# CURTAINs for your sliding window

Johnny Raine, Sam Klein, Debajyoti Sengupta*, Tobias Golling.
University of Geneva

Dark Interactions: New Perspectives from Theory and Experiment
November 16th, 2022

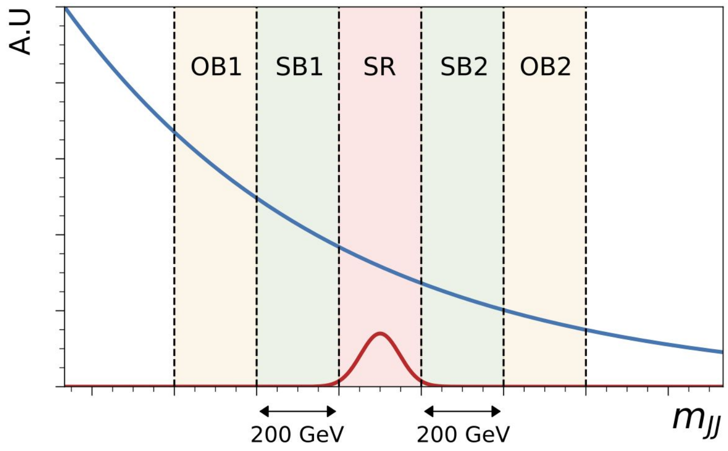
*debajyoti.sengupta@cern.ch

# A ‘standard’ bump hunt

Expected New physics ~ localised in smoothly falling feature
⇒ Bump in the spectrum

Method:

- Split spectrum into sliding ‘side bands’
- Fit the distribution in sidebands → Interpolate into the signal region
- Look for an excess
- Slide window and repeat



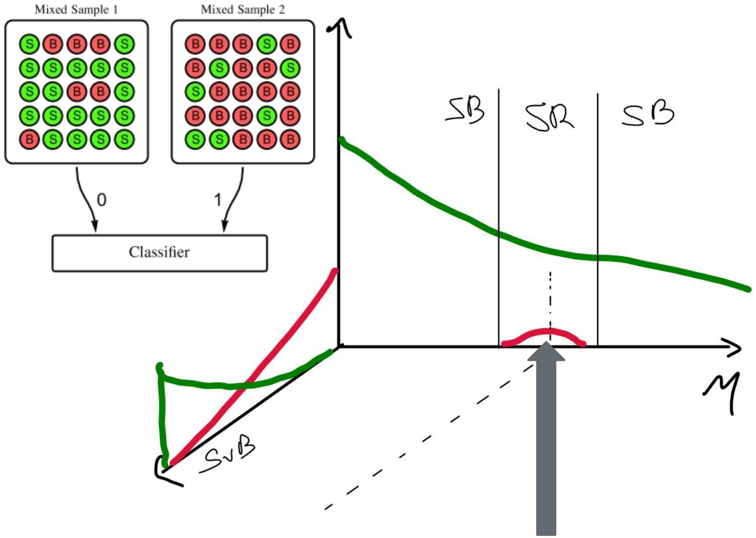

Limitations: New Physics is rare → Bump ~ background dominated.

- Utilise information other than just the resonant observable.

# Extending Bump hunts with CWoLa

- Identify features of interest that help Signal vs Background discrimination

- Train a classifier on SideBand vs Signal Region - Equivalent to Signal vs Background - The CWoLa limit
    - Collins et. al.




