Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Anomaly Detection with CMS

Gregor Kasieczka, Louis Moureaux, Manuel Sommerhalder
Chitrakshee Yede, Tore von Schwartz
(chitrakshee.yede@cern.ch)

Glühwein Workshop, KIT
16th Dec. 2024

I will be talking about

1. Model-agnostic search for dijet resonances with anomalous jet substructure in proton-proton collisions at $\sqrt{s} = 13$ TeV

   The CMS Collaboration*

   (CMS-EXO-22-026↗)

2. skCATHODE

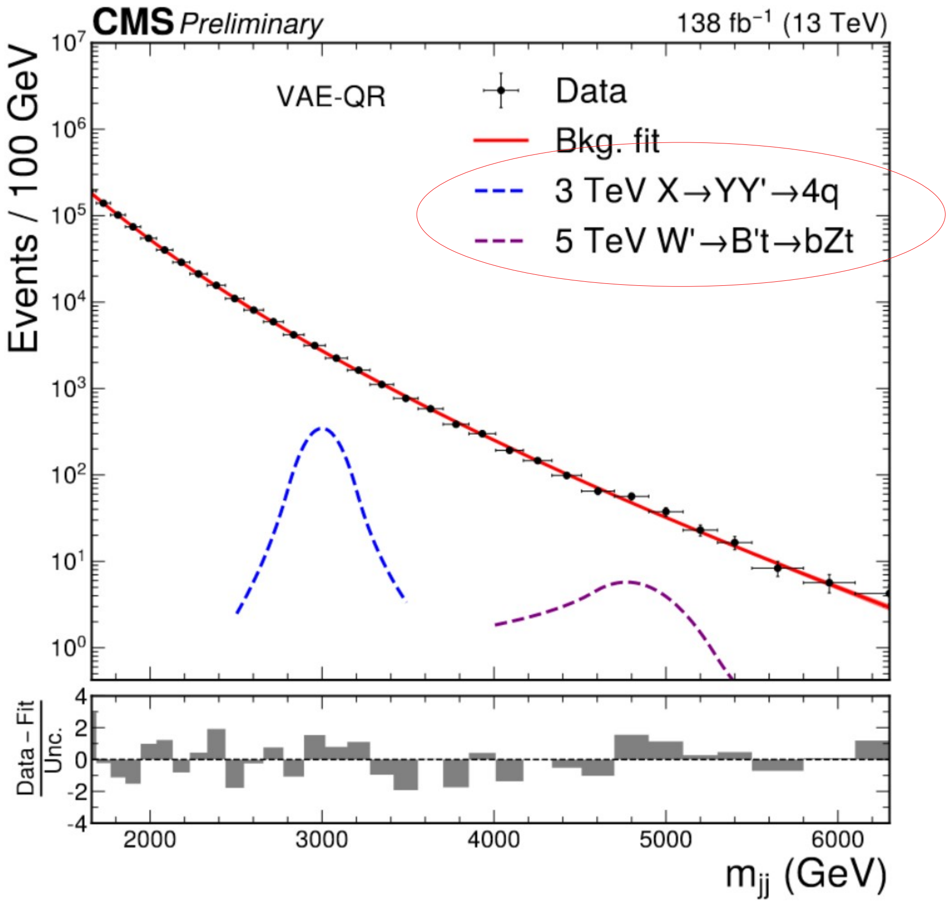
   (github_skCATHODE↗)

3. Lessons Learned!

Analysis Goal

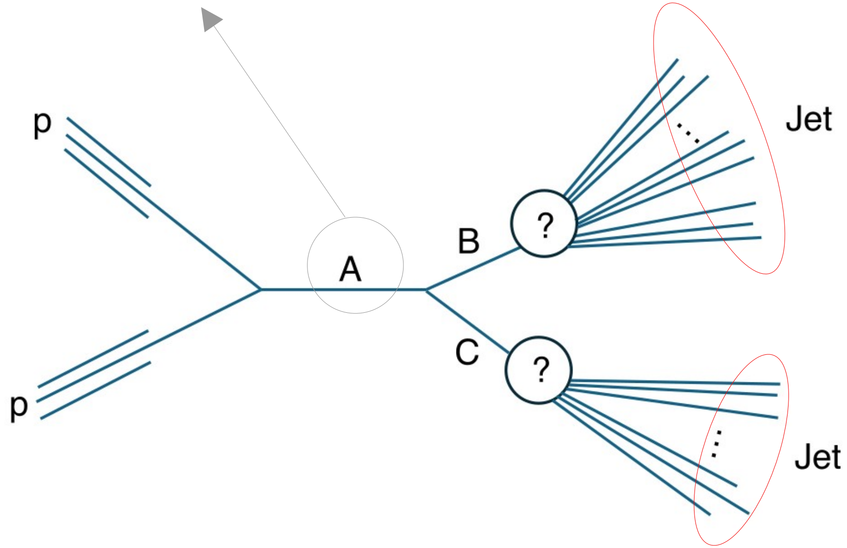
- look for resonances in the dijet mass spectrum

Signal shape shown for one of the methods



Signal Models

bump in the QCD background



$M_A = 2\text{-}5$ TeV

$M_B, M_C = 25\text{-}400$ GeV

Prongs	**A → BC**
(1+2)	$Q^* \to qW'$
(2+2)	$X \to YY' \to 4q$
(3+3)	$W' \to B't \to Bzt$
(2+4)	$W_{kk} \to RW \to 3W$
(5+5)	$Z' \to T'T' \to tZtZ$
(6+6)	$Y \to HH \to 4t$

