BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

Machine Learning-Driven Anomaly Detection in Dijet Events with ATLAS

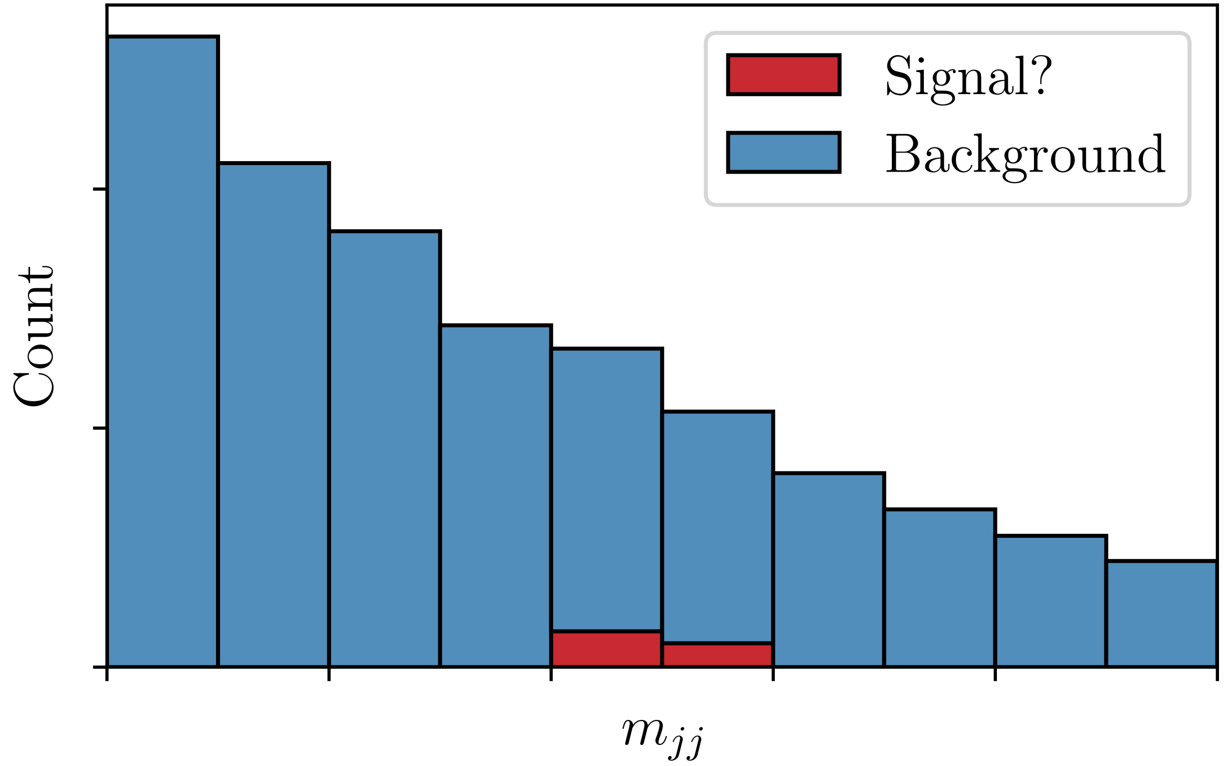
based on 2502.09770 (submitted to PRD)

Dennis Noll

June 16, 2025

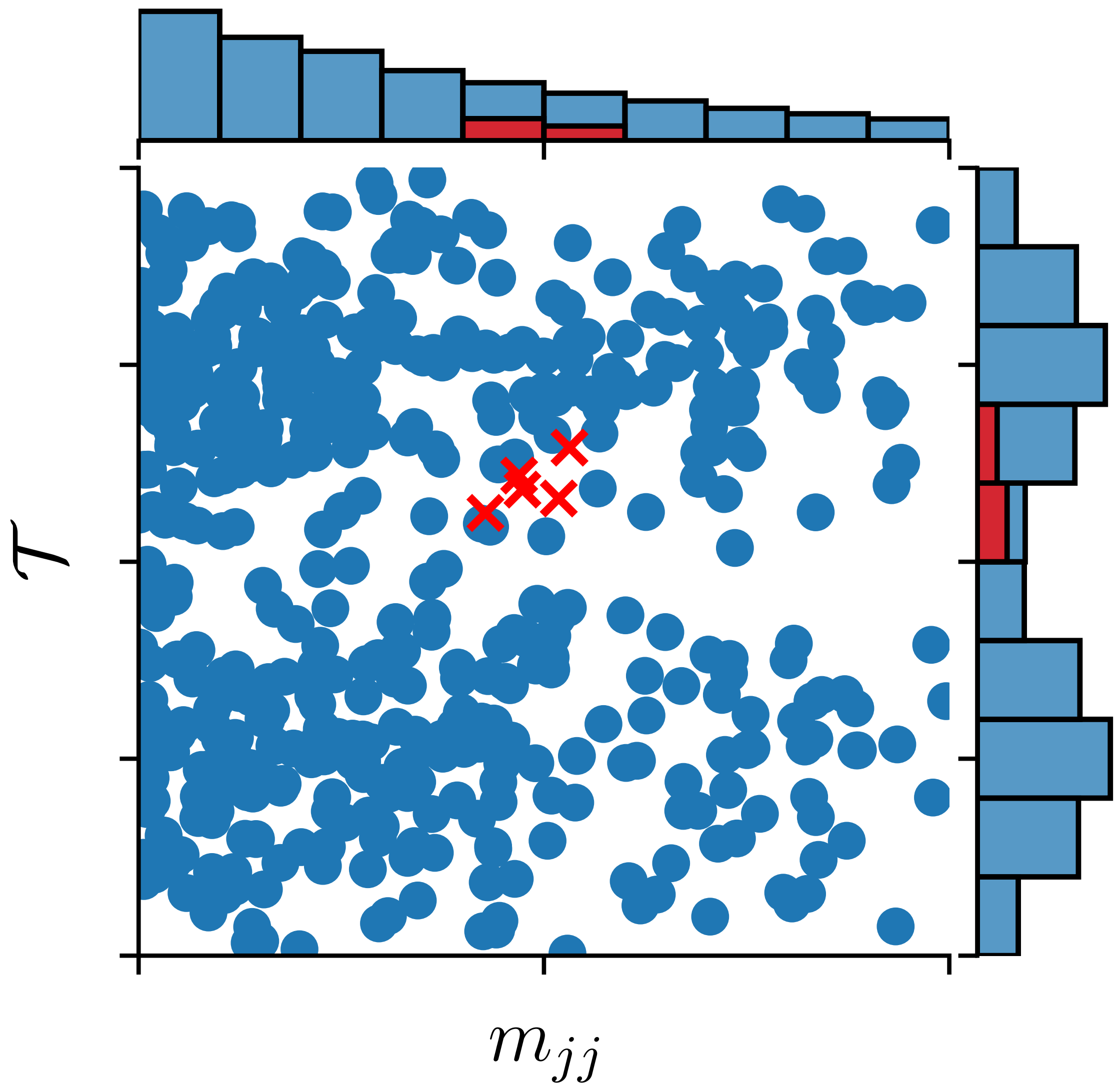
AD4HEP

Motivation



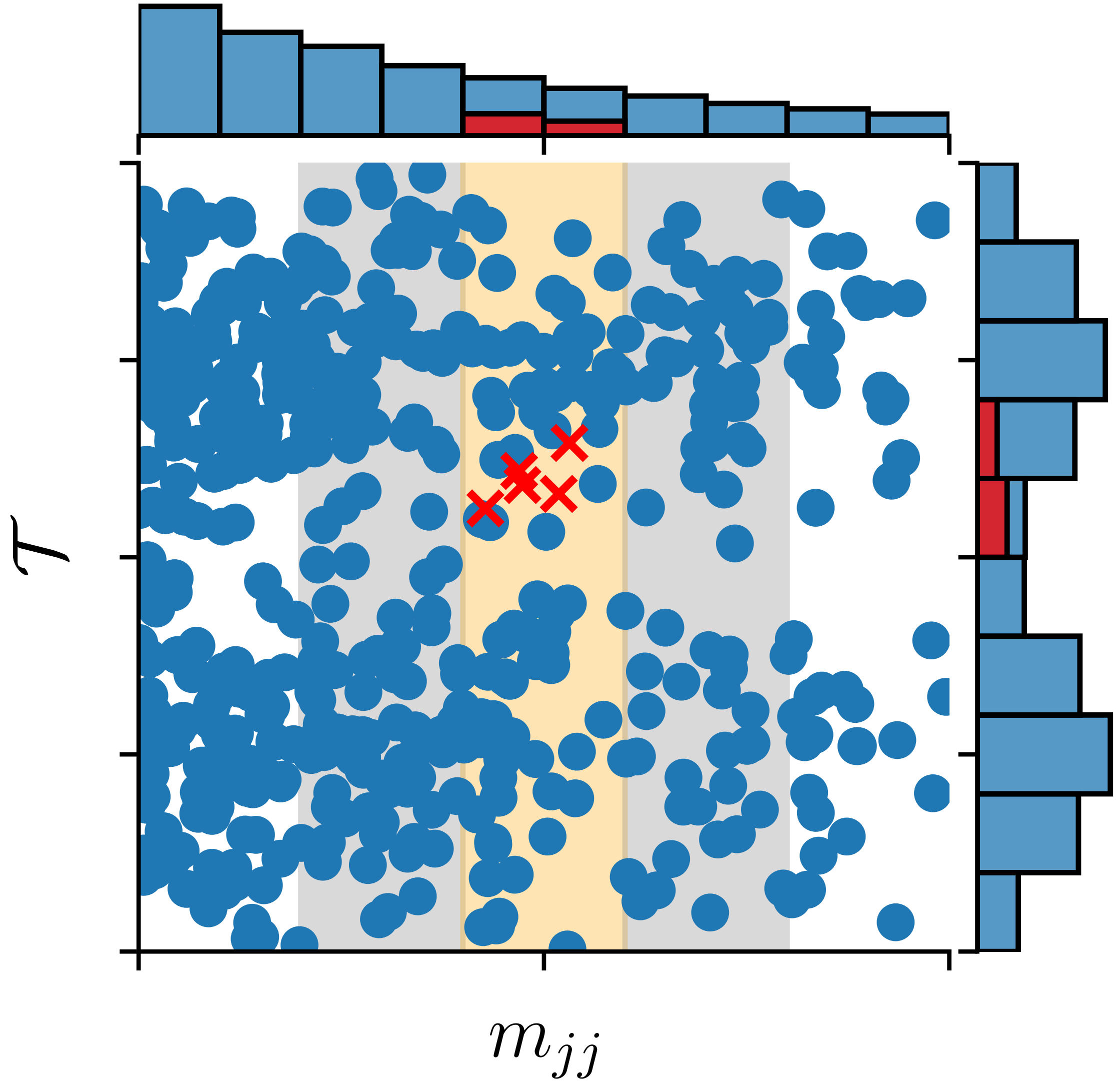
- Many more BSM models than possible analyses: **Need model-agnostic searches!**
- This analysis: **Signal-agnostic m_{JJ} Bump Hunt using Many Features**
- Previous analysis: ATLAS CWOLA Round 1 [2005.02983] (using less features)