Expected to be very small

Problem: ***What if the features are correlated with the resonant feature?***

# CURTAINs – Constructing Unobserved Regions by Transforming Adjacent Intervals

Learn a conditional transformation of datapoints

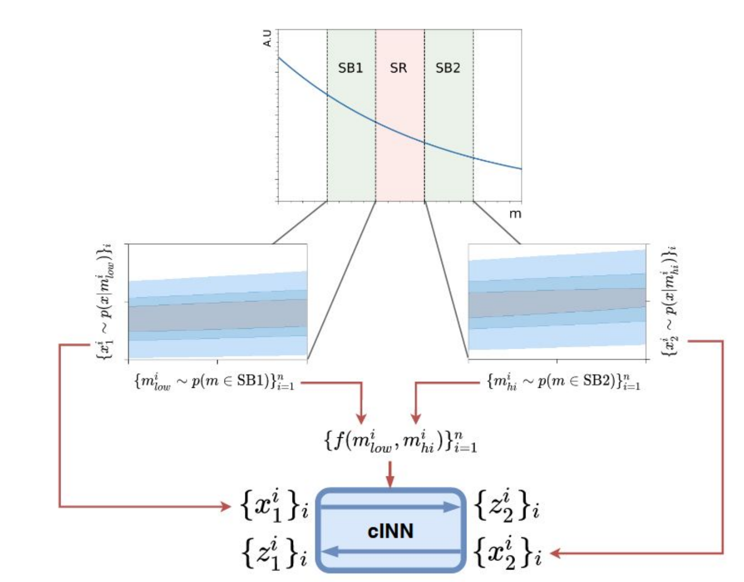
- *What would a datapoint look like if it had a different value of mass?*

Train an Invertible Neural Network (INN) to map between sidebands:

- Condition on (a function of) input and target mass.

Post training - Choose target mass values in SR → Transport sideband data to Signal Region

⇒ Signal region template made of conditionally transformed features!