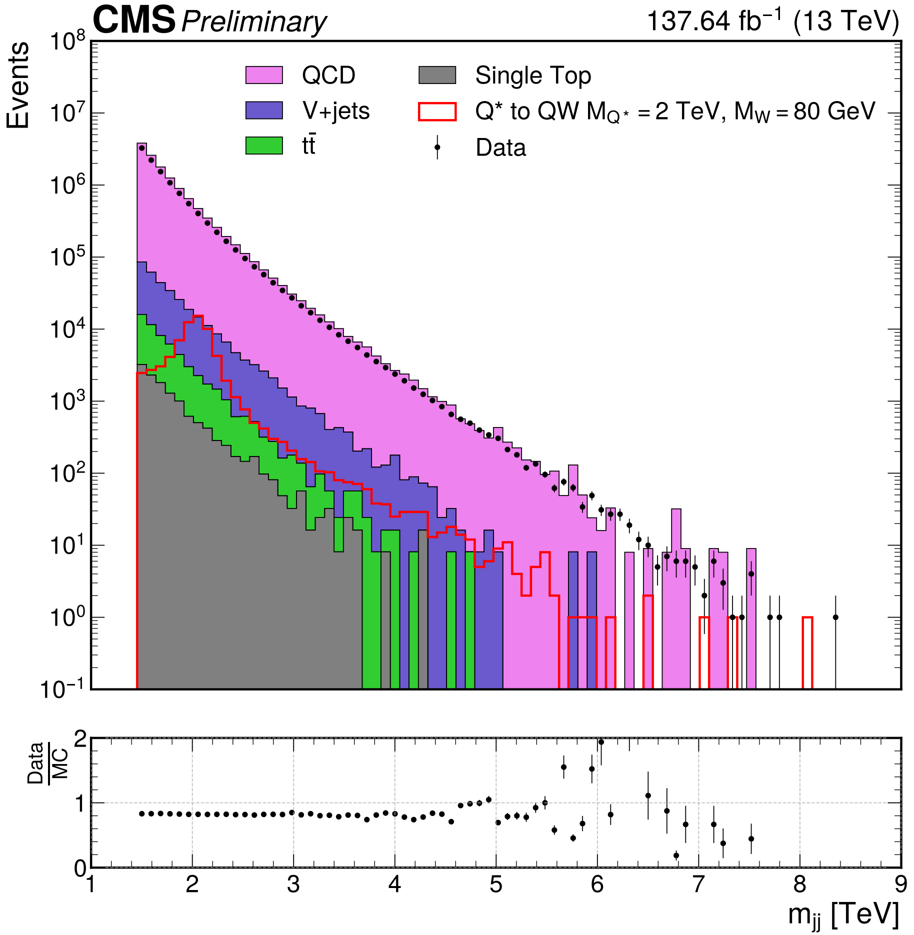
Datasets

Two datasets used in the analysis:

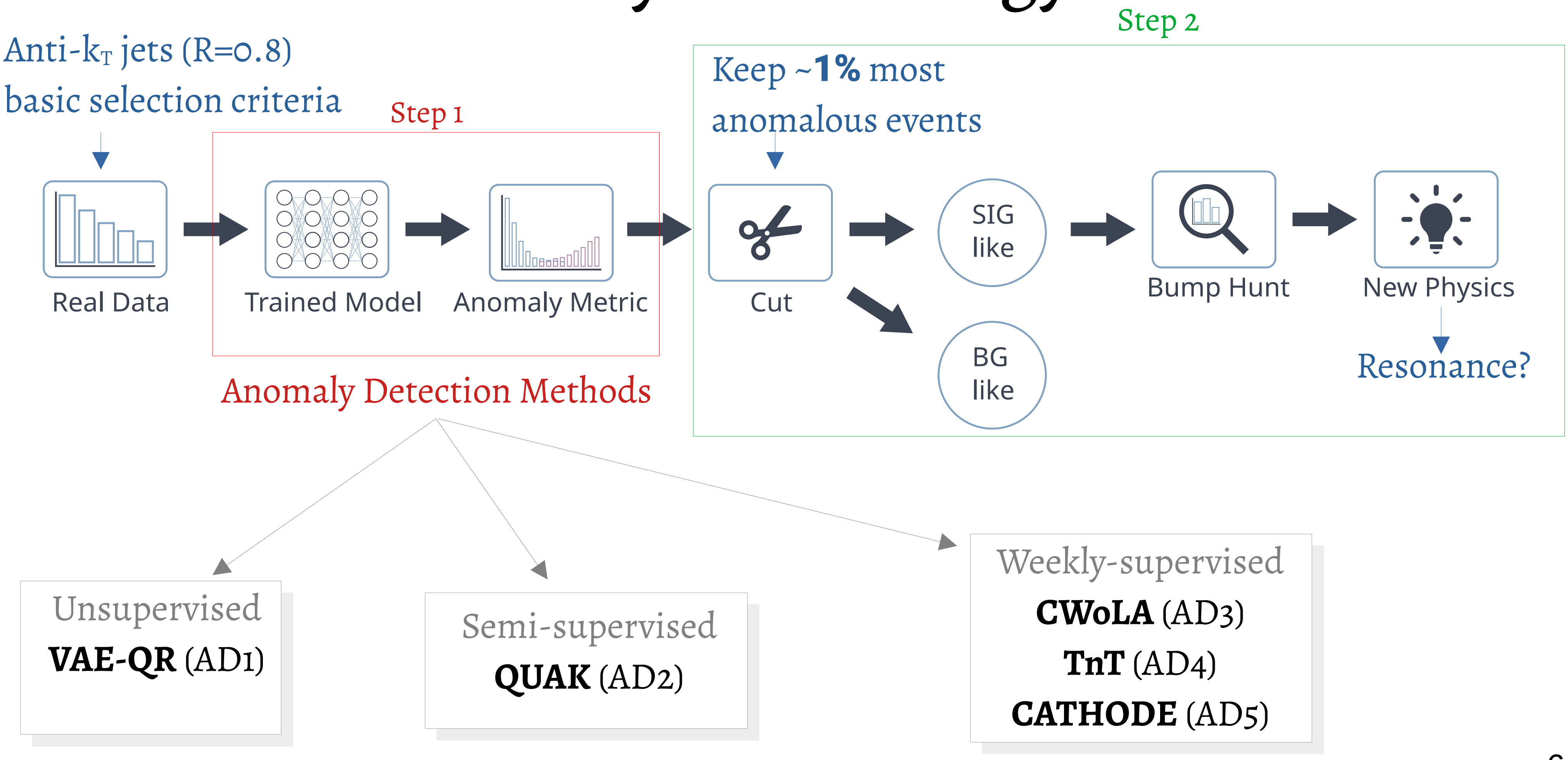
- MC “mock dataset”
  Unweighted sampling from QCD & minor backgrounds

- Full Run 2 data

Both selecting dijet ($R = 0.8$) events,
m_{jj} above trigger threshold



Analysis Strategy

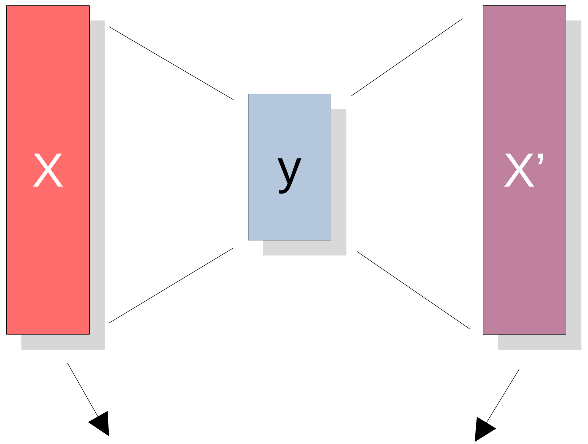


Unsupervised method

VAE -QR

(Variational Autoencoder-Quantile regression)

- Encodes up to 100 Particle Flow constituents per jet
- Trained with jets from a QCD-dominated sideband ($\Delta\eta > 1.4$)
- Background sculpting controlled with quantile regression



Final Anomaly score:

lowest reconstruction loss of the two jets

high – Anomalous!!