The 4-step Analysis Strategy



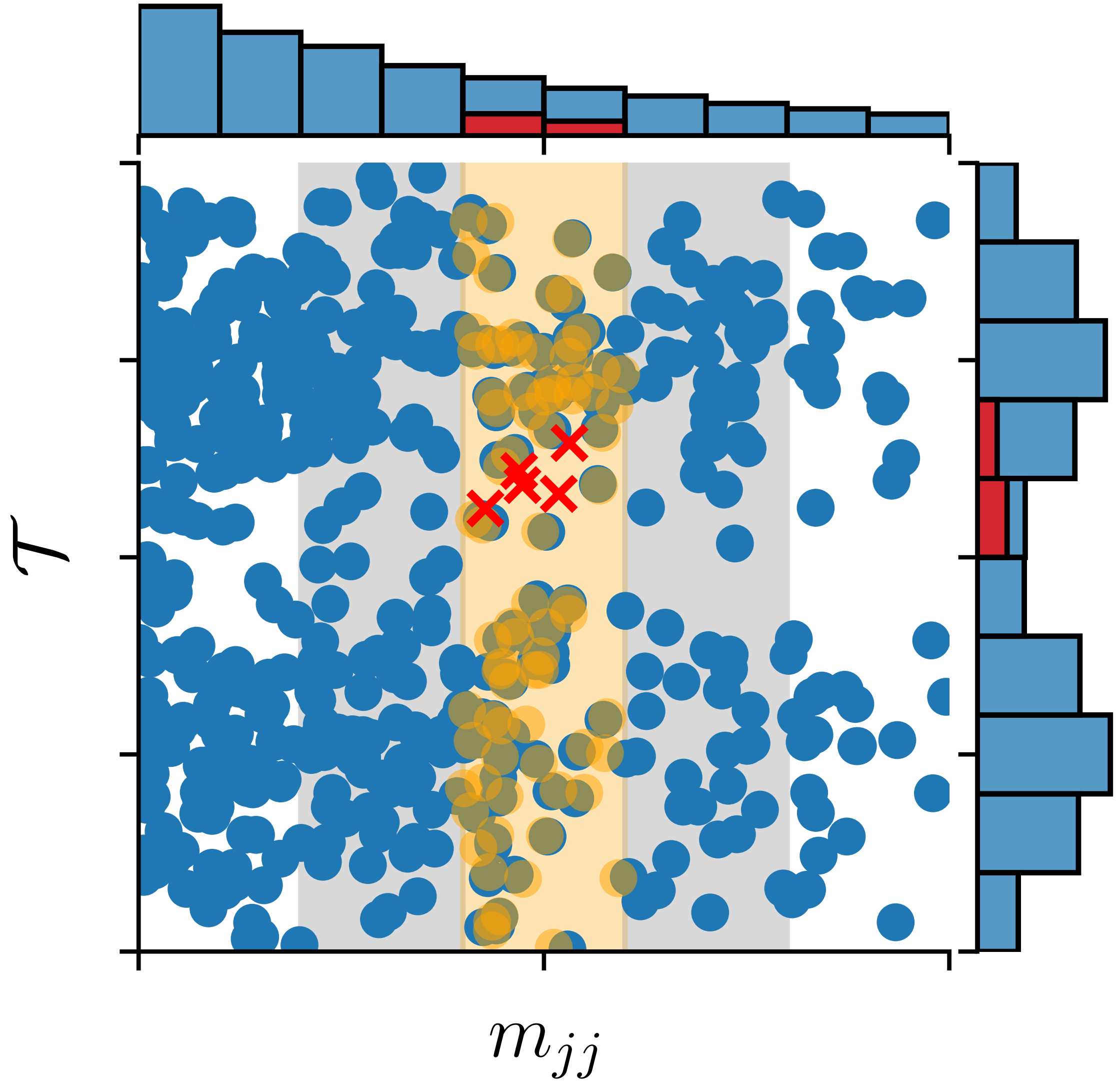
The 4-step Analysis Strategy

1. Definition of Regions



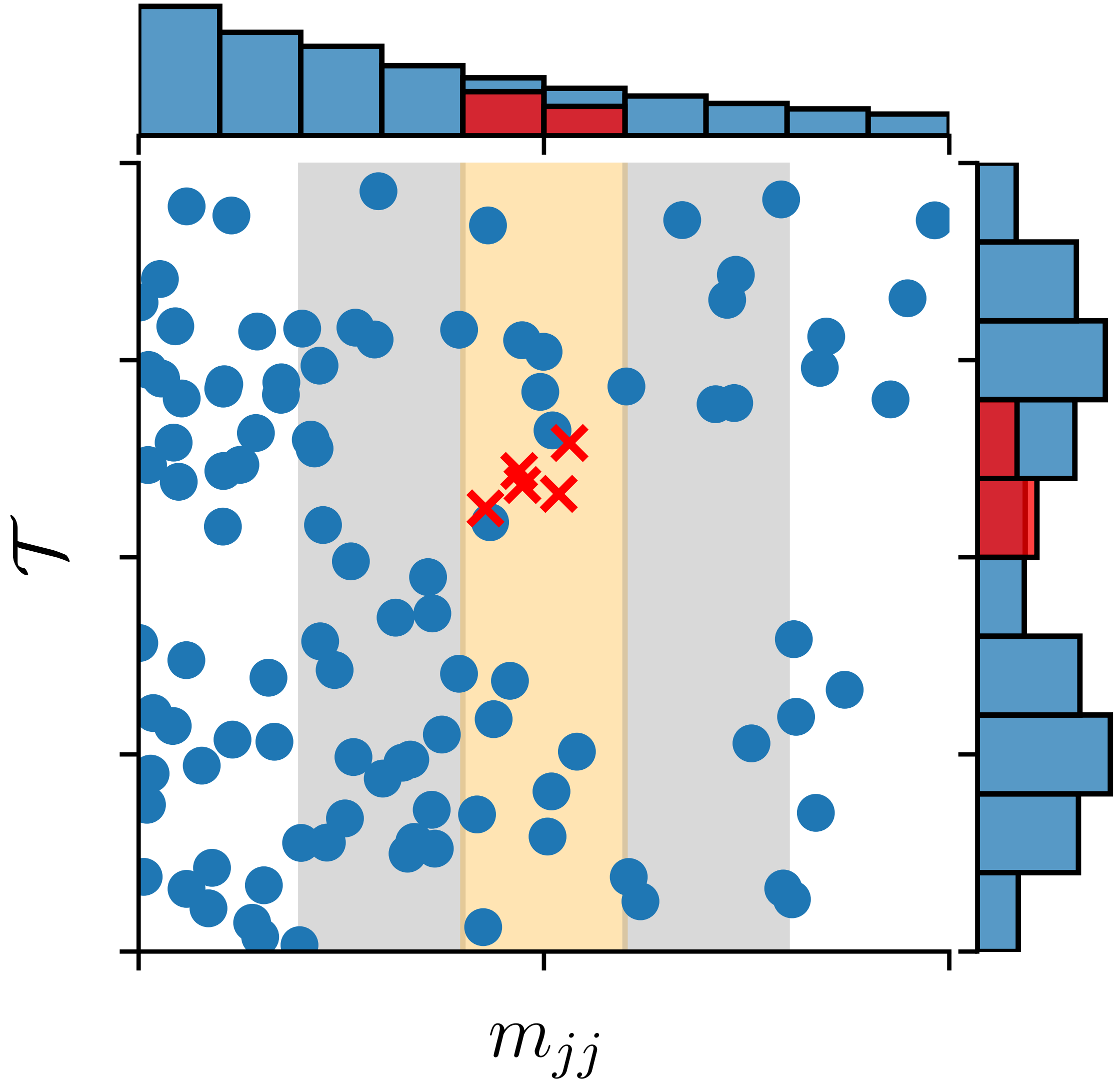
The 4-step Analysis Strategy

2. Background Estimation



