + \lambda D_{\mathrm{KL}}(z||z_t)$$

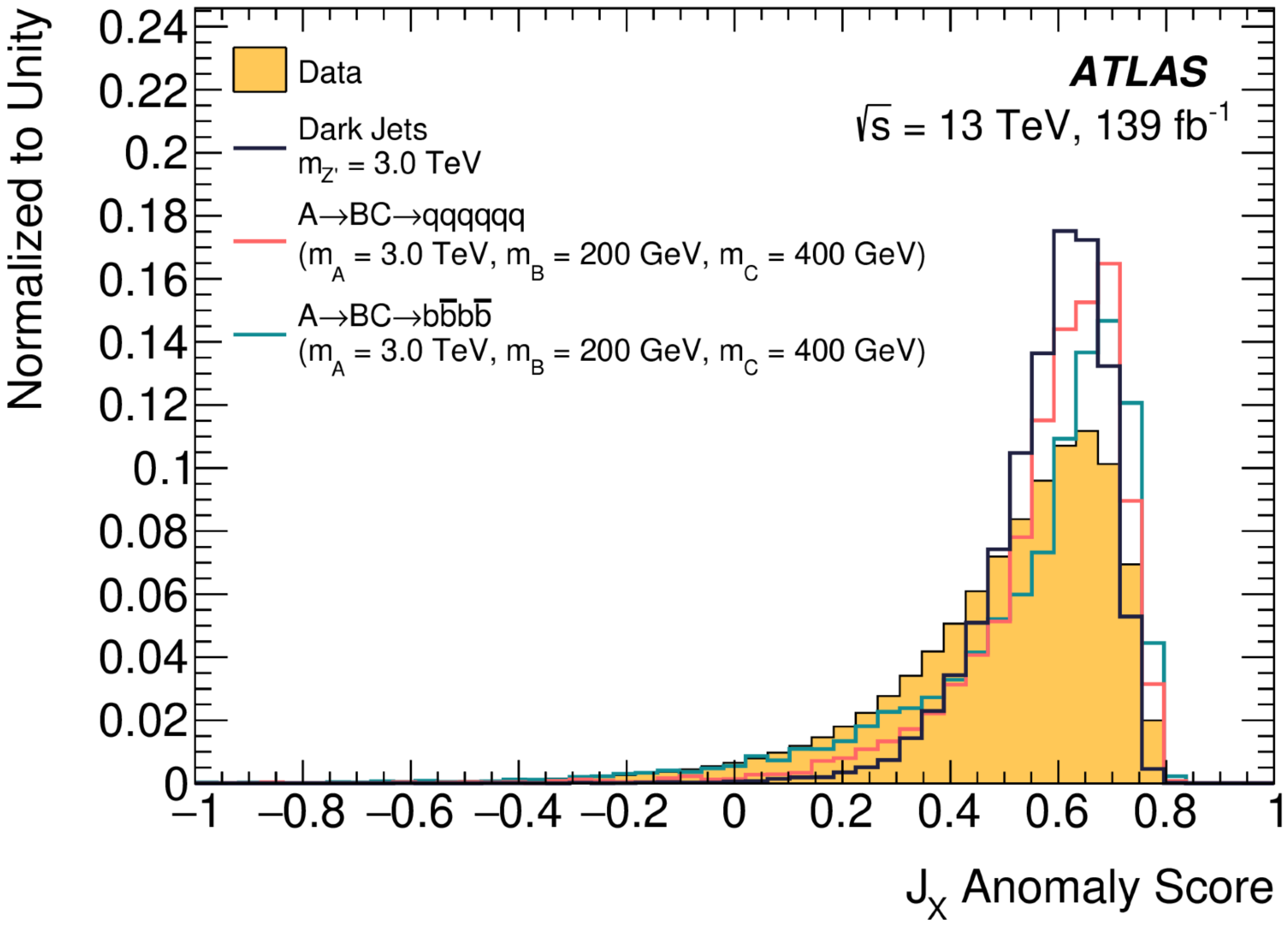


Analysis 2

Higgs + X Search: Demonstration

- Analysis is tested using different BSM scenarios
  - Generic $Y \rightarrow XH \rightarrow qqbb$ (left)
  - $X \rightarrow$ dark jet, 3 prong decay, heavy flavour (right)
- Real data shown as yellow filled histogram
- Jets are deemed anomalous if $J_X > 0.5$