# Training:

Draw data *x* from SB1 and SB2

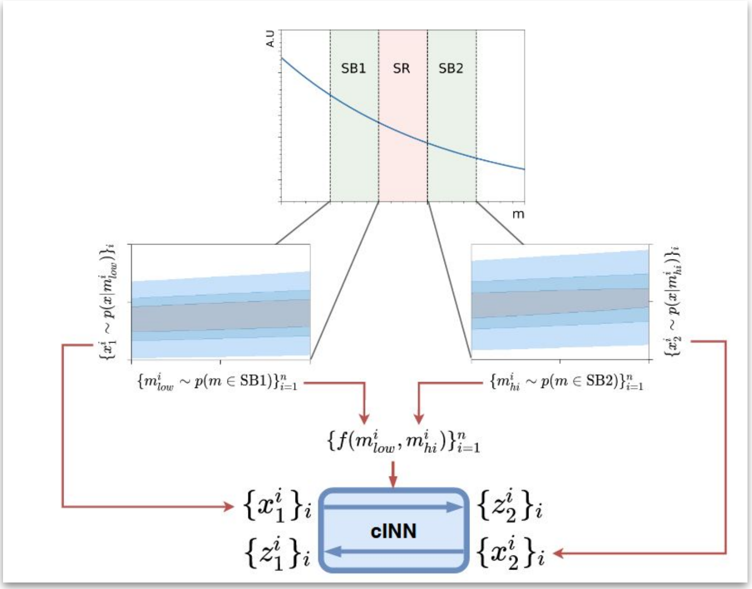
Assign target masses based on batch

Transport data from SB1<->SB2

- Can **train bidirectionally!**

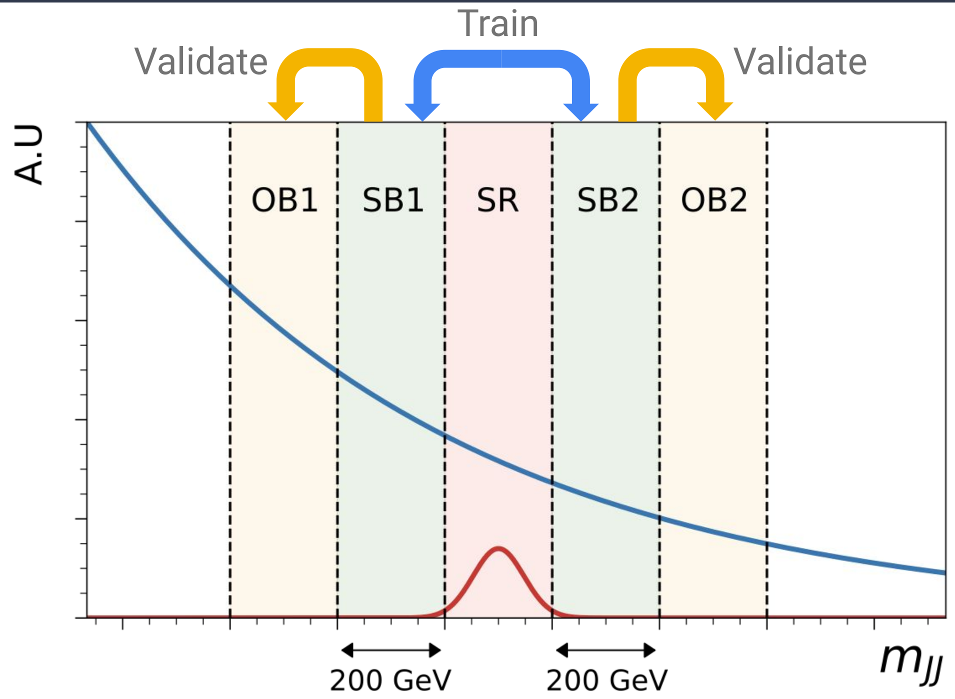
Measure the sinkhorn distance  between~  $z_2$ and $x_2$ and $z_1$ and $x_1$

Learned transformation ~ good approximation for true optimal transport.



# Validating Transformations

- Fix sidebands
- Define Outer-Band (OB) validation regions
- Train CURTAINs transformer
- Validate on OBs



# Results from Validation

CURTAINs applied to a dijet system: Training on the [LHCO R&D anomaly detection dataset]

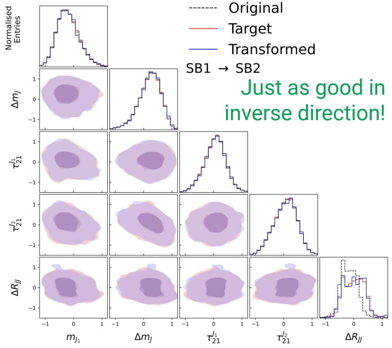
- Sideband 1: [3200, 3400] GeV
- Sideband 2: [3600, 3800] GeV

Five observables - $M_{J1}$, $M_{J1}$- $M_{J2}$, $\tau^{21}{}_{J1}$, $\tau^{21}{}_{J2}$, $\Delta R_{JJ}$ (due to correlation to Mjj)

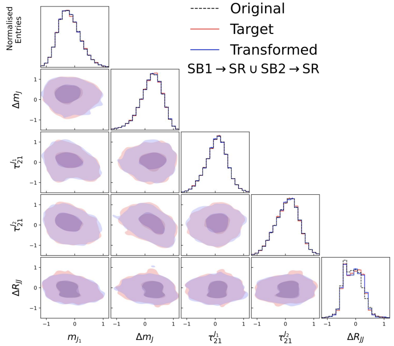
Excellent agreement across all features!

Compare SR template and SR data (not possible in analyses):

- Excellent agreement → Near perfect background template!




Just as good in inverse direction!



# Comparisons

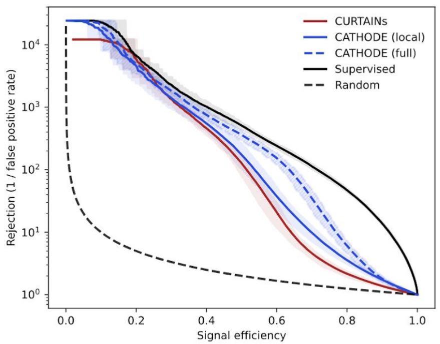
1. CATHODE: [2109.00546] ~ Learn the conditional density in the sidebands - interpolate and generate samples in the Signal region.
    a. Full - use everything other than SR as sidebands.
    b. Local - same window definition as CURTAINs.