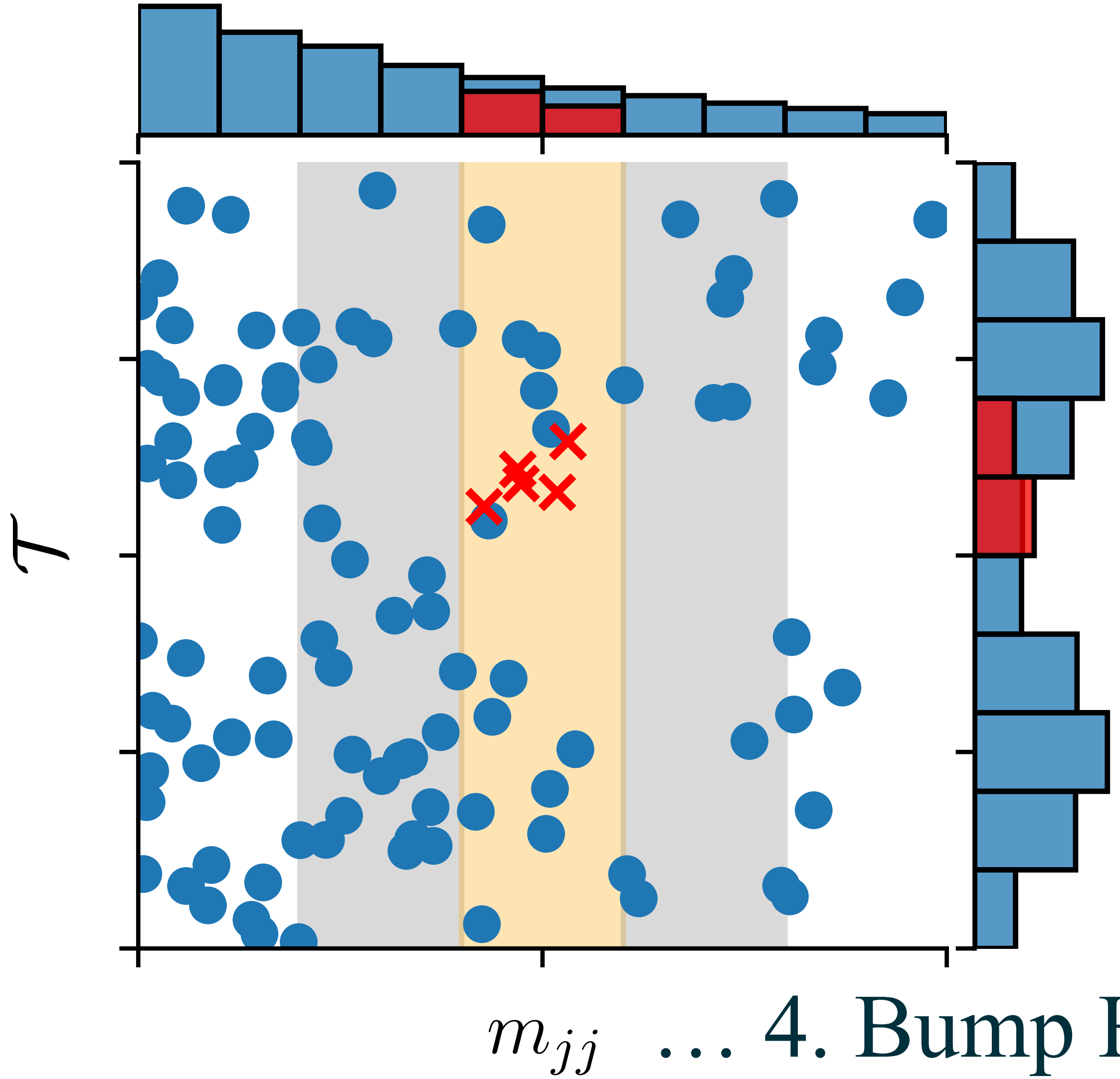
The 4-step Analysis Strategy

3. Classification of Signal



The 4-step Analysis Strategy

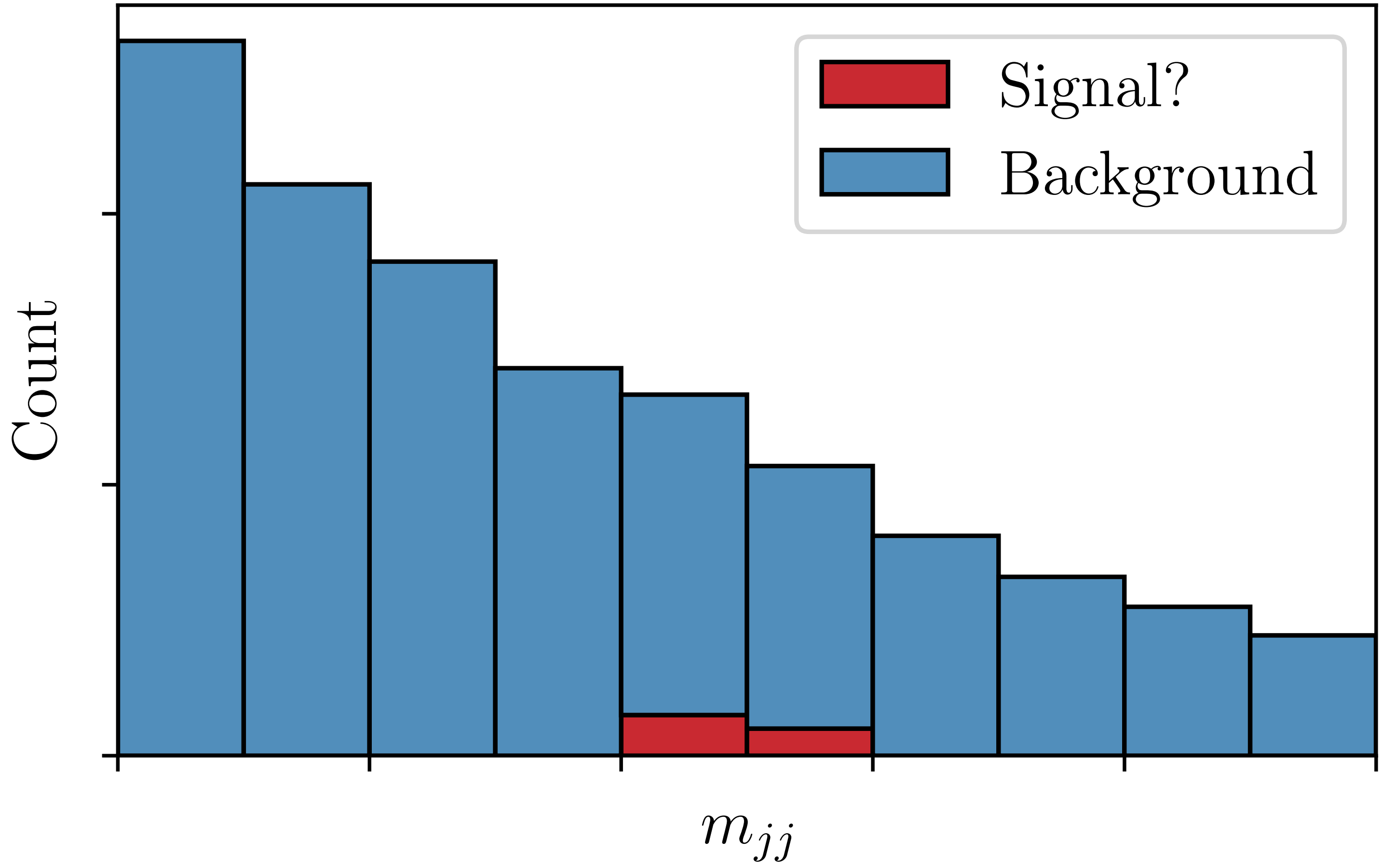
3. Classification of Signal



… 4. Bump Hunt

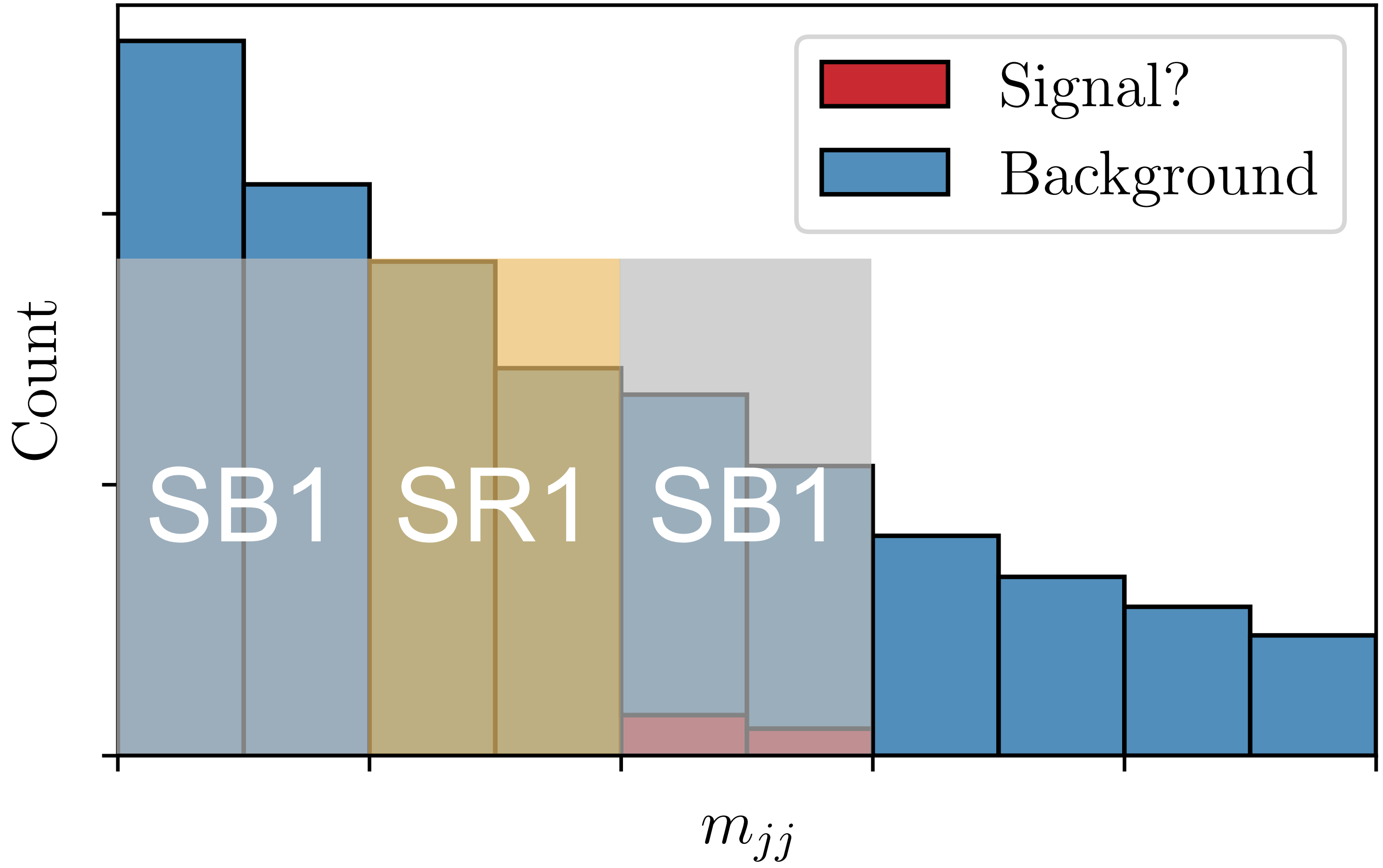
Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



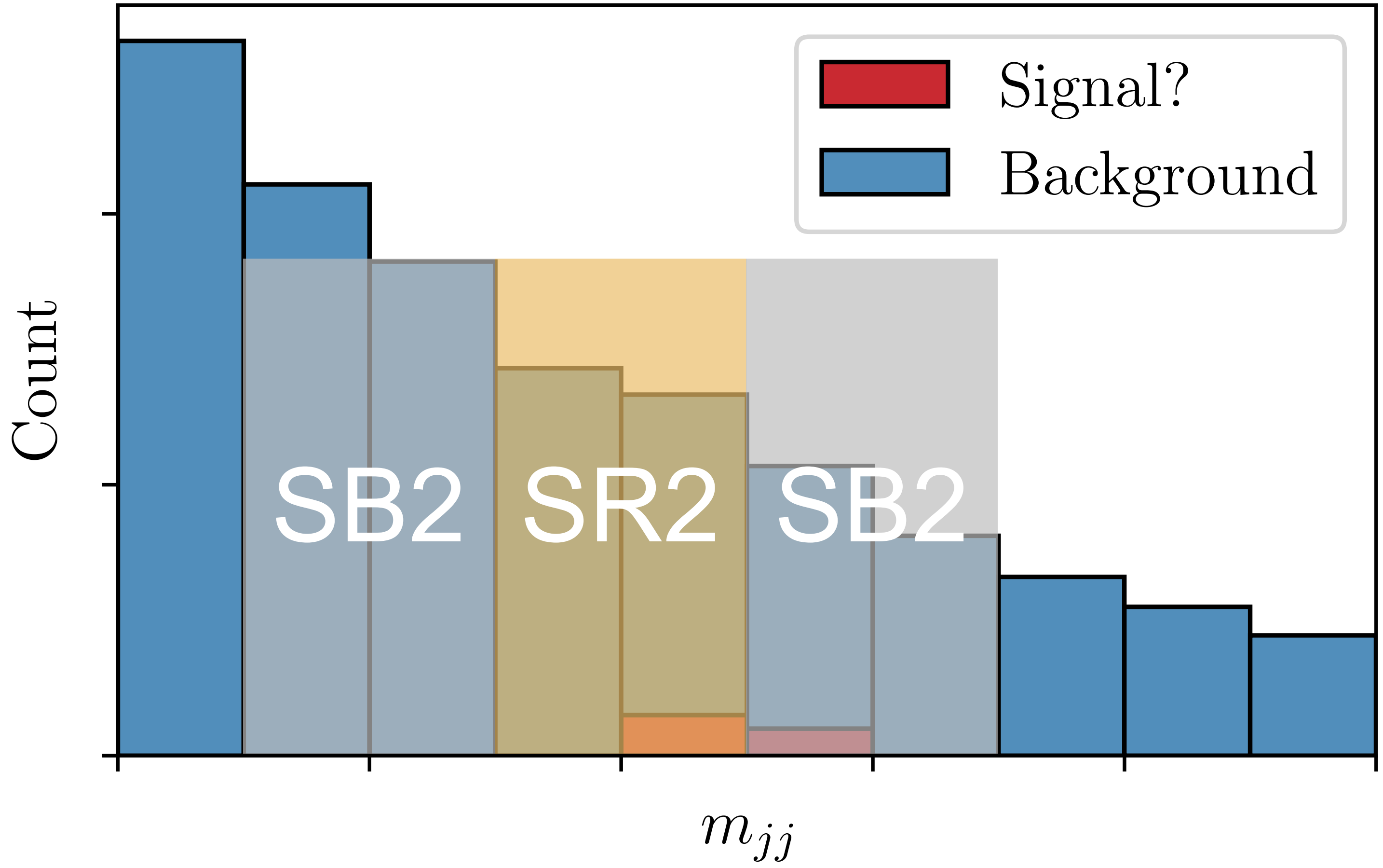
Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



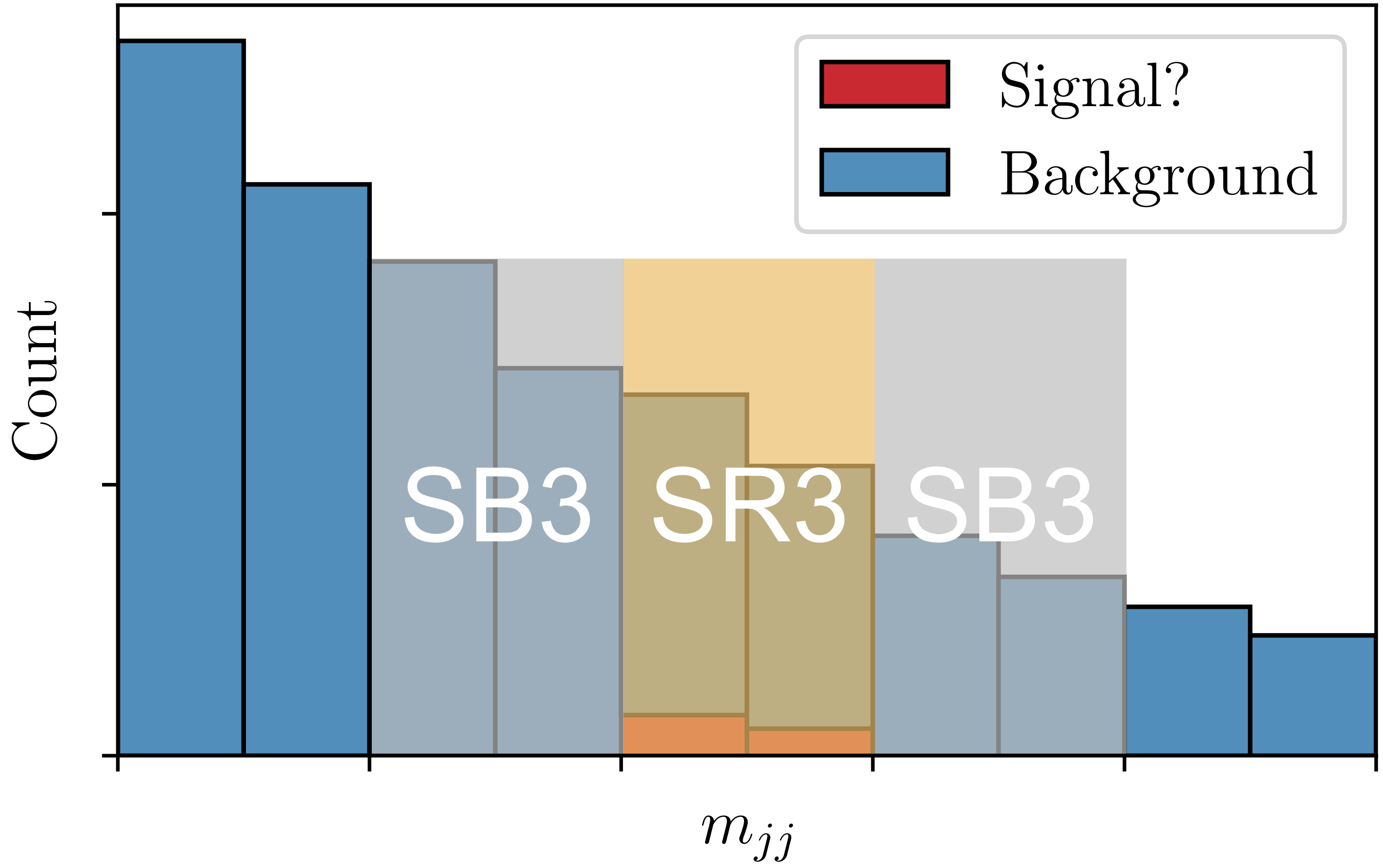
Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