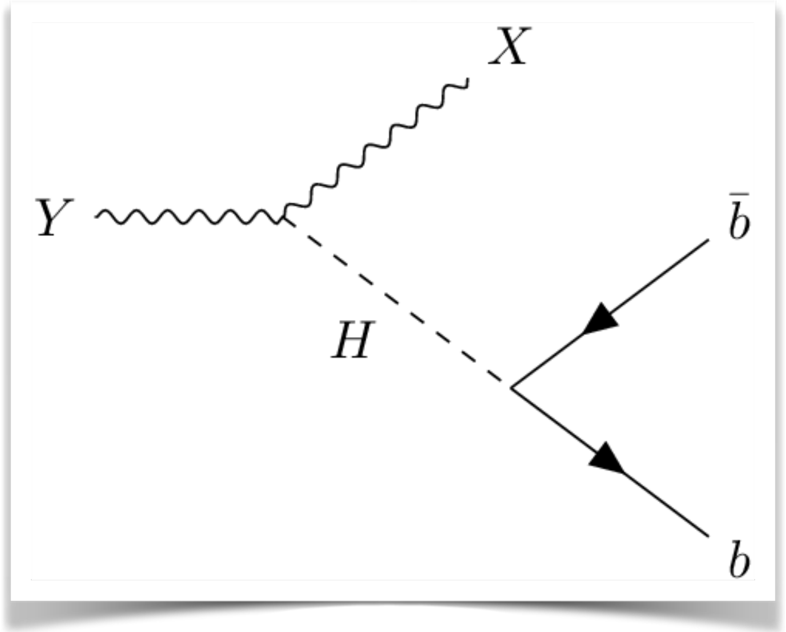




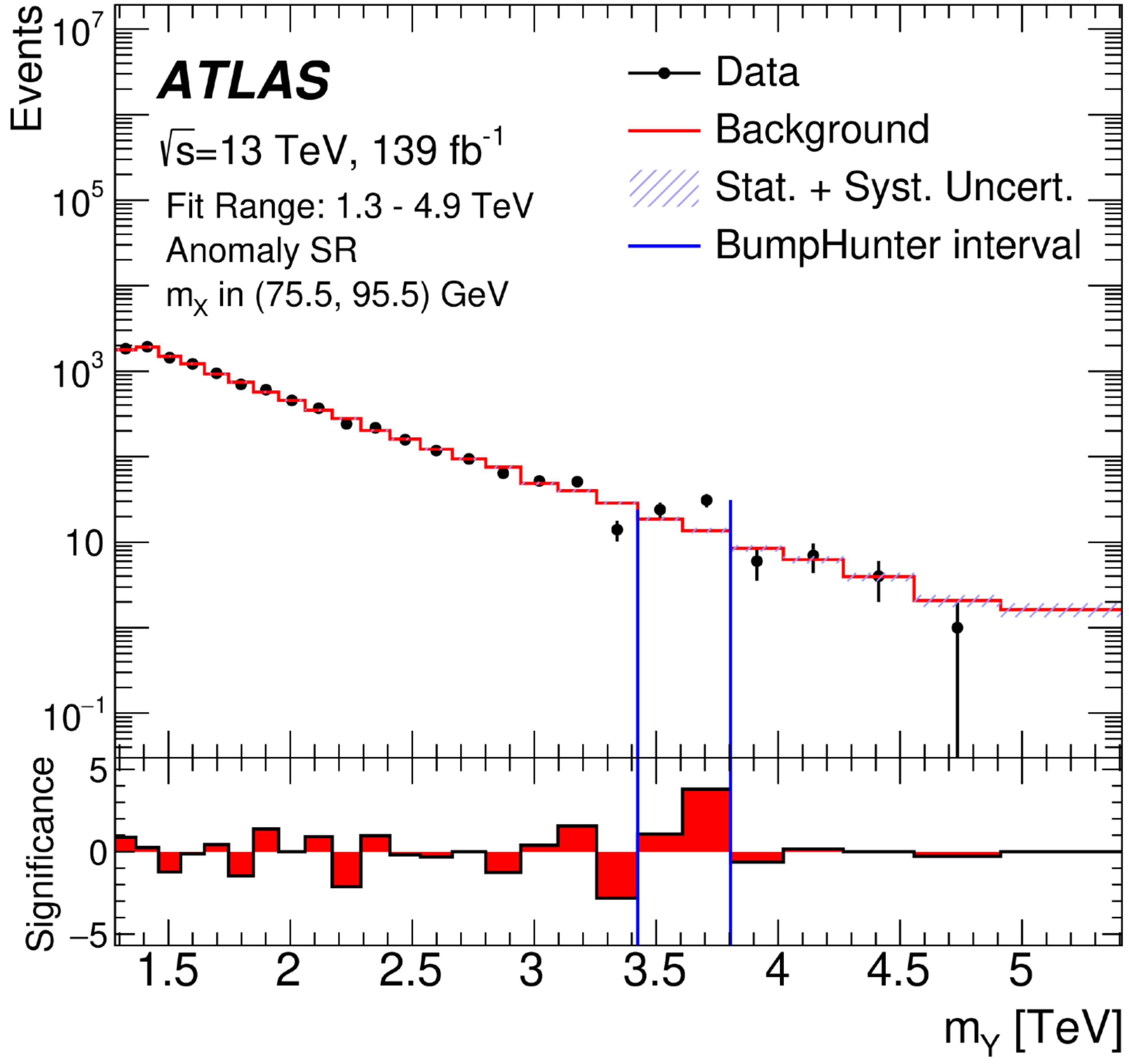
Analysis 2

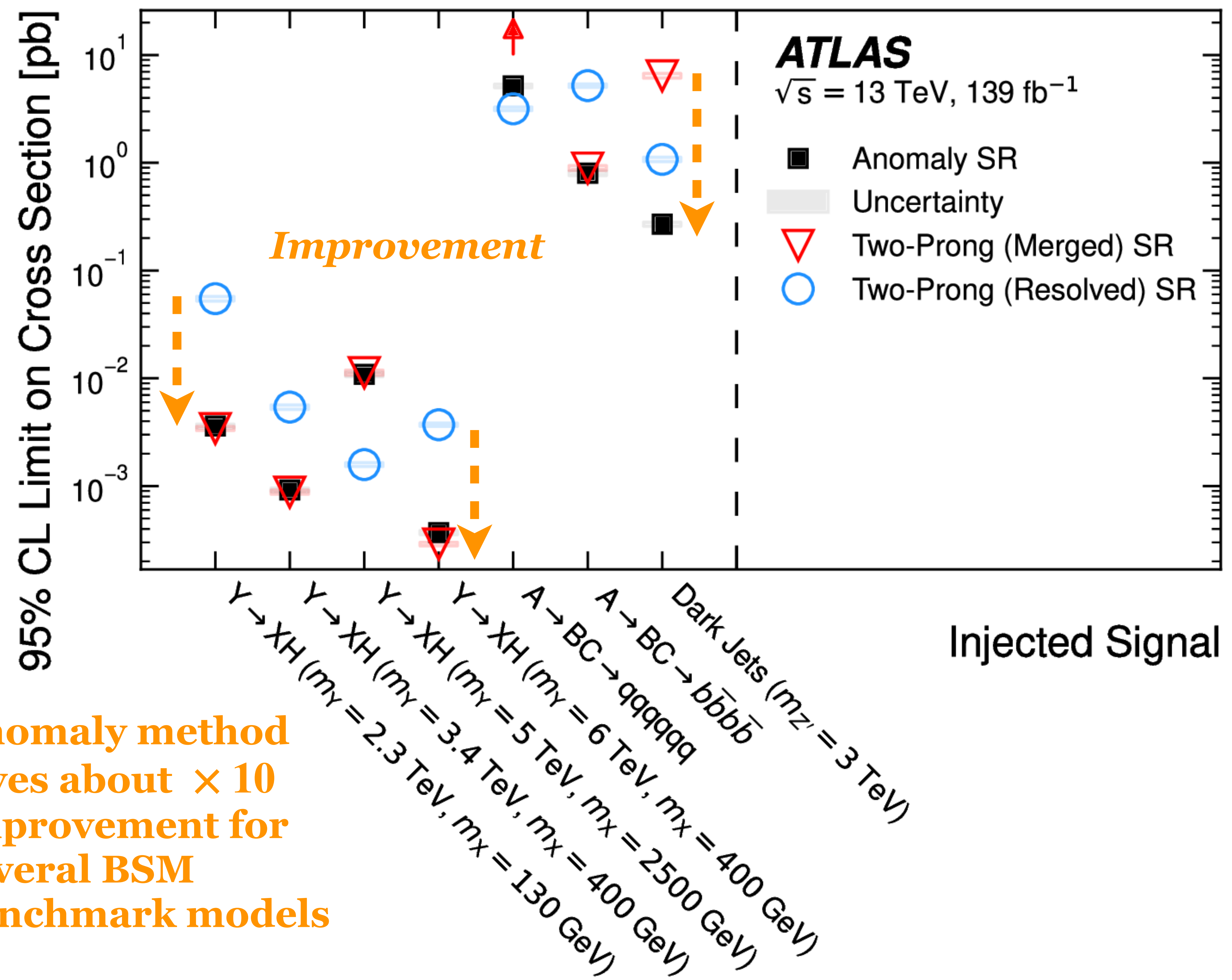
PRD 108 (2023) 052009

Higgs + X Search: Results



- In the anomaly SR ($J_X > 0.5$), a bump-hunt is performed using a (m_X, m_Y) grid
- Largest excess for $m_X \approx 85$ GeV, $m_Y \approx 3.7$ TeV (see below). But global significance (only) 1.4σ.
- Upper limits set on benchmark models by injection of MC signal. Compared to traditional two-prong tagging





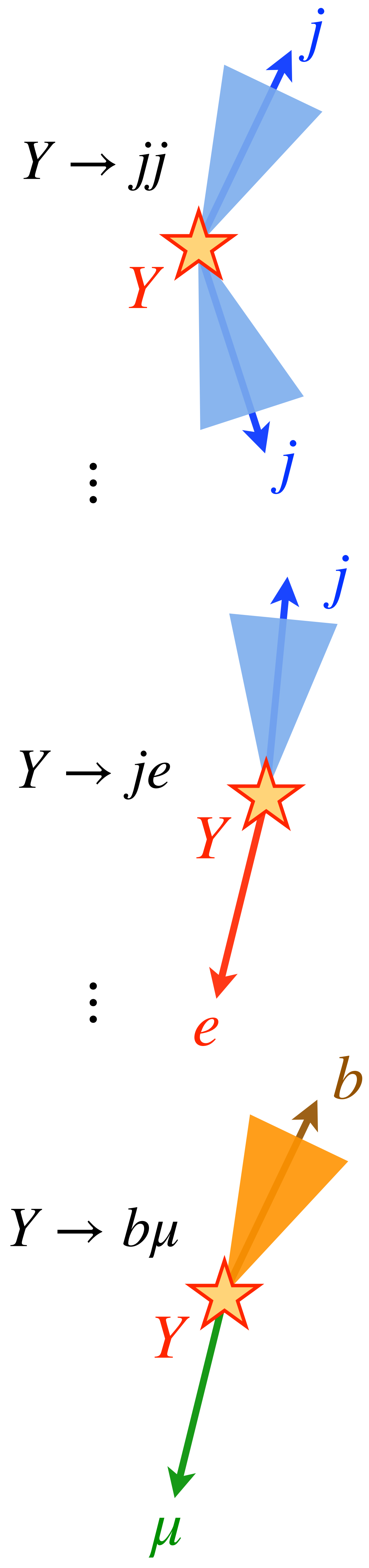
Anomaly method gives about ×10 improvement for several BSM benchmark models

PRL 132 (2024) 081801

Jet + X search: Overview

j = light jet
b = b-tagged jet
e = electron, …

- Search for generic mass resonance in a BSM 2-body decay:
  - 9 final states considered: $jj, jb, bb, je, be, j\gamma, b\gamma, j\mu, b\mu$
- All events are required to contain a lepton (e or μ) with $p_T > 60$ GeV (trigger)
  - Reduces QCD multijet background
  - Unbiased selection of (b)jets: gives access to low jet p_T region (down to 30 GeV!)
  - *Large phase-space probed (kinematics+particle type and multiplicity)*
- Use unsupervised learning to identify **anomalous** events
  - Input large range of kinematic features (more later)
  - Anomalous detectrion baed on **autoencoder**
- Mass spectrum scanned for bump in the nine probed final states, i.e.
  - $m_{jj}, m_{jb}, m_{bb}, m_{be}, \ldots, m_{b\mu}$