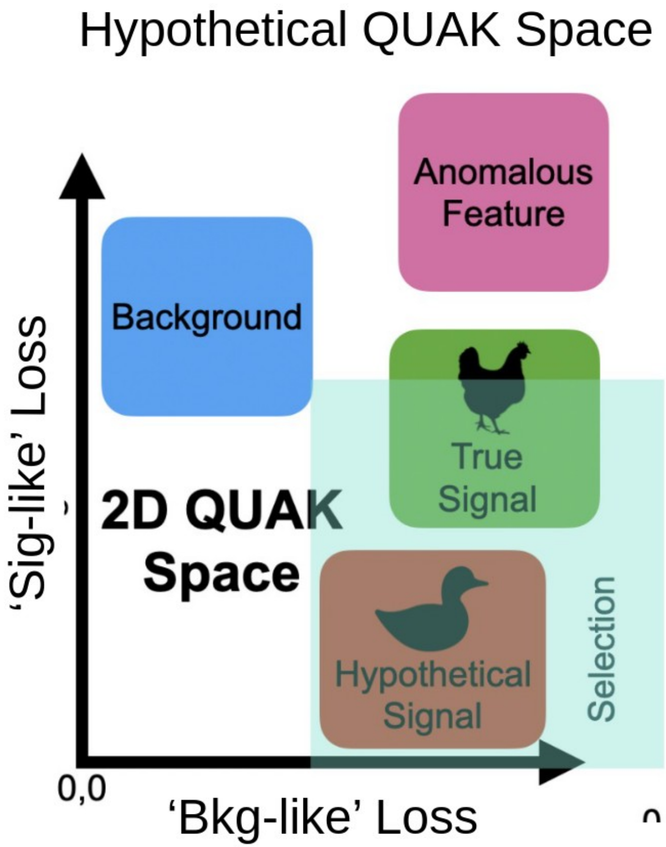
Semi-supervised method

QUAK
(Quasi Anomolous Knowledge)

- Hybrid approach, encoding a prior on signal-like features
- Train two normalizing flows:

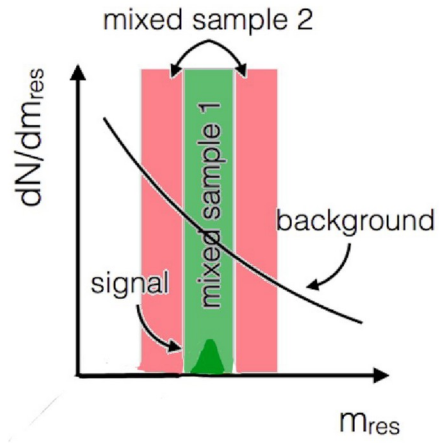
  -on a mixture of signal MCs

  -on background MC
- The losses define a 2D QUAK space
- The signal is somewhere in that space…



Weakly supervised methods

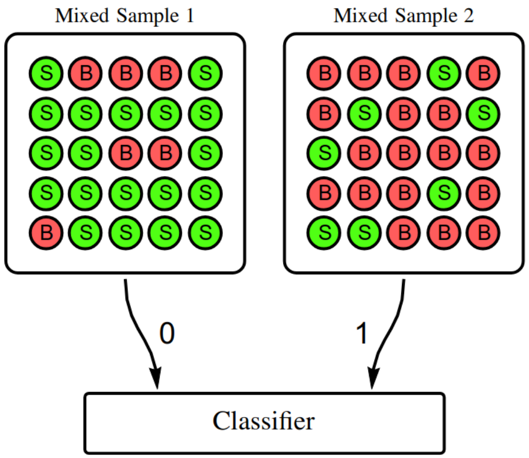
- Assume signal is a narrow resonance and choose a mass window that is defined as SR (signal region)
- Define the SB (side bands)
- Train a classifier between data and a **background-like sample**



$$\boxed{\frac{p_{\rm data}(\vec{x})}{p_{\rm bkg}(\vec{x})}} = \frac{(1-\epsilon)p_{\rm bkg}(\vec{x}) + \epsilon p_{\rm sig}(\vec{x})}{p_{\rm bkg}(\vec{x})} = 1 - \epsilon + \epsilon \boxed{\frac{p_{\rm sig}(\vec{x})}{p_{\rm bkg}(\vec{x})}}$$

data vs background

signal vs background



Weakly supervised methods

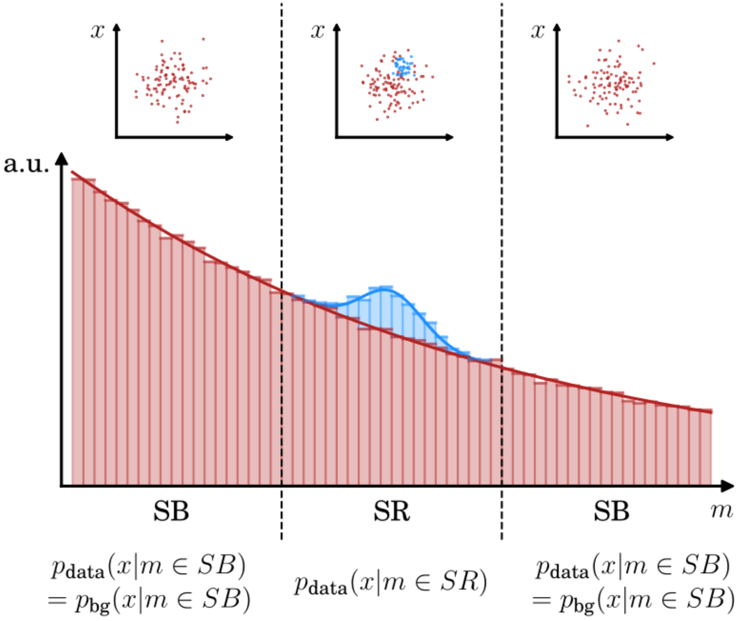
CWoLA (Classification WithOut LAbels)

- Background is taken from the side bands

TnT (Tag N' Train)

- Autoencoder preselection, targets events with **two** anomalous jets

CATHODE (Classifying Anomalies THrough Outer Density Estimation)



- background interpolated from sidebands
- Density estimation -> Masked Autoregressive Flow (MAF)
- Train it using the conditional feature, usually m and auxiliary features X only from the SBs

Input Variables

- CWoLa Hunting / TNT (one model per jet):

$$m_{\mathrm{sd}}, \quad \tau_{21}, \quad \tau_{32}, \quad \tau_{43}, \quad n_{\mathrm{PF}}, \quad \mathrm{LSF}_3, \quad \mathrm{DeepB}$$