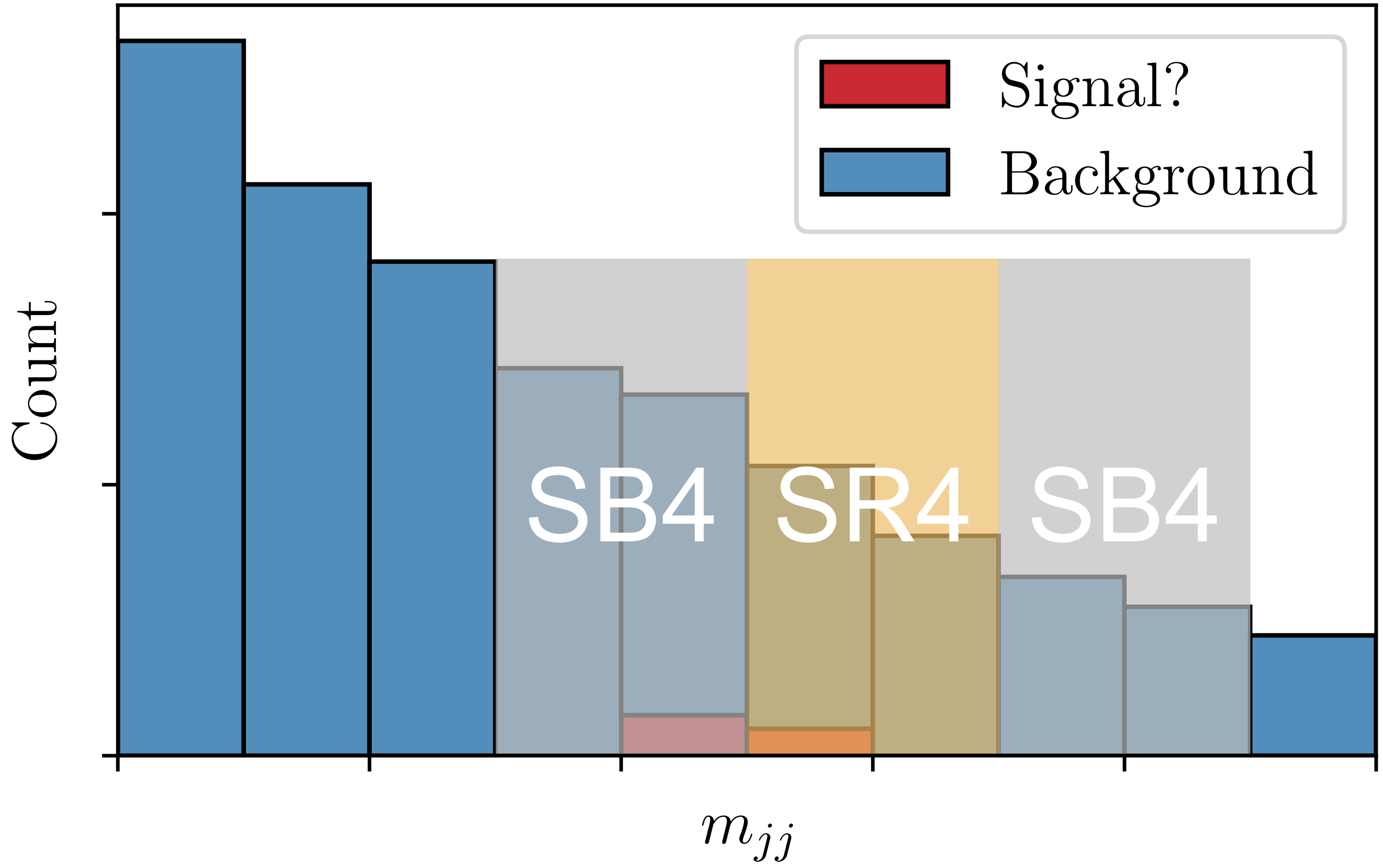
Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



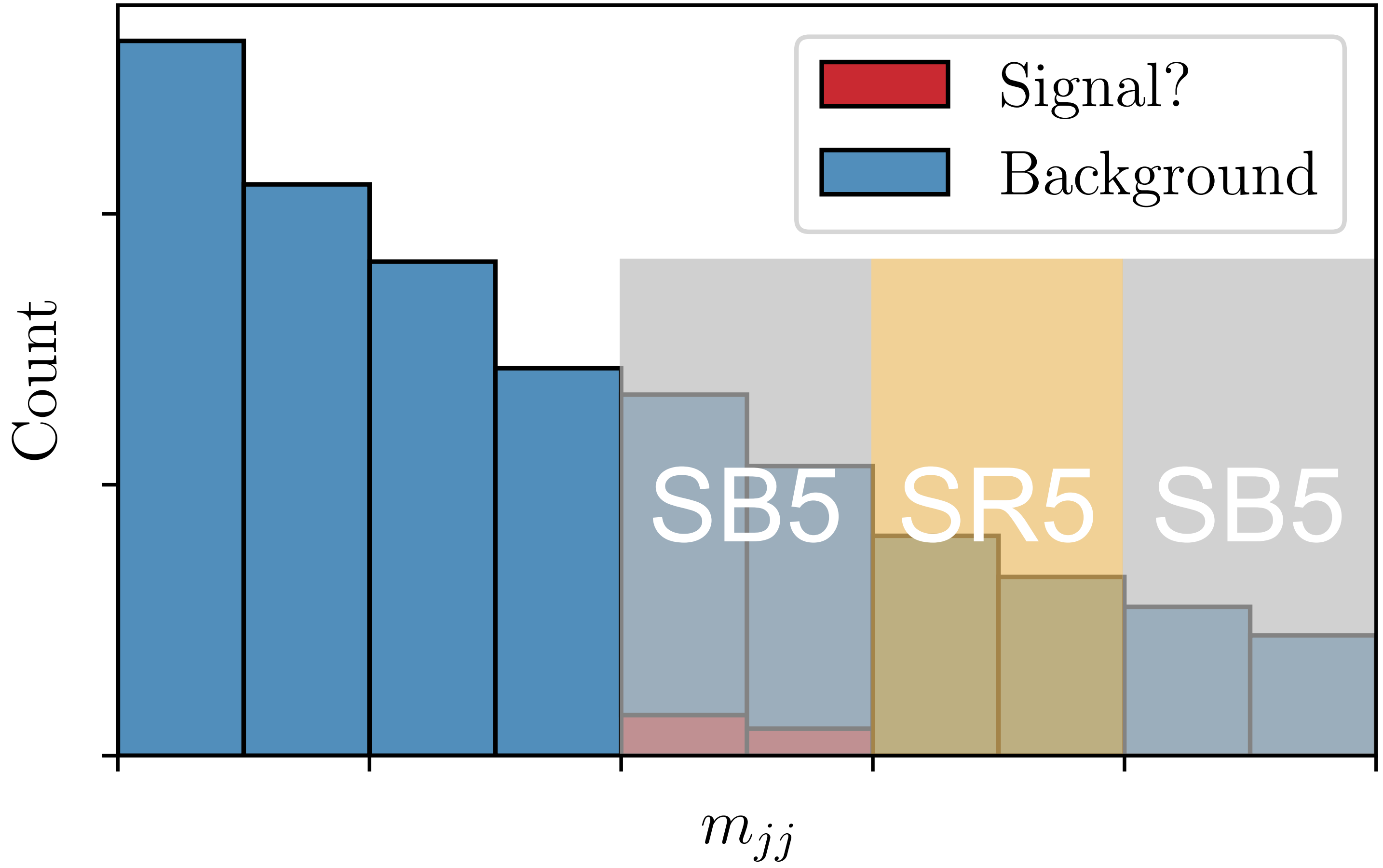
Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



Definition of Signal and Background Regions

- Events with ≥ 2 large radius jets (anti-kt, R=1) with low rapidity difference
- Signal agnostic scan with 300GeV step size over whole m_{jj} range
- Signal Regions (SR): 600GeV; Background regions (SB): 600GeV below & above