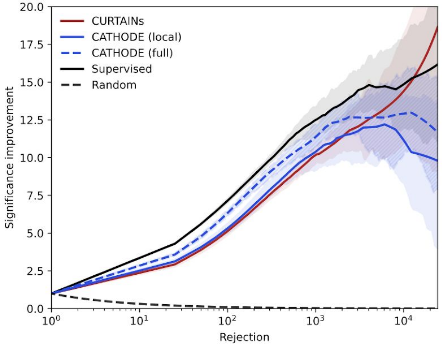
Benchmarks:

a. Supervised: Train a classifier on labelled nominal vs labelled signal.
b. Idealised: Train a classifier on data vs a perfect background simulation.

# Anomaly Detection: Performance

Results for SR Window centralised on and containing the Signal peak:

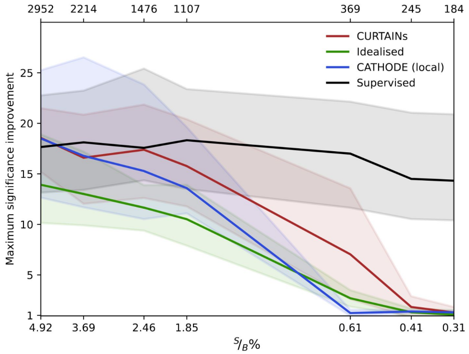



Measure Significance Improvement (SI) = $\dfrac{\text{signal efficiency}}{(1 - \text{background efficiency})^{0.5}}$

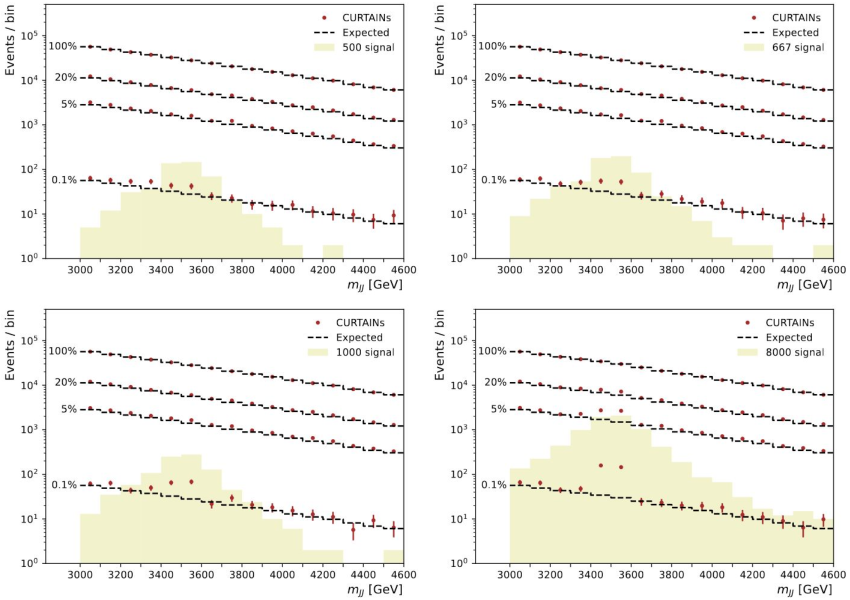
Measure the maximum SI as a function of injected signal.

CURTAINs outperforms CATHODE (local) at low signal fractions.



# Anomaly Detection: The Bump Hunt

- Define SB-SR-SB → Train CURTAINs and construct template in SR.
- Train a CWoLa style classifier on template vs SR data.
- Slide window and repeat.




Bump hunt performance as a function of injected signal → more signal ⇒ prominent bump!

# Flows4Flows

Use normalizing flows to parametrise base distribution! No more approximate OT loss!