- CATHODE (one model per event):

$$m_{j1}, \quad \Delta m_{j1j2} = m_{j1} - m_{j2}, \quad \tau_{41,j1}, \quad \tau_{41,j2},$$

- CATHODE-b (one model per event):

$$m_{j1}, \quad \Delta m_{j1j2} = m_{j1} - m_{j2}, \quad \tau_{41,j1}, \quad \tau_{41,j2}, \quad \mathrm{DeepB}_{j1}, \quad \mathrm{DeepB}_{j2}.$$

- VAE-QR (one model for both jets):

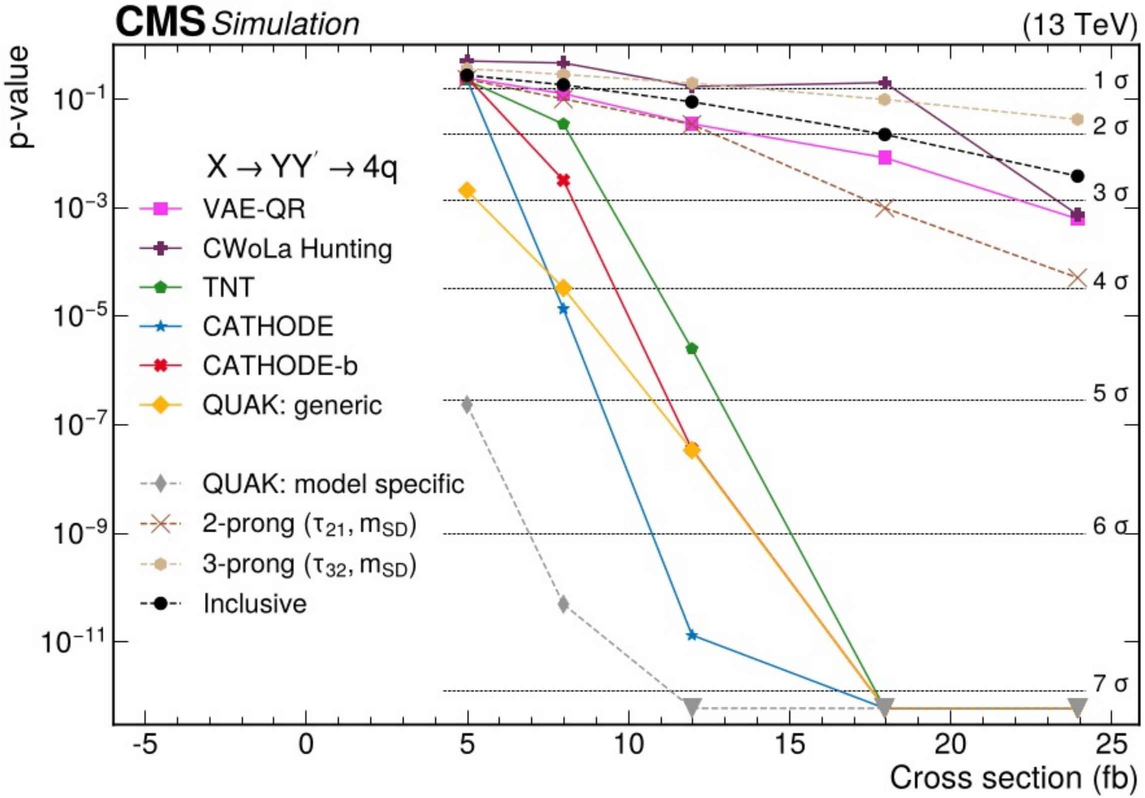
  Jet images

Performance on simulation

- Tested for process

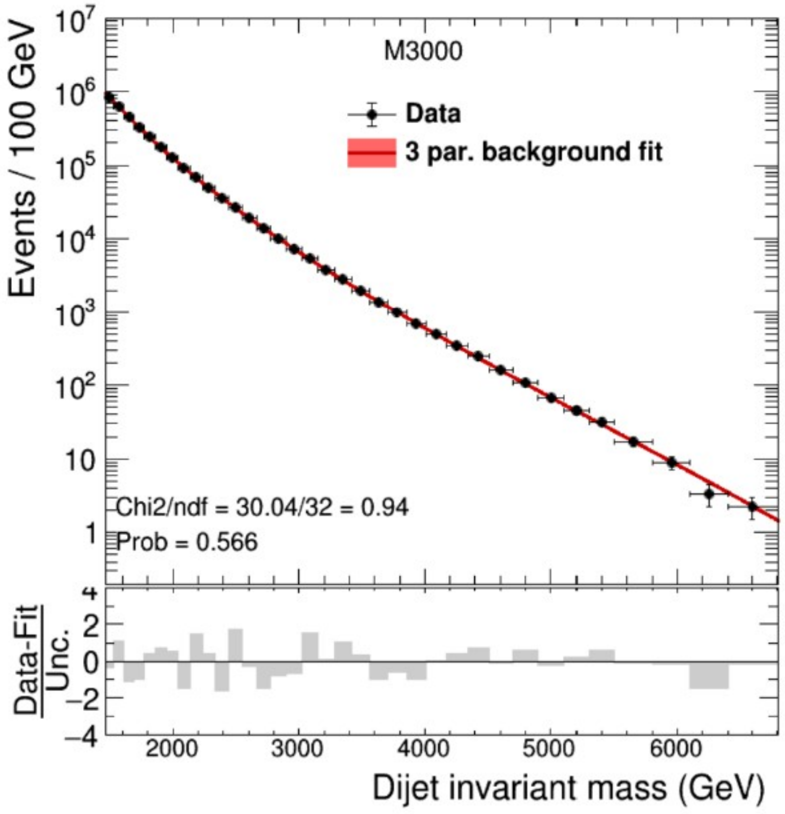
  X(3000) → YY' → qq qq

- Signal detection already possible for low cross sections



Getting Results

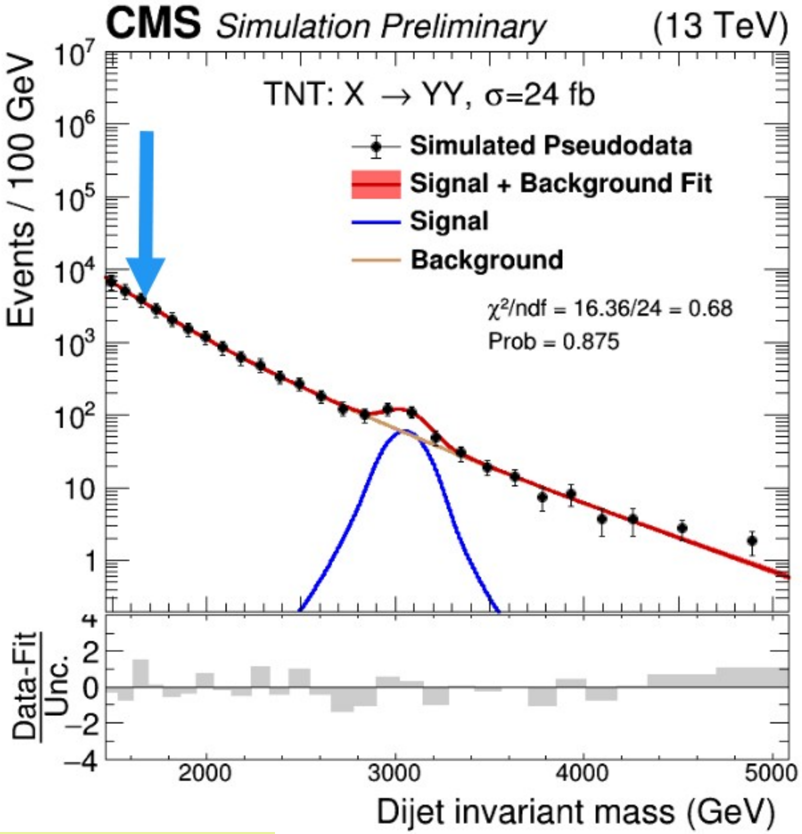
- Select the ~1% most anomalous events



MC mock dataset

Getting Results