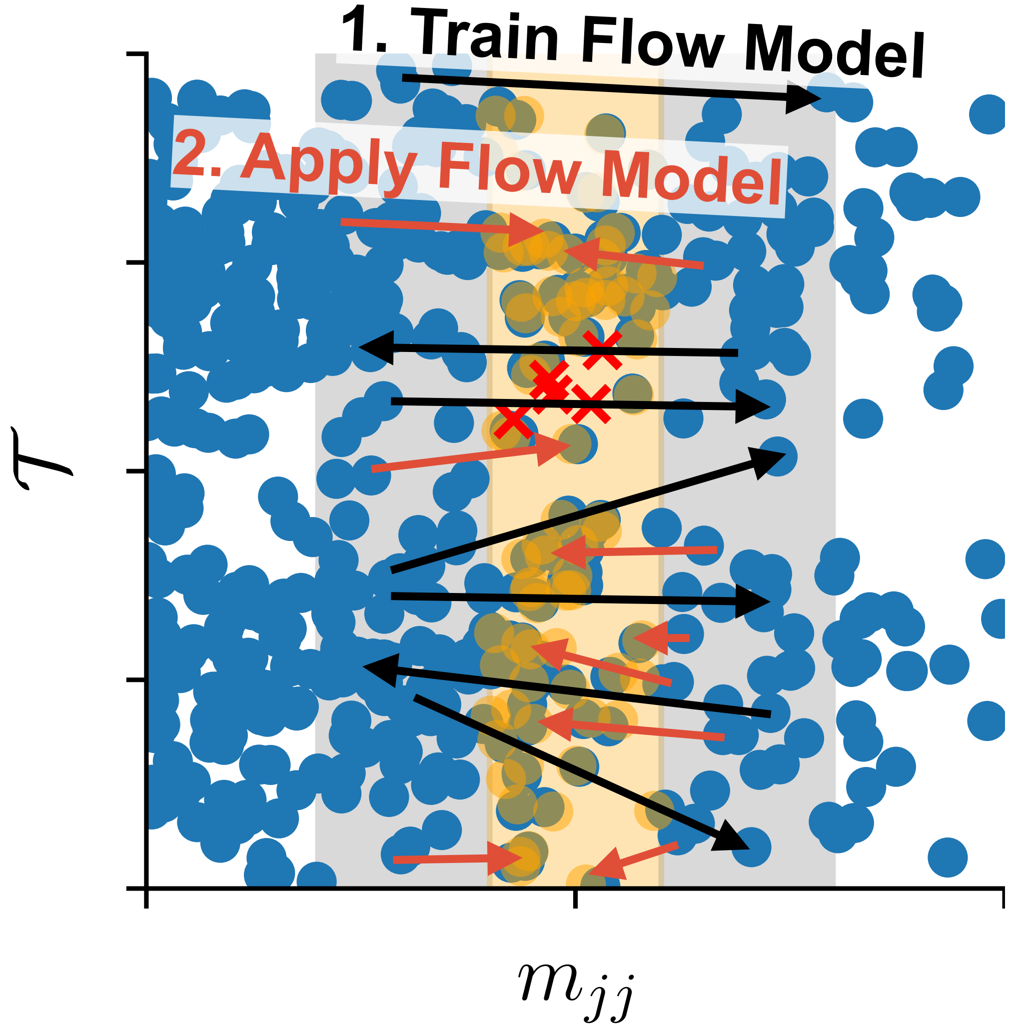
Background Estimation Two different methods

SALAD [2001.05001]

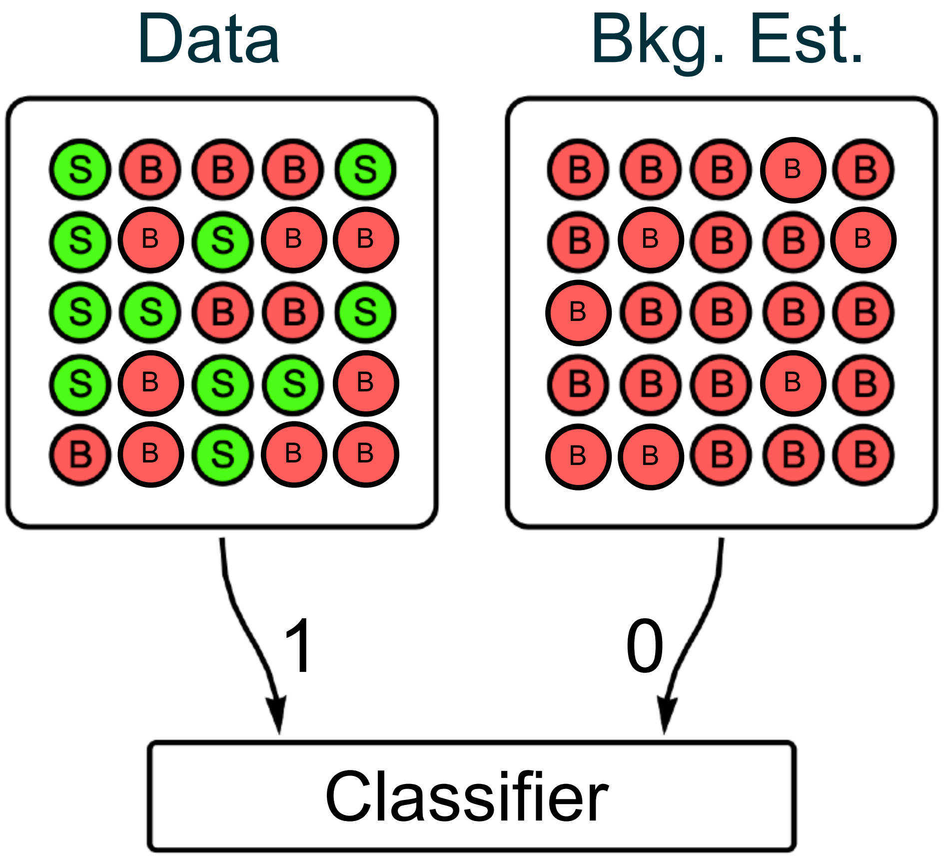
Background Estimation *Two different methods*

SALAD [2001.05001]

CURTAINS [2203.09470]