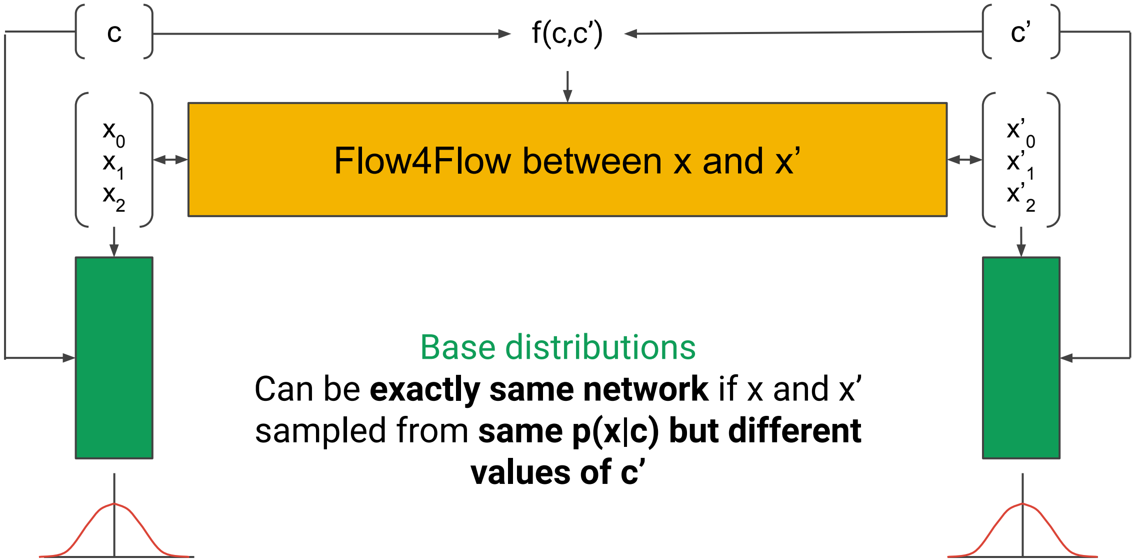
**Flows4flows**

Train a flow between arbitrary distributions

Simply another change of variables for p(z) in normalizing flows!

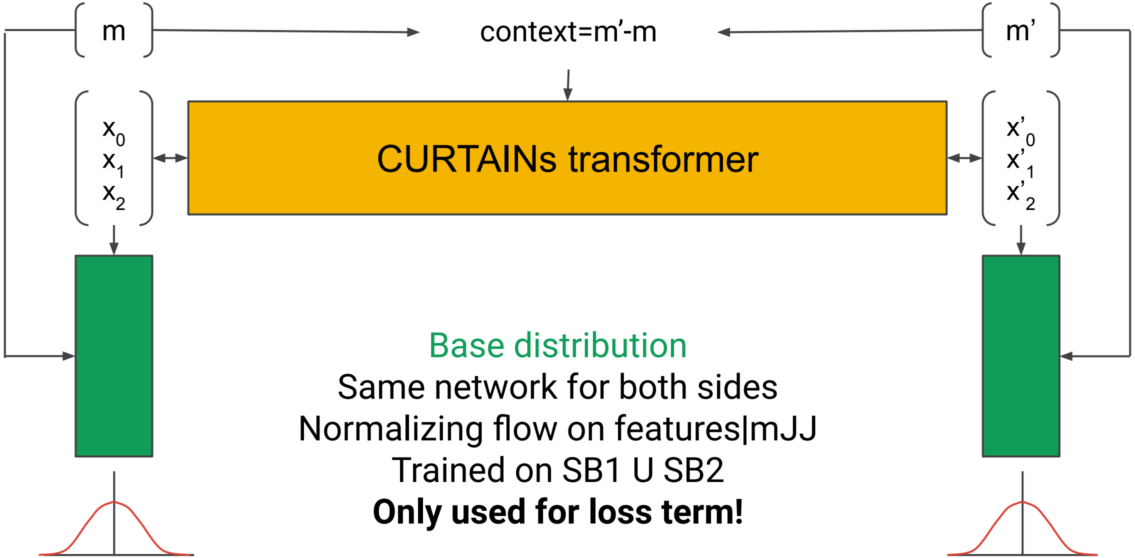
Pretrain base distribution(s)