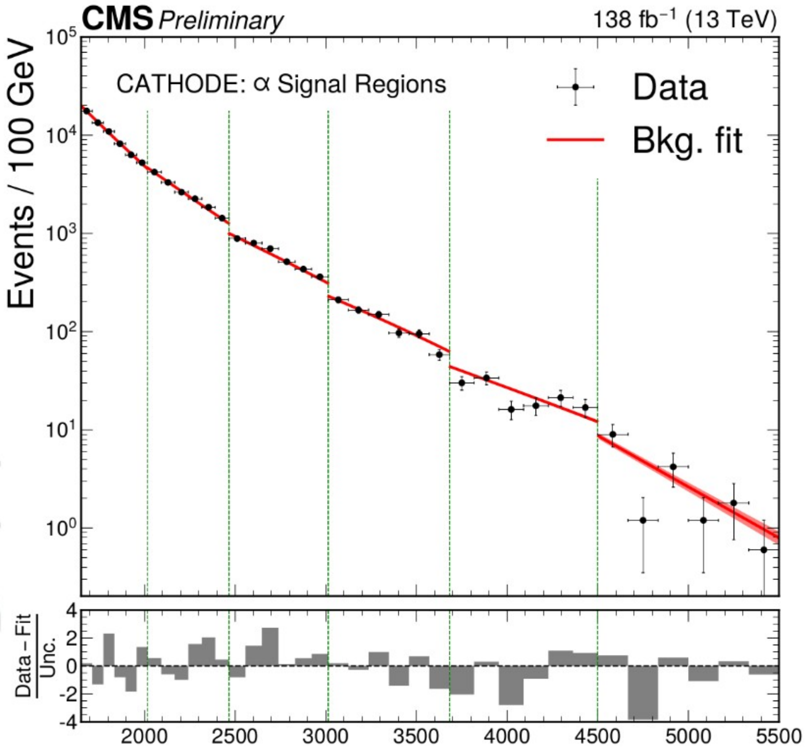
- Select the ~1% most anomalous events
- Choose a working point, select events
- Fit dijet mass spectrum with analytic function
- Perform Bump hunt
- Derive p-value



MC mock dataset

Getting Results

- Select the ~1% most anomalous events
- Choose a working point, select events
- Fit dijet mass spectrum with analytic function
- Perform Bump hunt
- Derive p-value
- Apply method to complete mass range, redefine SRs
- Scan over the dijet invariant mass



Full Run 2 data

Limits

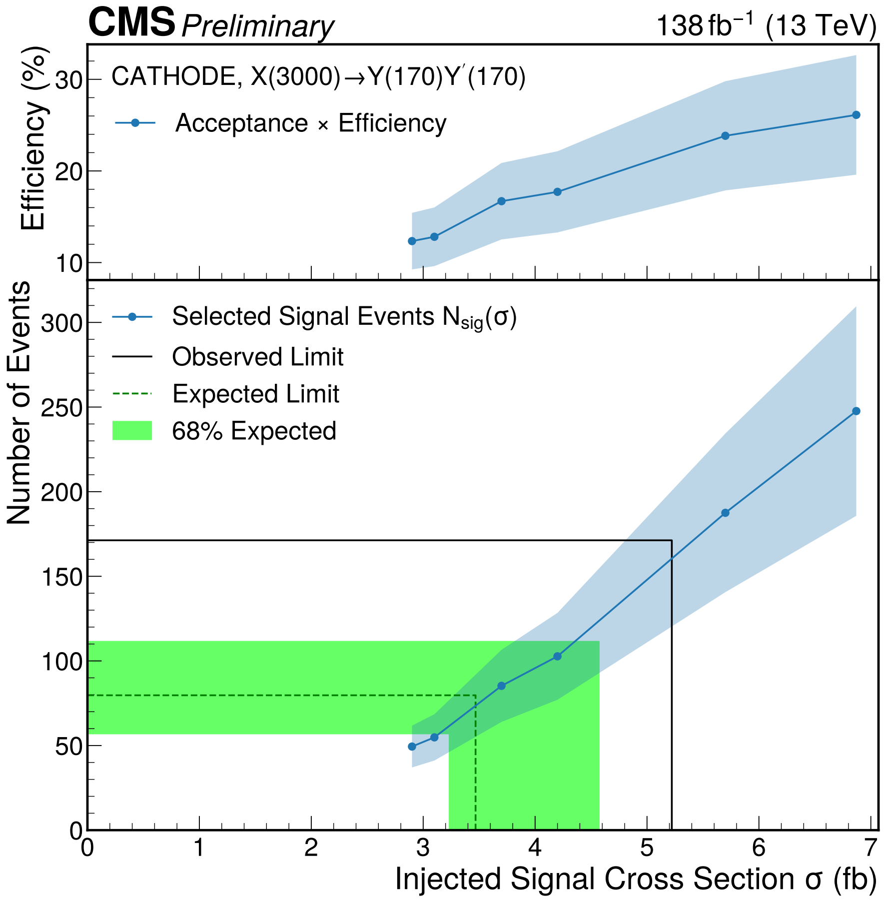
For weakly supervised methods:

- The signal efficiency depends on the number of signal events

$$N_{\rm sig}(\sigma) = \sigma \times \mathcal{L} \times A \times \epsilon(\sigma)$$