Classification of a Potential Signal

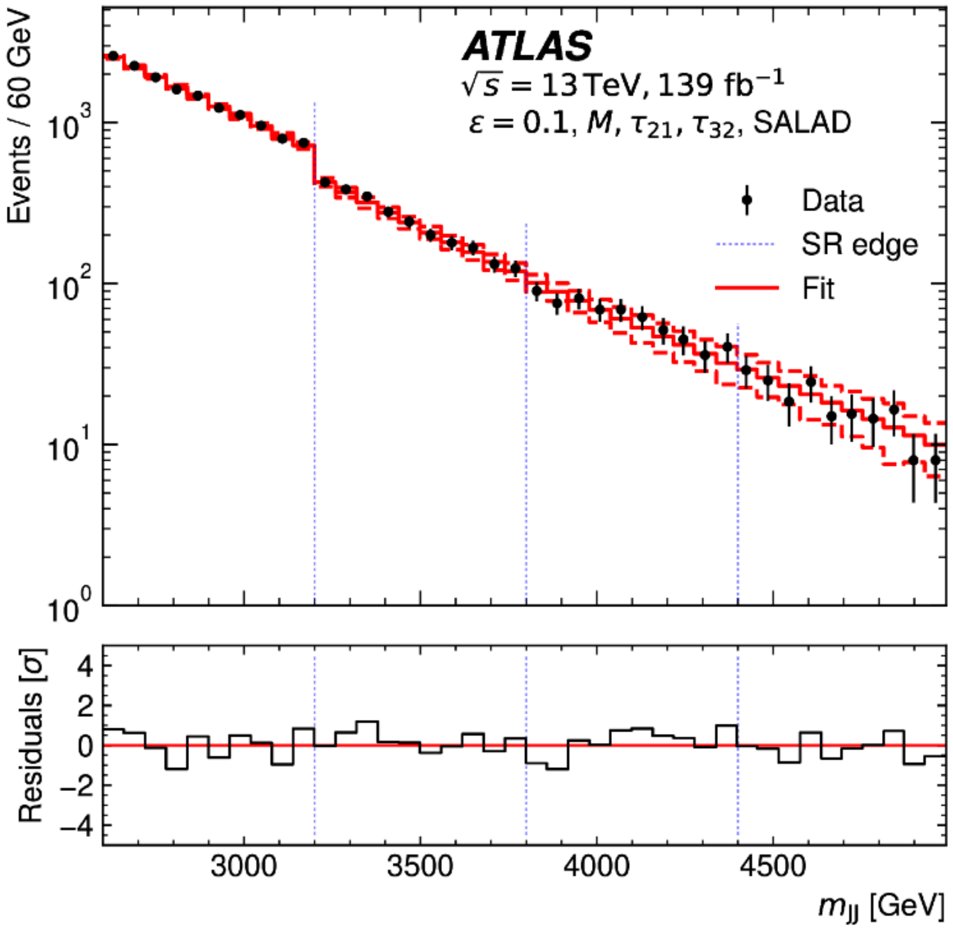


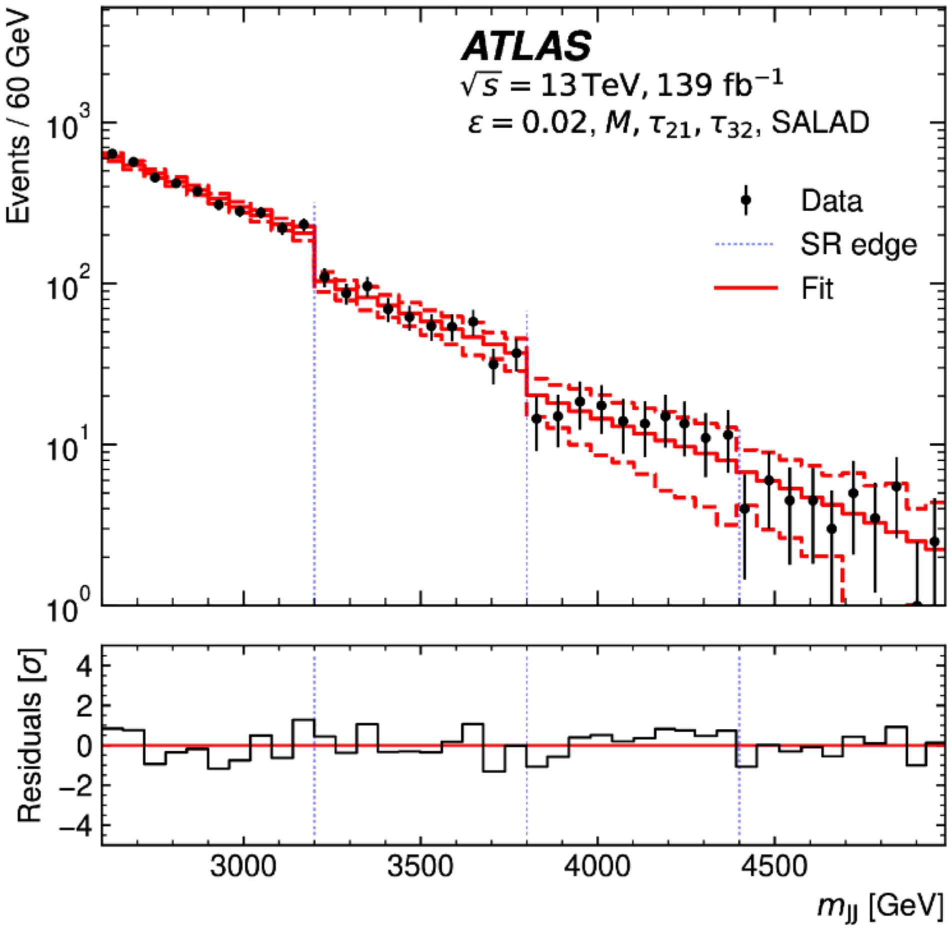
- CWOLA method [1708.02949]: Training of mixed samples (S+B vs B) converges to optimal classifier
- Used Features: Jet mass (M_j) & Jet substructure (τ_{21j}, τ_{32j}) of two AK10 jets (j)

Scan three subsets: $\{M_j\}$ $\{M_j, \tau_{21j}\}$ $\{M_j, \tau_{21j}, \tau_{32j}\}$

Cut on the classifier and Distribution

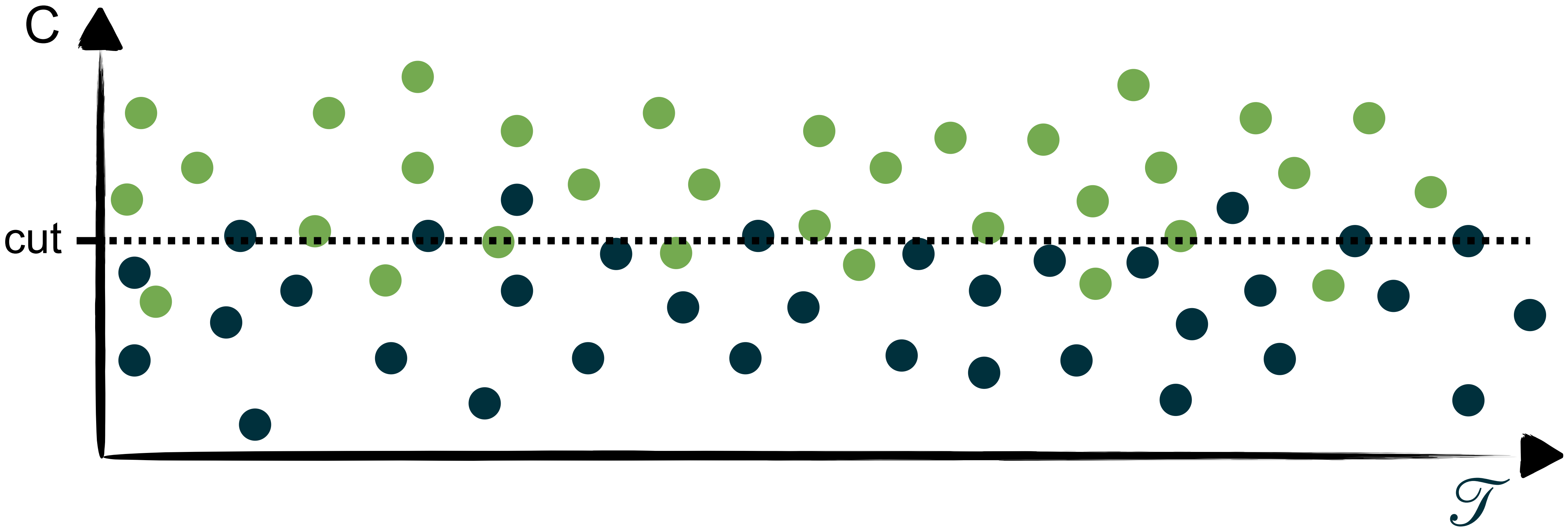
- Cut on classifier output to increase signal purity / suppress background
- Cut on different background efficiencies (FPR) using 2% and 10%





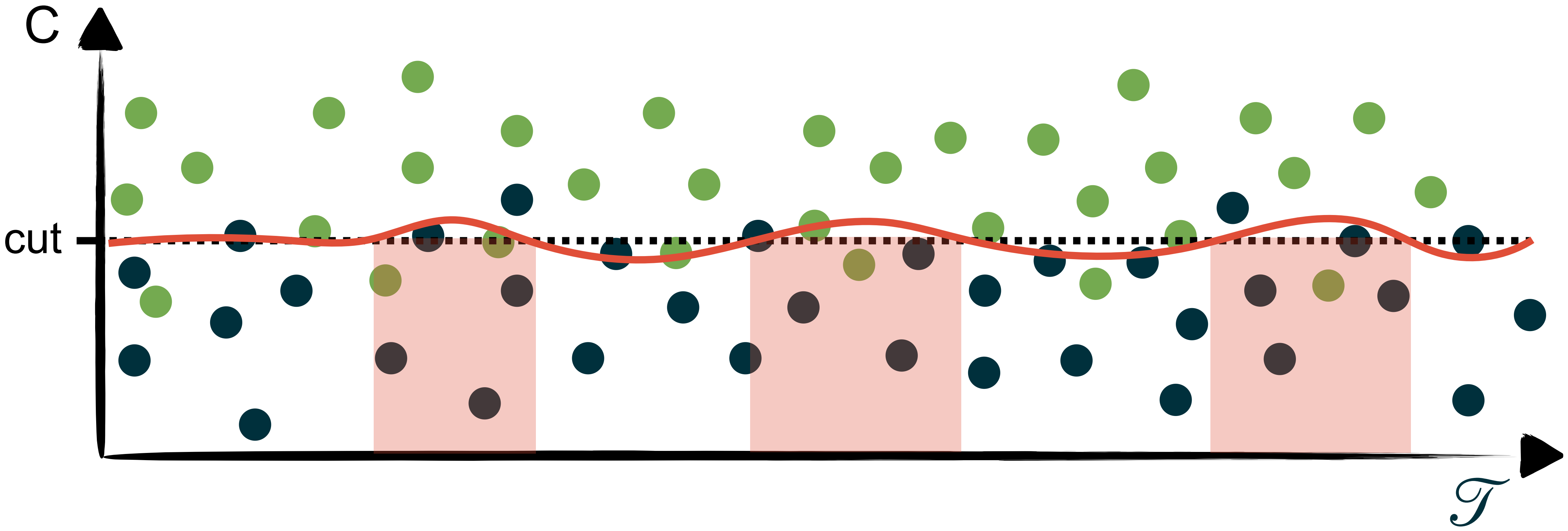
Ensembling strategy

- Random fluctuations of single CWOLA classifier lead to disjunct phase space after cut
- Use ensemble of 10 classifiers, make cut on each classifier, take mean of number of events