Use base distribution for loss in exact maximum likelihood




**Can now train CURTAINs with exact maximum likelihood!**

# CURTAINs: Flows4Flows




*During training:*
*Randomly assign target mass by shuffling input values in batch*

*For application use SR mass values*

*Flow direction depends on if target > input mass*

Loss: $\log(p(x)) = \log\det|J(f(x|m,m'))| + \log\det|J(g(f(x|m,m')|m') + p(g(f(x|m,m')|m'))$

# Comparisons with CURTAINs v1

**Significant improvement** with new loss!

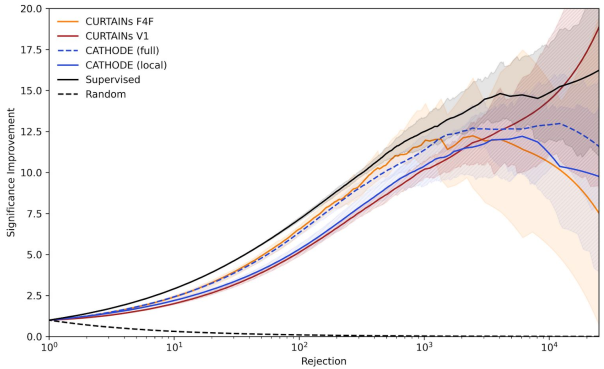
Much faster to train (factor 5-6), *including* base density

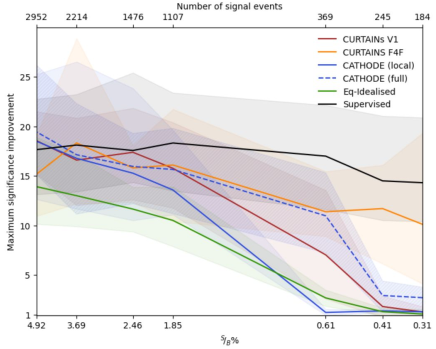
Still trained on a very local window

- Only 200GeV either side of SR

Compared to CURTAINs v1

- **Simpler** to set up and train
- Higher max SI at low signal fractions!

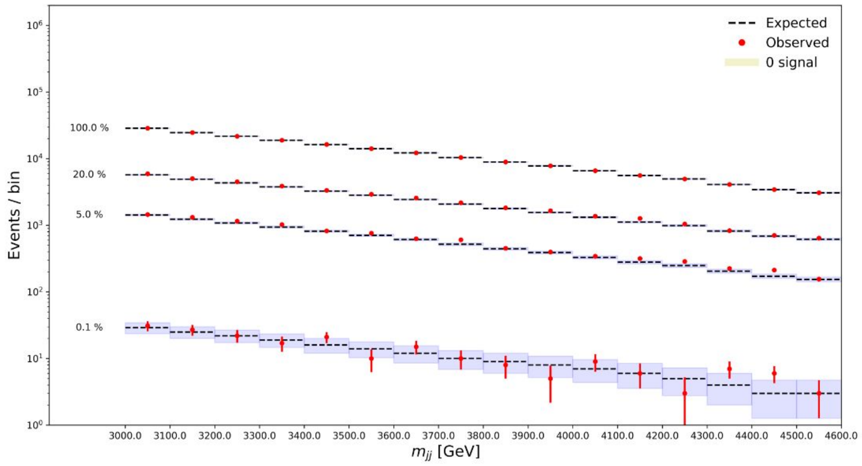


# Bump Hunt with Flows4Flows

Compared to CURTAINs v1

- **Simpler** to set up and train
- Still **robust** to case where there is no signal



# Concluding remarks

- CURTAINs - powerful and robust technique - can be applied to any 1D resonance searches!

- Improved: CURTAINs Flows4Flows → much faster and performant out of the box!
    - Pre-print in works.

# Backup

# Process and stats details:

- Background: dijet QCD
- Signal: W’ → X ( → qq) Y ( → qq);
  - mW’ = 3.5 TeV, mX = 500 GeV, mY = 100 GeV
- Jets reconstructed with R=1.0 antikT algo.
- pT requirements: Atleast one jet > 1.2 TeV.