- Special limit-setting procedure
- Explained in Appendix B

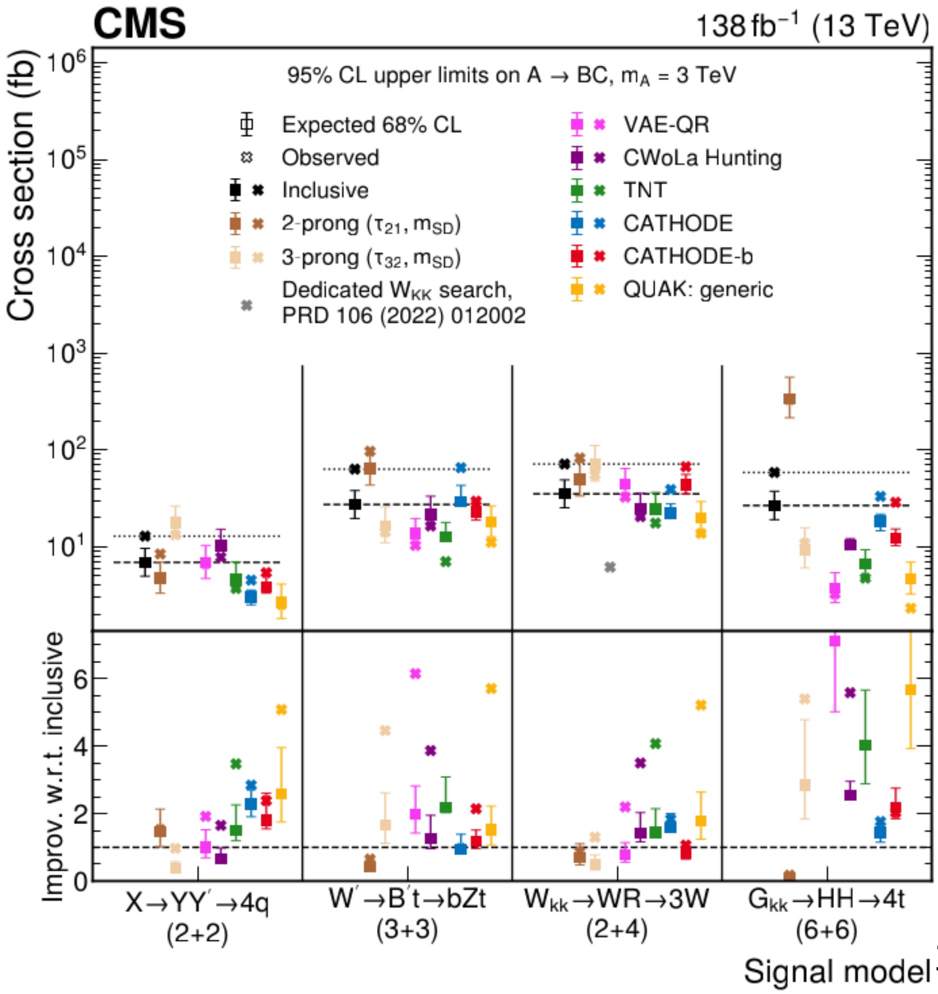
Appendix B, EXO-22-026↗



Full Run 2 data

Results

- No anomaly detected
- Improved limits compared to inclusive search

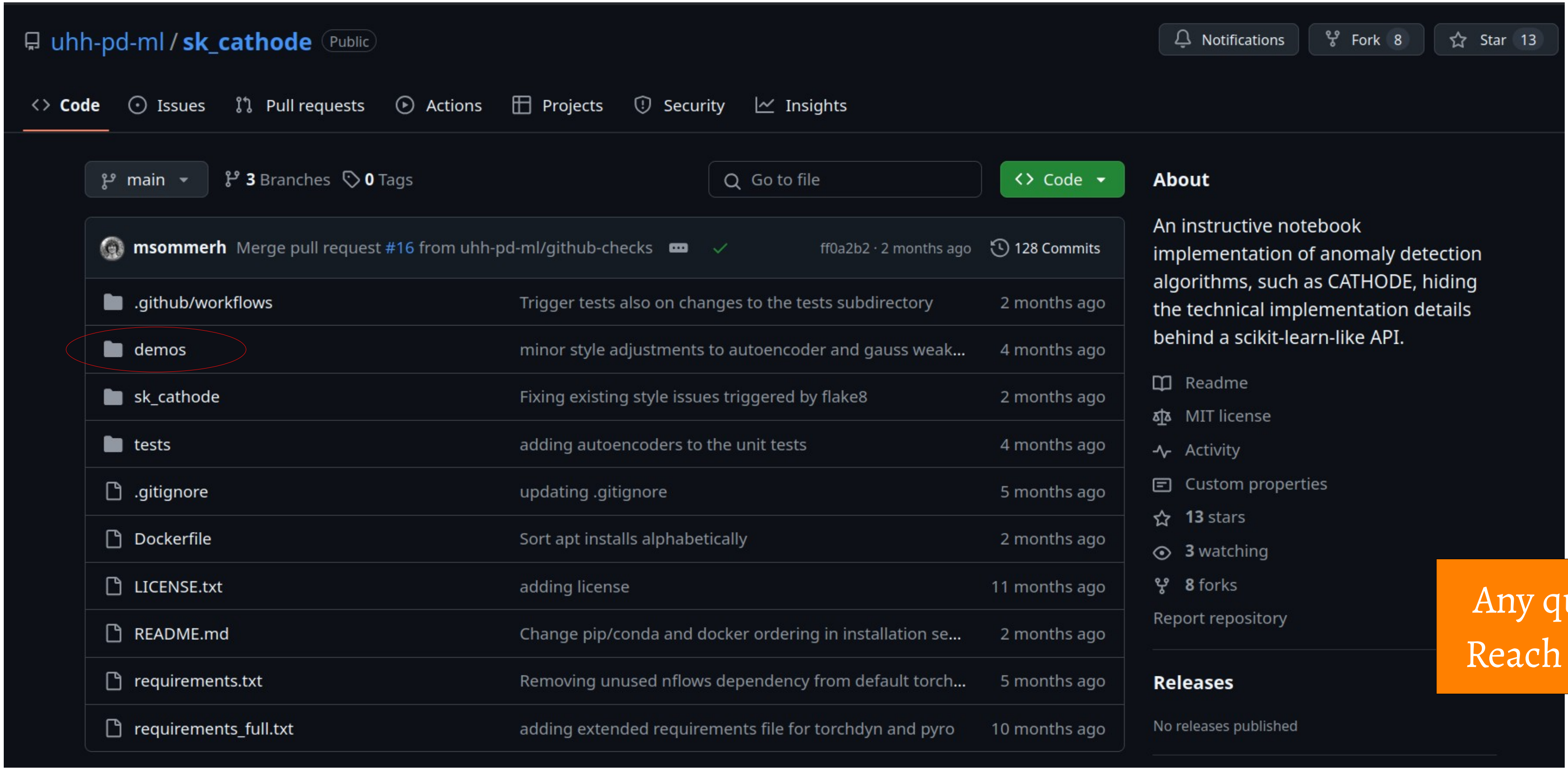


skCATHODE

Check it out

(github:skCATHODE↗)

This code base provides easy-to use modules for building and testing CATHODE



Any questions?
Reach out to us!