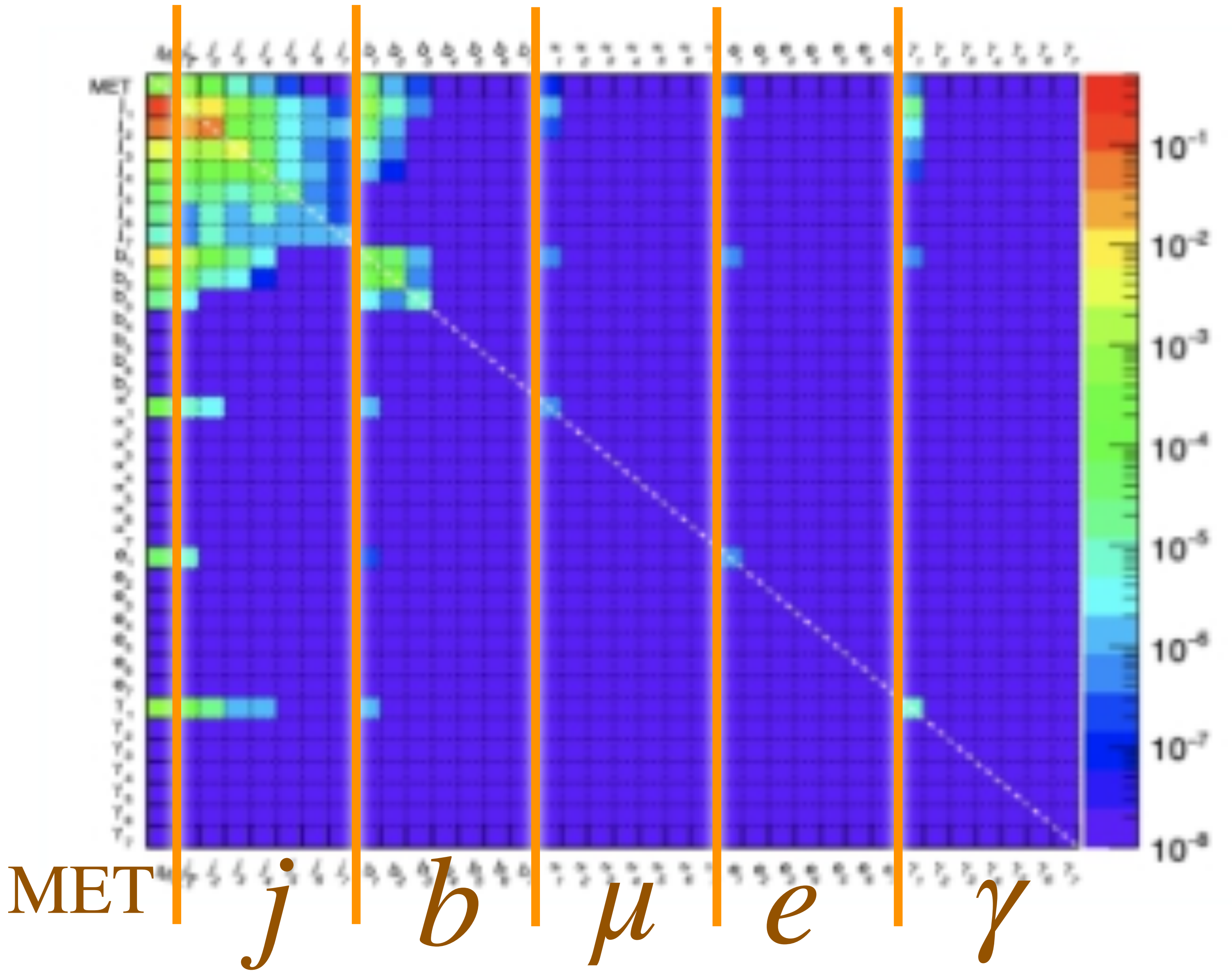
PRL 132 (2024) 081801

Jet + X search: Kinematic input

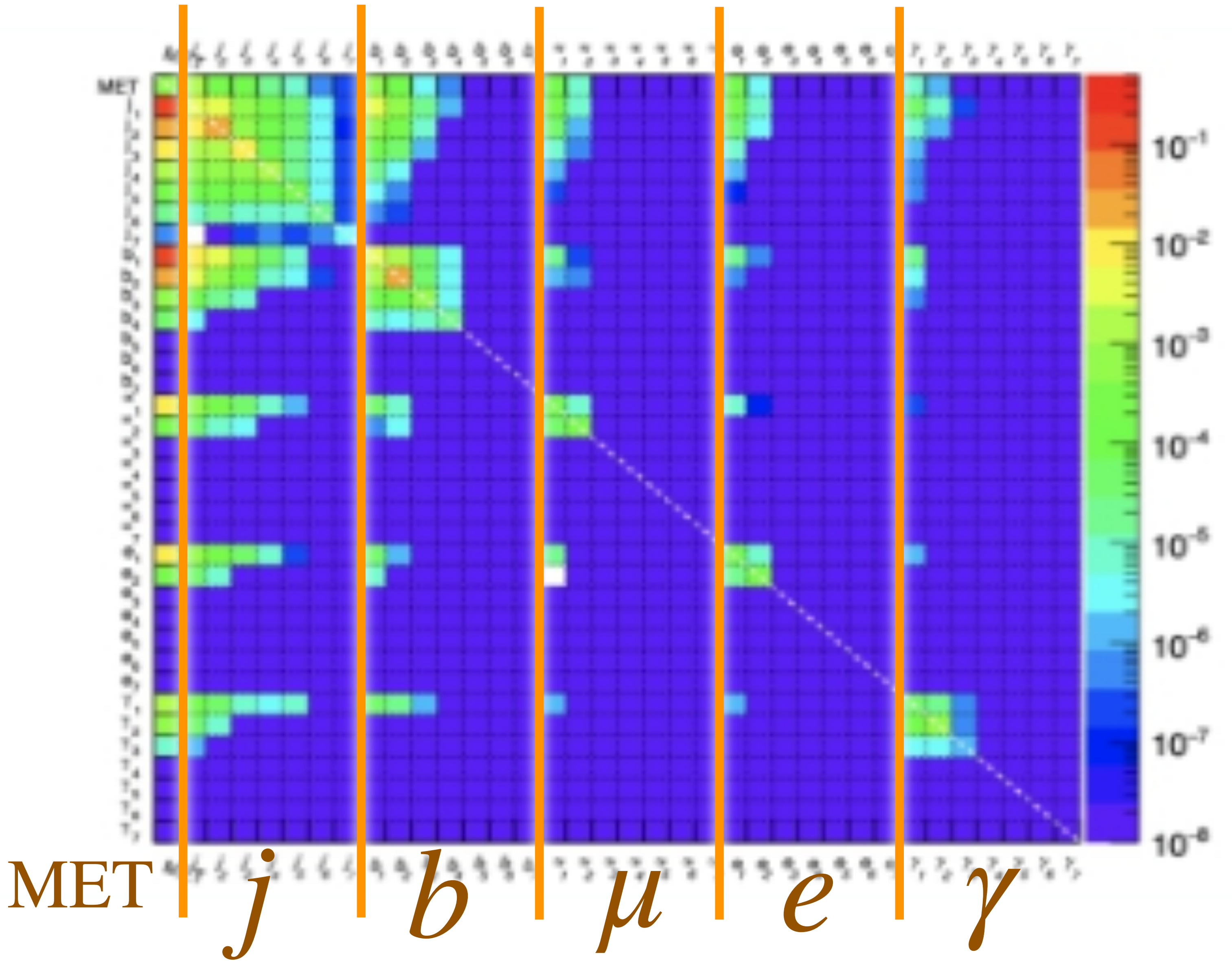
- Kinematics of each event is encoded in the **Rapidity Mass Matrix**
- Contains key kinematic variables suitable for exotic searches—more robust performance than four-momenta
  - Up to 36 physics objects considered for each event: 10 jets, 10 b-jets, 5 e, 5 μ, 5 γ + $E_{\mathrm{T}}^{\mathrm{miss}}$
- Holds metrics of: y, m_{T} for each object, diagonal: p_{T}-imbalance, Δy + inv. mass m_{ab} for each obj. pair ab
  - Exclude the 9 invariant masses probed in the end (leading objects only)
  - Number of elements: $36^2 - 9 = 1287$
- No object → elements = 0. Pre-processing: elements $\in [0,1]$