- 1 million QCD dijet events.
- A total of 100,000 signal events.

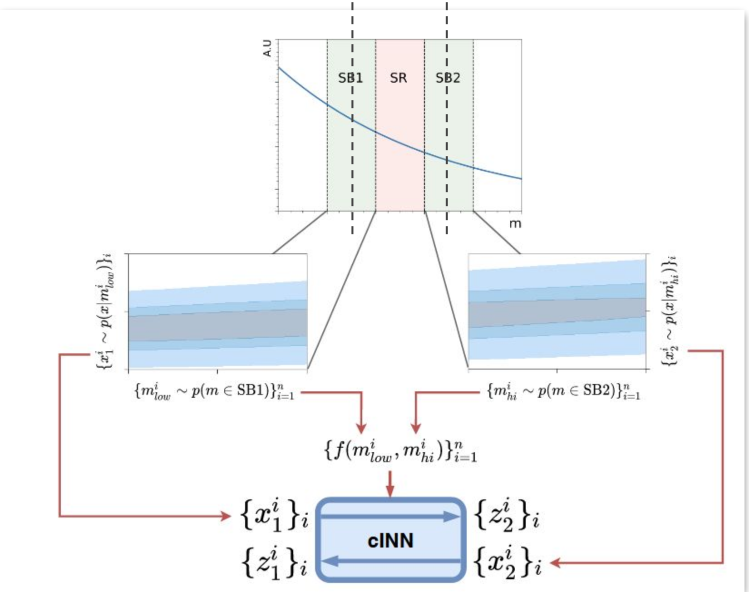
# Conditional Transformations: Technicality

Using f(m,m')=m'-m

- During training **min value is width of SR**
- To transport data from SB->SR **min value is 0**
- Outside of training domain...

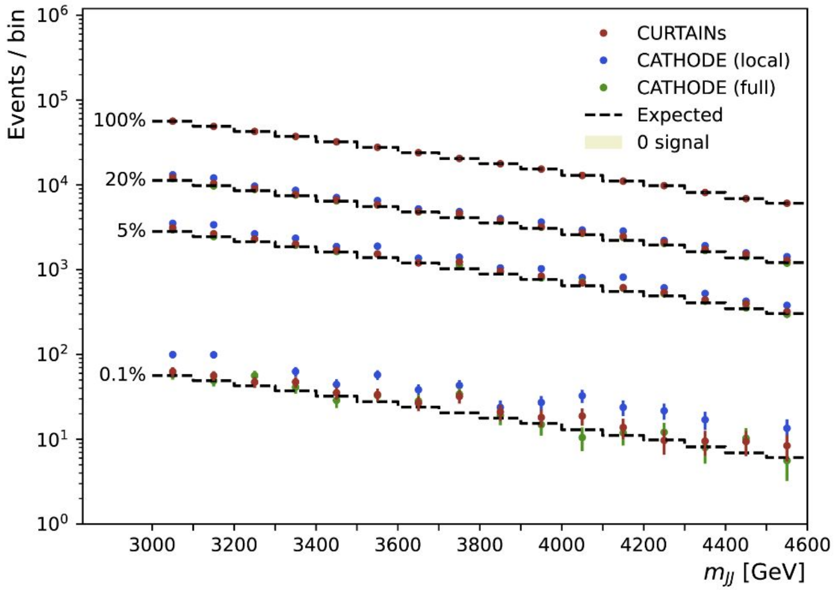
Solution: Split sidebands into two

- Train between lower and upper half of SB1/2
- Now min value is also 0



# V1 CURTAINs (+CATHODE) bump hunt

- Robust when there are no signal events present.



# Flows

Flow = Base Density + Invertible Function

$$\log p_{\theta,\phi}(x) = \log p_\theta(f_\phi^{-1}(x)) - \log \left|\det(J_{f_\phi^{-1}(x)})\right|$$

data distribution $p_D(x)$

base density $p_\theta$

# CURTAINs on OuterBands

- Outer Band agreement, just as good.