Lessons Learned!

- Five different methods → complimentary
- Limit setting is unconventional (for weak supervision)
- …

↓

Machine Learning Paper

Performance for all signals

- Best limit improvement for all 22 signals (3 TeV)

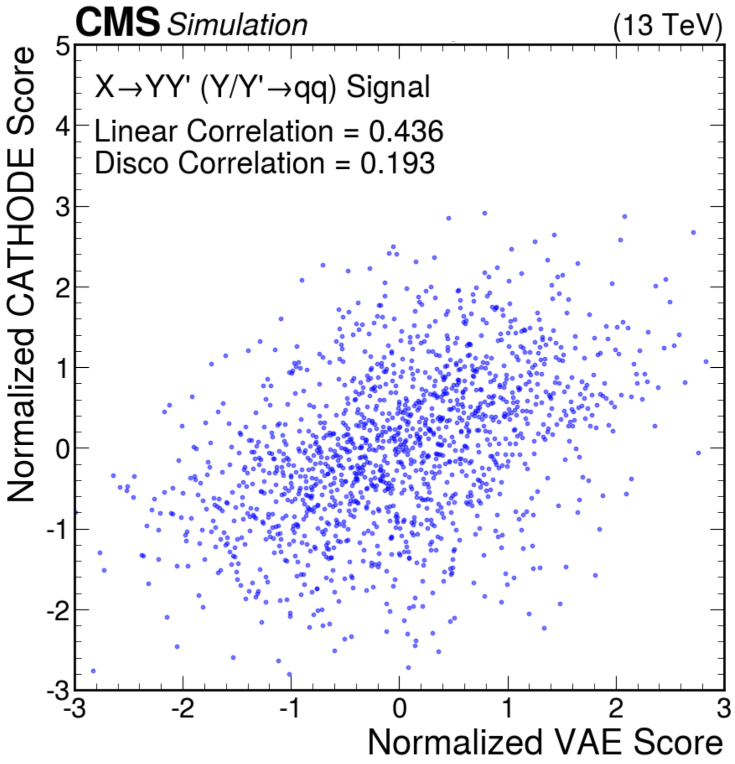
- All methods represented

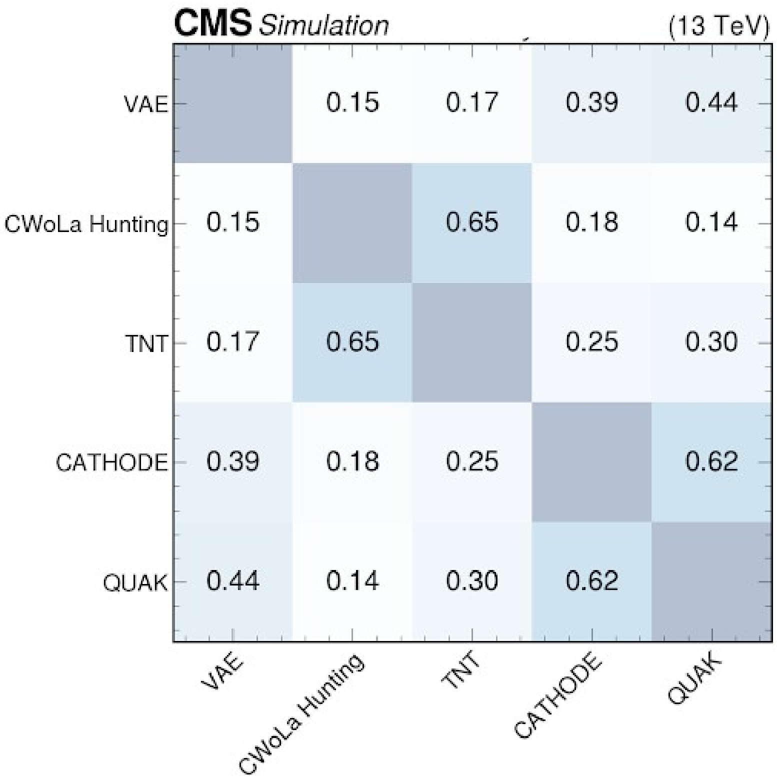
  ➔ Complementarity

Model $A \to BC$ (m_A = 3 TeV)	Daughter Masses (GeV)	Method	Exp. (Obs.) Limit (fb)	Improvement wrt Inclusive
$Q* \to qW'$	25	*CWoLa Hunting*	61.1 (30.1)	0.3
	80	*CATHODE*	50.0 (95.2)	0.4
	170	*VAE-QR*	52.5 (37.5)	0.4
	400	*CWoLa Hunting*	45.8 (24.3)	0.5
$X \to YY' \to 4q$	25/25	*CATHODE*	8.0 (9.9)	0.9
	25/80	*CATHODE*	7.6 (13.2)	0.9
	25/170	*CATHODE*	10.3 (18.4)	0.7
	25/400	*VAE-QR*	13.6 (12.5)	0.6
	80/80	*CATHODE*	4.2 (8.0)	1.6
	80/170	*CATHODE*	5.7 (11.4)	1.2
	80/400	*CATHODE*	6.0 (7.3)	1.2
	170/170	*CATHODE*	3.7 (6.8)	1.9
	170/400	*VAE-QR*	4.4 (4.0)	1.7
	400/400	*VAE-QR*	2.1 (1.9)	4.2
$W' \to B't \to bZt$	25	*TNT*	25.2 (17.4)	1.5
	80	*TNT*	22.3 (14.6)	1.5
	170	*TNT*	12.2 (7.3)	2.1
	400	*VAE-QR*	15.2 (11.4)	1.8
$W_{KK} \to RW \to 3W$	170	*TNT*	25.1 (20.1)	1.4
	400	*CWoLa Hunting*	23.8 (25.0)	1.5
$Z' \to T'T' \to tZtZ$	400	*QUAK*	28.3 (13.9)	2.7
$Y \to HH \to 4t$	400	*QUAK*	7.7 (3.7)	3.5