$$
\begin{pmatrix}
\mathbf{e}_{\mathbf{T}}^{\mathrm{miss}} & m_T(j_1) & m_T(j_2) & \dots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \dots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e}_{\mathbf{T}}(\mathbf{j}_1) & m(j_1,j_2) & \dots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \dots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e}_{\mathbf{T}}(\mathbf{j}_2) & \dots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \dots m(j_2,\mu_N) \\
\dots & \dots & \dots & \dots, & \dots & \dots & \dots \\
h_L(j_N) & h(j_1,j_N) & \dots & \dots\delta\mathbf{e}_{\mathbf{T}}(\mathbf{j}_N) & m(j_N,\mu_1) & m(j_N,\mu_2) & \dots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \dots h(\mu_1,j_N) & \mathbf{e}_{\mathbf{T}}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \dots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e}_{\mathbf{T}}(\mu_2) & m(\mu_2,\mu_N) \\
\dots & \dots & \dots & \dots & \dots & \dots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \dots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e}_{\mathbf{T}}(\mu_N)
\end{pmatrix}
$$

Multi-jet QCD process



Higgs process



Example RMMs from analysis
36 × 36 matrix.
Quite different depending on final state!
Objects are p_{T} sorted. Higher p_{T} to the left.

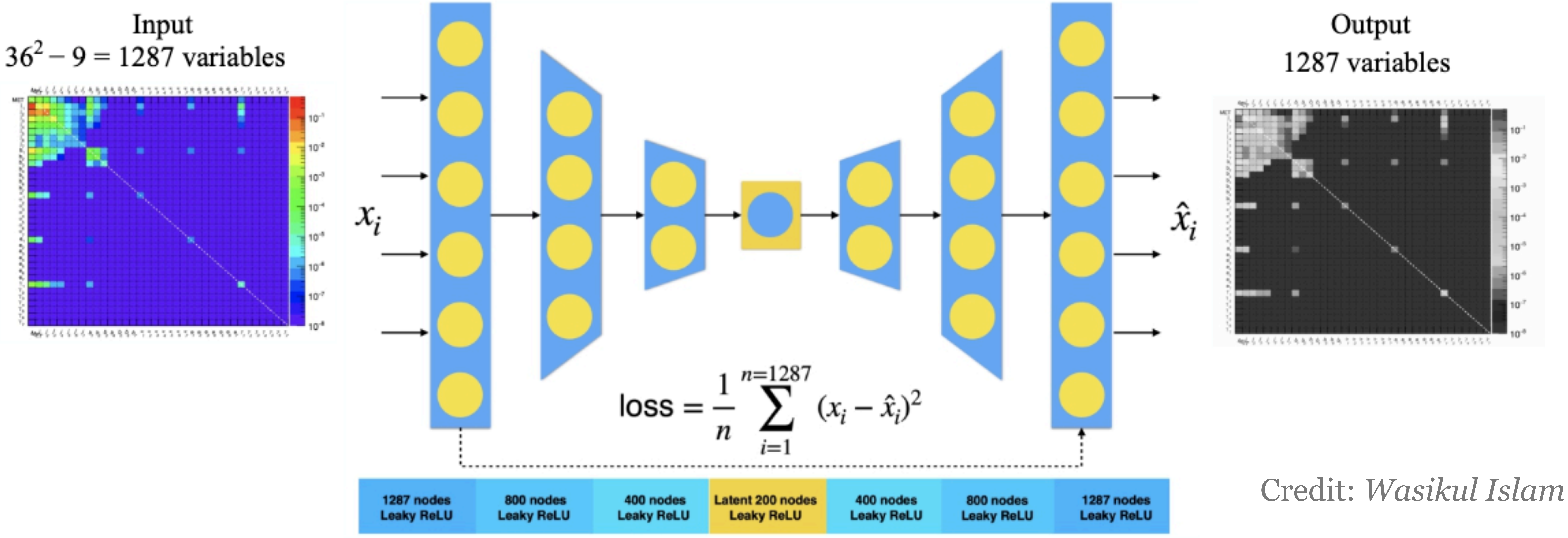
arXiv:1810.06669

PRL 132 (2024) 081801

Jet + X search: Method

- Autoencoder implemented using TensorFlow. Encoder+Decoder. Structure displayed below
- 1% of the data is randomly selected for training
  - Probability of finding BSM signal in this subset considered negligible
  - Provides sufficient training stats, further split 7:3, training:validation
- Key quantity: loss calculated between input and output.
  - An atypical event will get a large loss → deemed anomalous

Training repeated 50 times with different random initializations
→ 50 separate AEs
Generally similar performance.
Median loss AE used for end analysis.

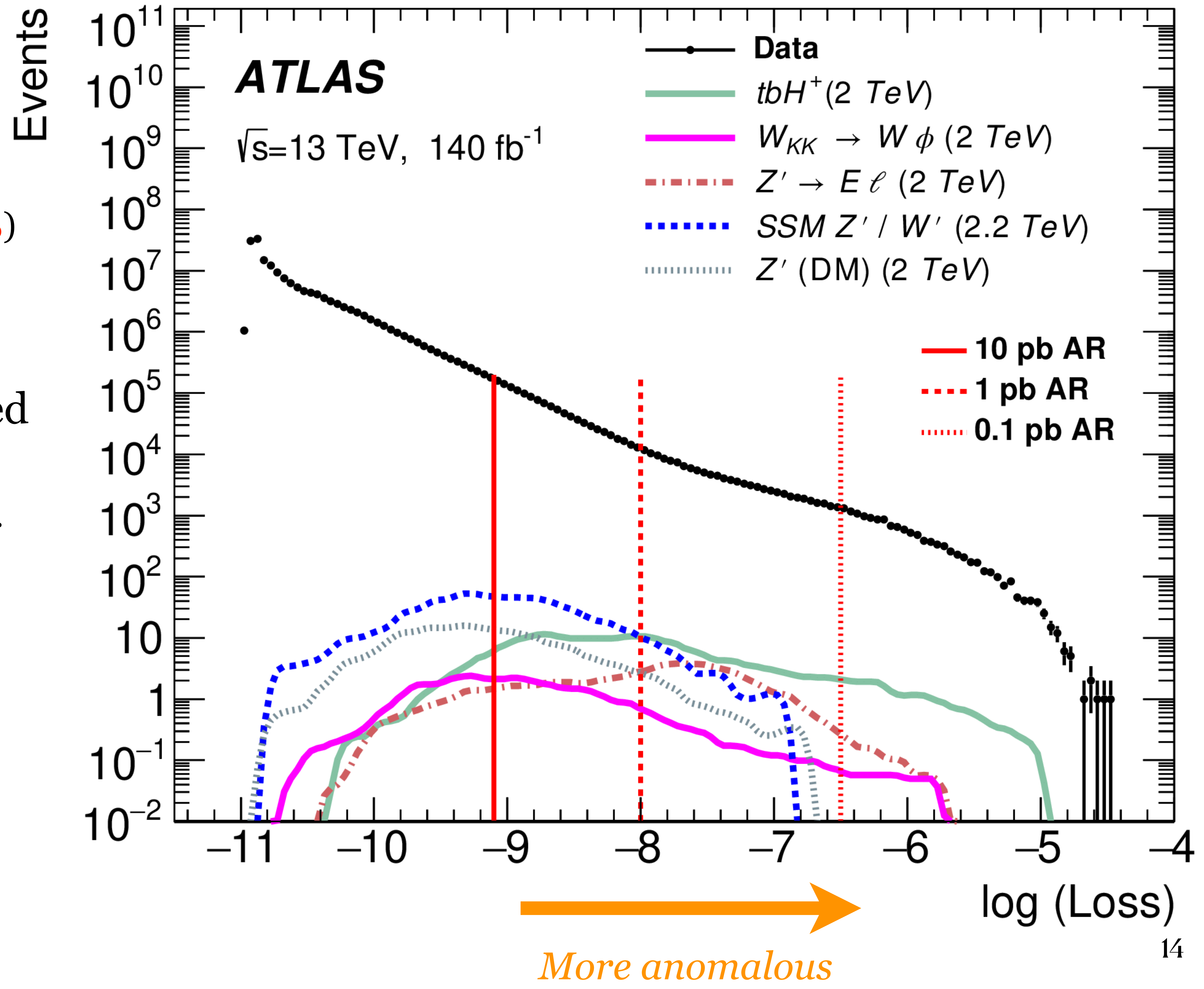


Credit: *Wasikul Islam*

PRL 132 (2024) 081801

Jet + X search: Result + validation

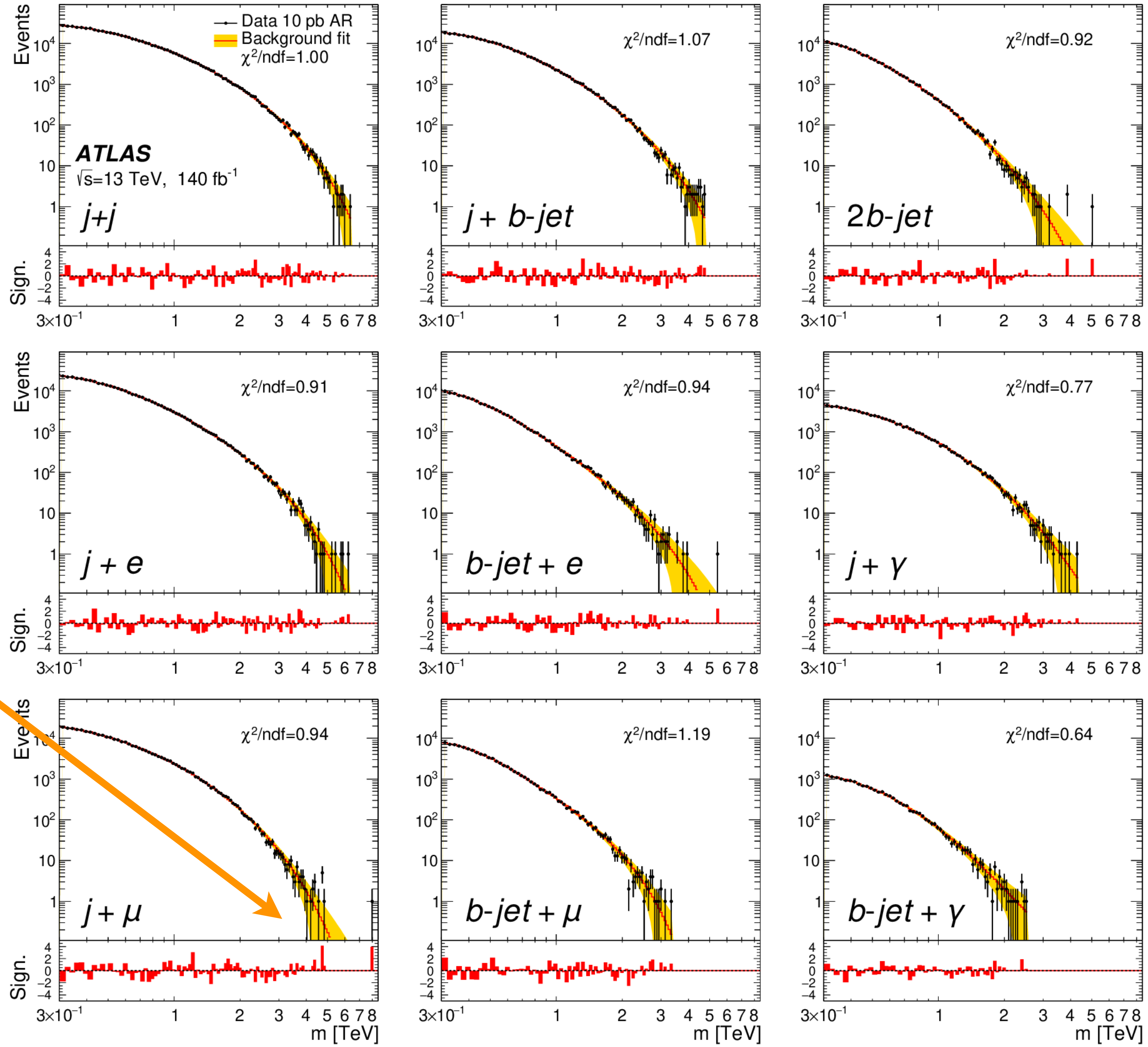
- Plot of anomalous score for all events
  - Data peaks at ≈ -11
- Three anomalous regions (ARs) defined:
  - > -9.1, keeps 10 pb/140 fb of data (0.0071%)
  - > -8.0, keeps 1 pb/140 fb of data
  - > -6.7, keeps 0.1 pb/140 fb of data
- MC predictions from BSM scenarios overlayed
  - Significantly shifted to high scores
- Anomalous score found robust as function of time and beam/pileup conditions
- Each region analyzed separately
  - Loosest region (10 pb AR) generally give bests sensitivity
  - Shown on next slide



Analysis 3

Jet + X search: Result