Complementarity

- Small correlations between anomaly scores ➔ Complementarity

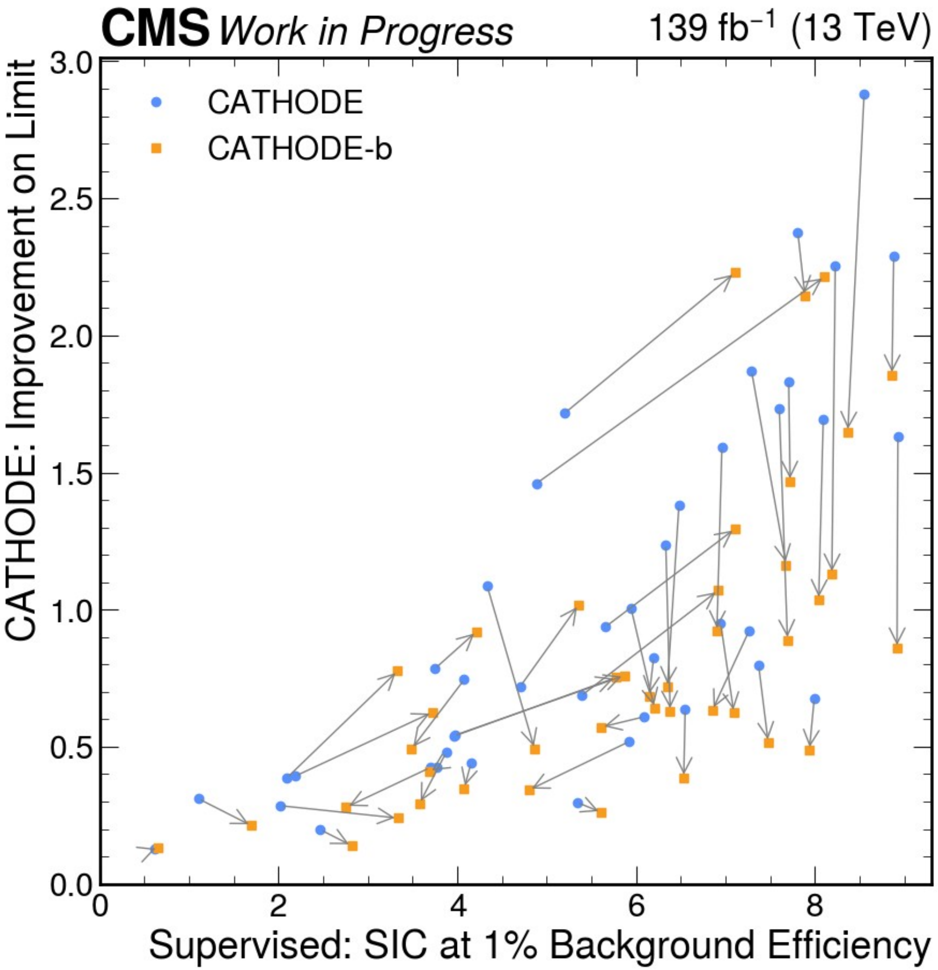




MC mock dataset – CMS Note

Cathode Vs. Cathode-b

- Testing more signals with CATHODE(-b)
- Mixing 3 and 5 TeV mass points
- Using the supervised classifier score as a proxy for input strength



- CATHODE (one model per event):

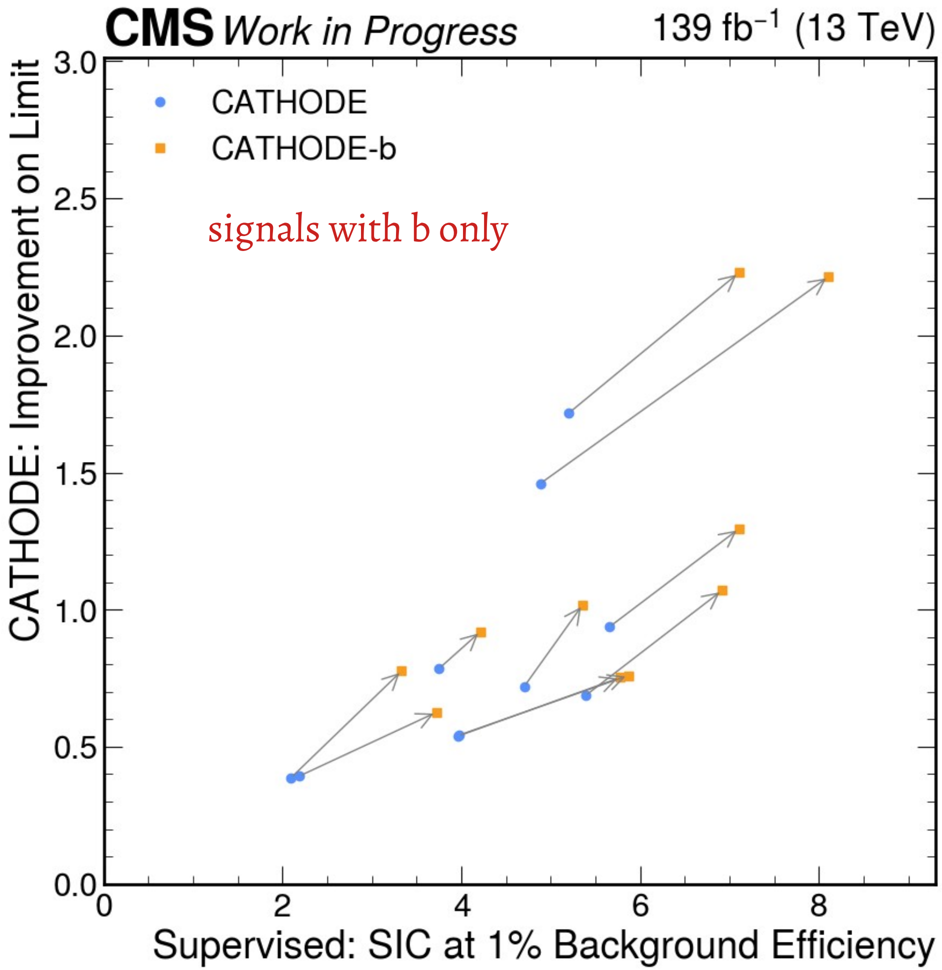
$$m_{j1},\quad \Delta m_{j1j2} = m_{j1} - m_{j2},\quad \tau_{41,j1},\quad \tau_{41,j2},$$

- CATHODE-b (one model per event):

$$m_{j1},\quad \Delta m_{j1j2} = m_{j1} - m_{j2},\quad \tau_{41,j1},\quad \tau_{41,j2},\quad \mathrm{DeepB}_{j1},\quad \mathrm{DeepB}_{j2}.$$

Cathode Vs. Cathode-b

- Testing more signals with CATHODE(-b)
- Mixing 3 and 5 TeV mass points
- CATHODE-b better on b-tagged signals, worse on others



Summary

- We presented the first CMS results on anomaly detection
- There are 5 different methods that improves performance as compared to the inclusive search
- The five methods are complimentary, there's no single method performing the best. But different methods do well with different signal.
- We are investigating about these complementarity and differences in the methods, currently observing different input features
- We will do more studies, we are happy to hear suggestions !

CMS

Compact Muon Solenoid