- The 9 mass distributions plotted in the three ARs (27 spectra).
- Bin width mimics detector resolution
- Here 10 pb AR
- Bump hunt performed.
- Largest access found for $Y \rightarrow j\mu$ at $m_{j\mu} = 4.8$ TeV
  - Local significance: 2.9 sigma
  - Not observed in 1 pb AR nor the 0.1 pb AR

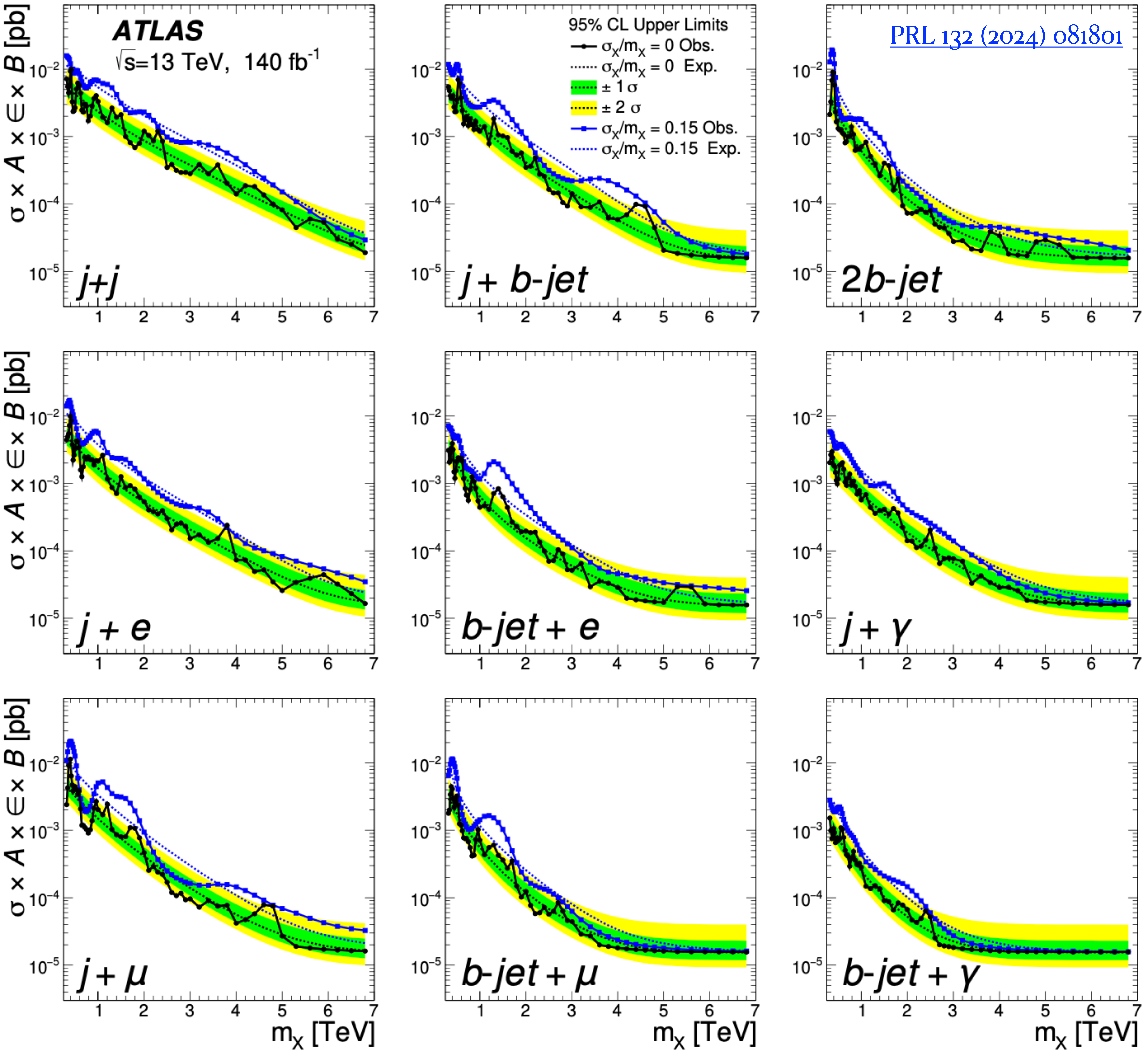


Analysis 3

Jet + X search: Limits

- Limits are placed based on Gaussian mass peaks with assuming different intrinsic width
  - Zero width (black points)
  - $\sigma_m/m = 0.15$ (blue points)
- The narrower width gives better limits (as expected)
- Zero width means the signal is assumed to give a Gaussian shape with a $\sigma =$ detector resolution
- Largest deviation

PRL 132 (2024) 081801

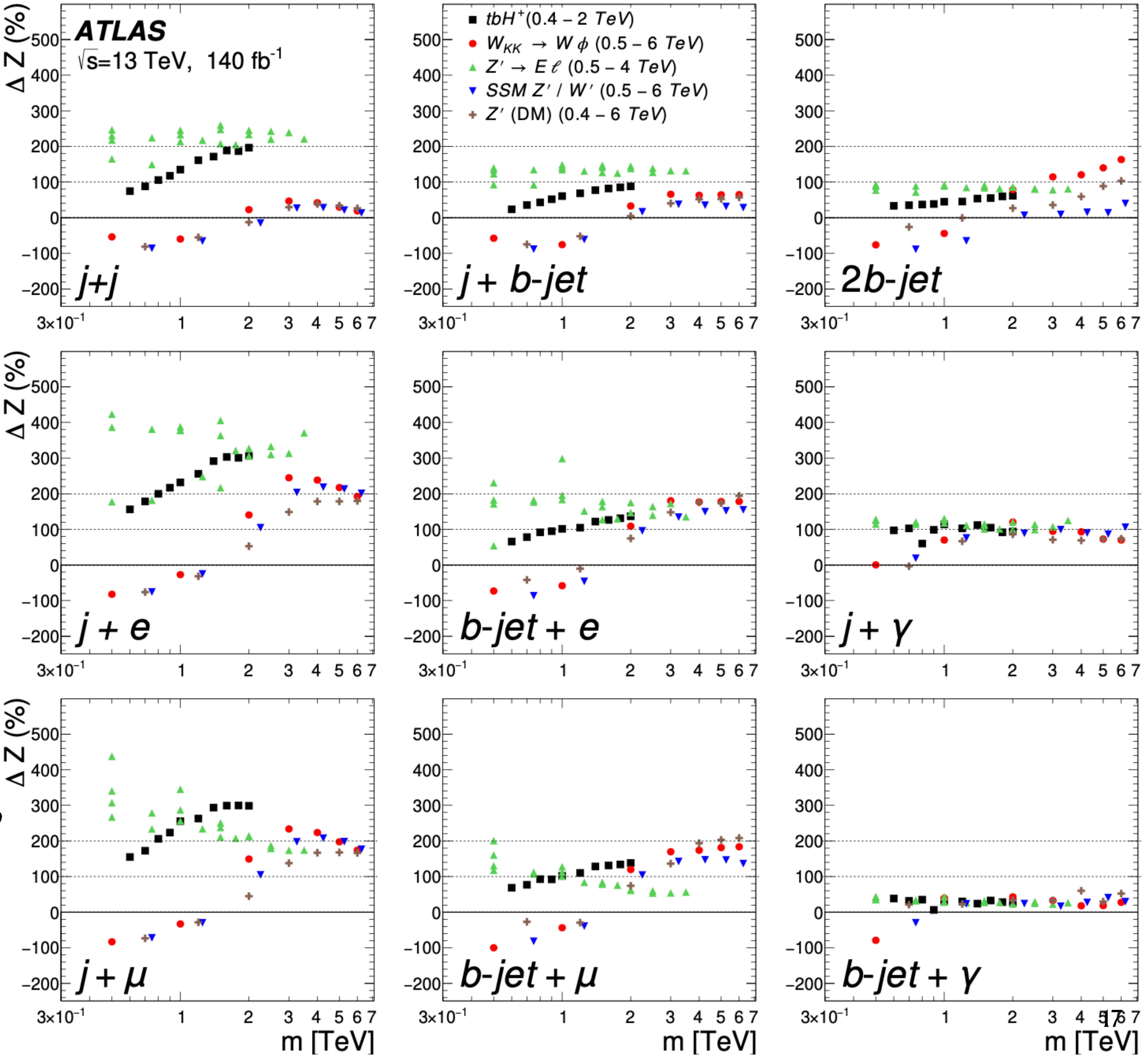


Jet + X search: AE improvement

- Here, results using the 10 pb AR are compared with all data
  - I.e. *using vs not using AE*
- Different benchmark BSM models are tested, and the significance often increase substantially.
  - +200% means significance $\times$ 3 (e.g. 1σ excess $\rightarrow$ 3σ excess)
- Some models with low mass with signature close to SM degrade
  - Loss in stats hurts more than gain in s/b

$$\Delta Z = ((Z_{\mathrm{AE}}/Z) - 1) \times 100\%$$

$$Z = \sqrt{2\left((s+b)\ln(1+s/b) - s\right)}$$



Conclusion

- ATLAS has published three analyses that use anomaly detection
  - All look for a mass resonance in a subset of events deemed anomalous
- Final states probed
  - Dijet search $Y \rightarrow AB$ (CWoLa method) **Analysis 1**
  - Dedicated $Y \rightarrow H + X, H \rightarrow bb$ **Analysis 2**
  - $Y \rightarrow j + X$, in events with at least one charged lepton (e or μ) **Analysis 3**
- In each case, the results are model independent
- No excess content with a mass resonance found
- Limits are placed under different assumption on the mass resonance width Γ_Y
- Anomaly detection is a powerful way to cast a wide net in the search for new physics

Backup