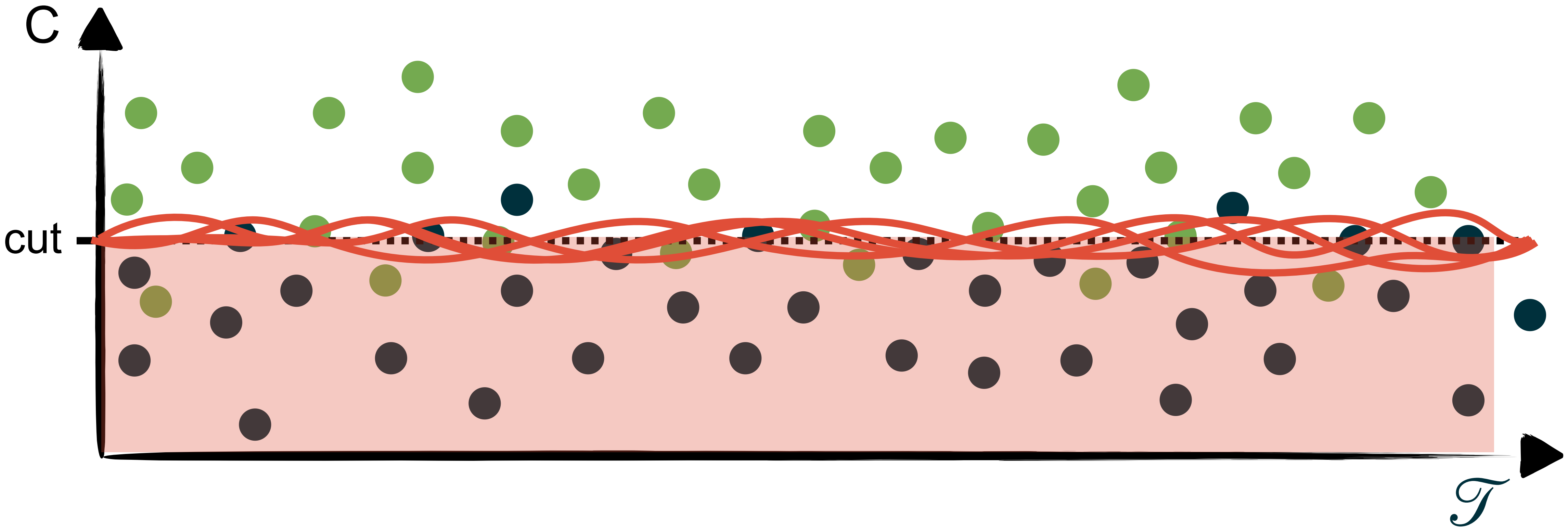
Ensembling strategy

- Random fluctuations of single CWOLA classifier lead to disjunct phase space after cut
- Use ensemble of 10 classifiers, make cut on each classifier, take mean of number of events

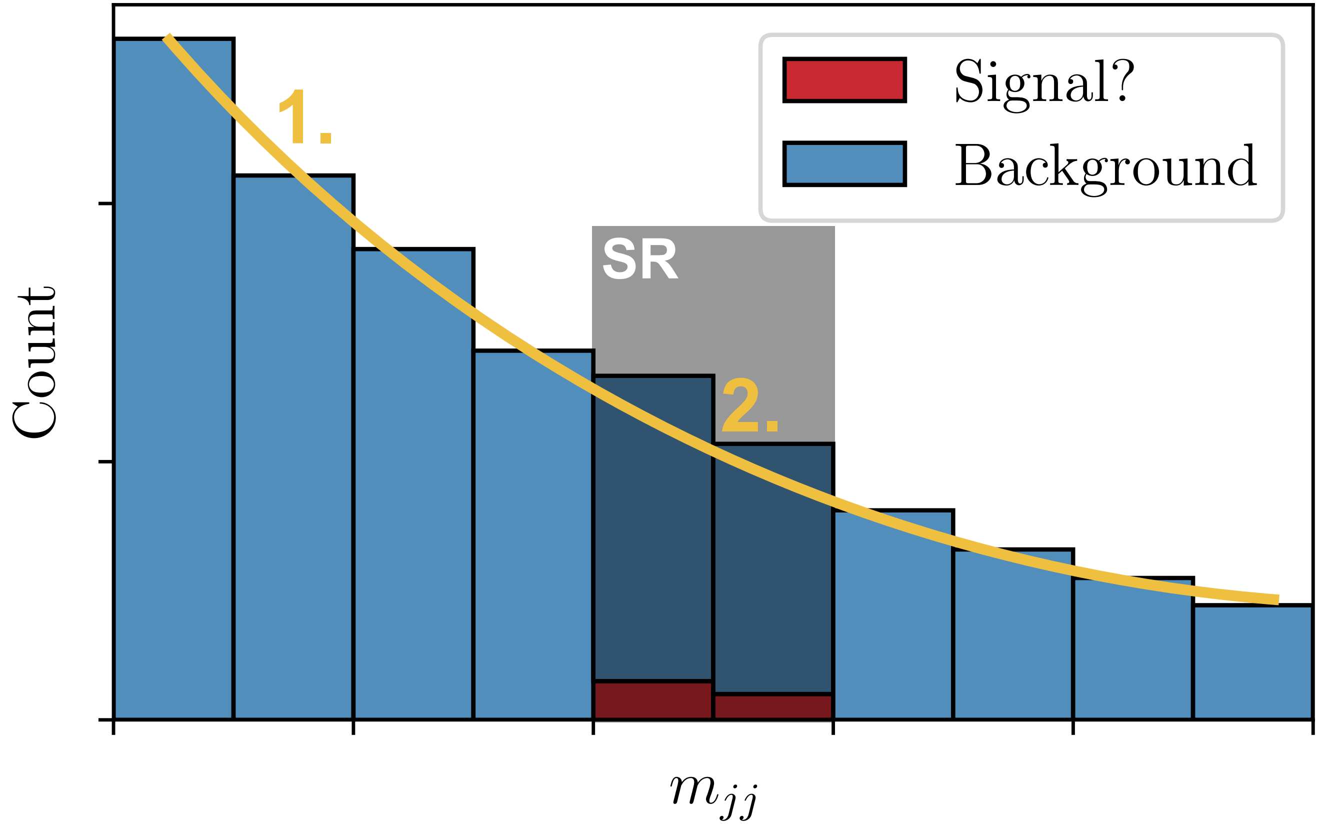


Ensembling strategy

- Random fluctuations of single CWOLA classifier lead to disjunct phase space after cut
- Use ensemble of 10 classifiers, make cut on each classifier, take mean of number of events



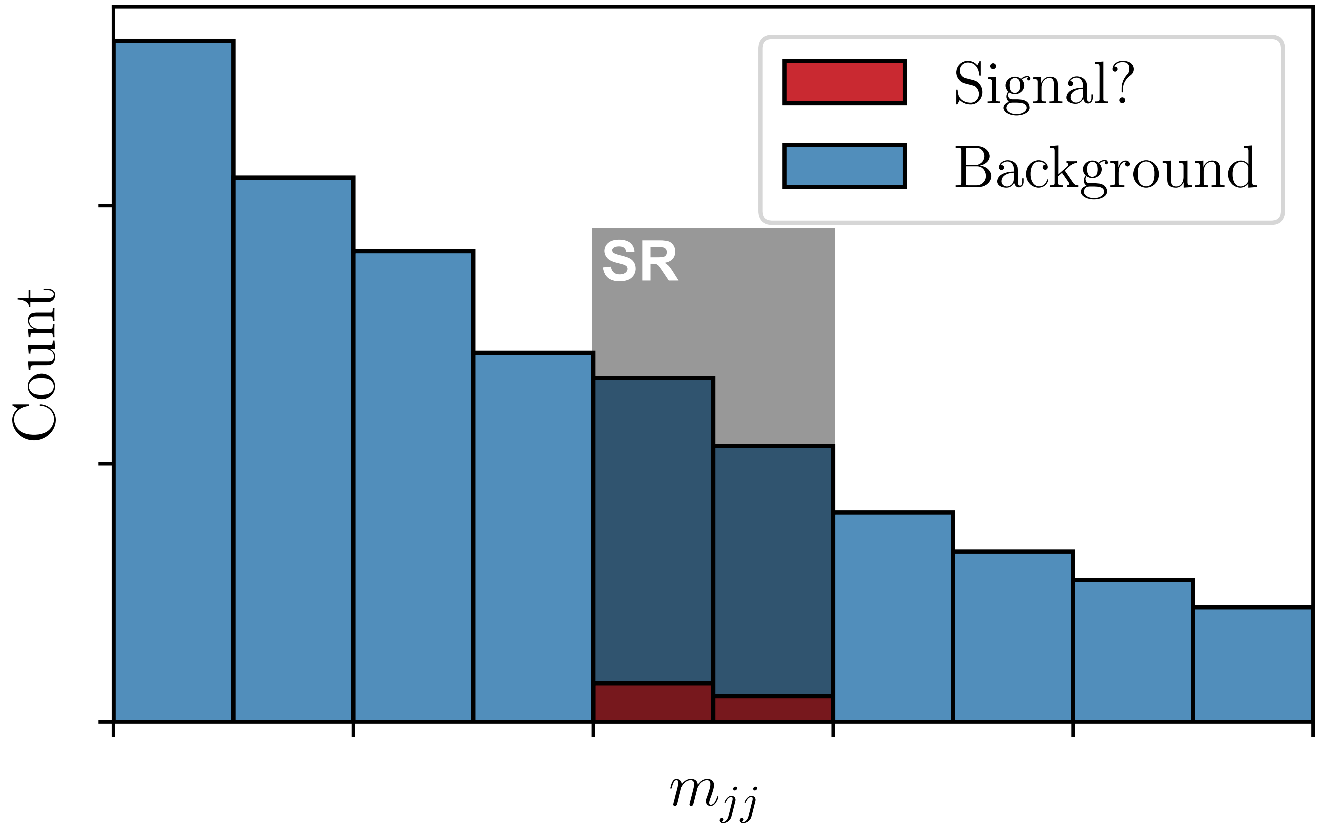
Inference: Fit and Likelihood



- If CWOLA increases purity of an existing signal, there will be a bump in the m_{jj} spectrum:
  1. Fit exponential in side bands (up to four-parametric exponential, iterative procedure)
  2. Compare sum of counts in signal region (SR)

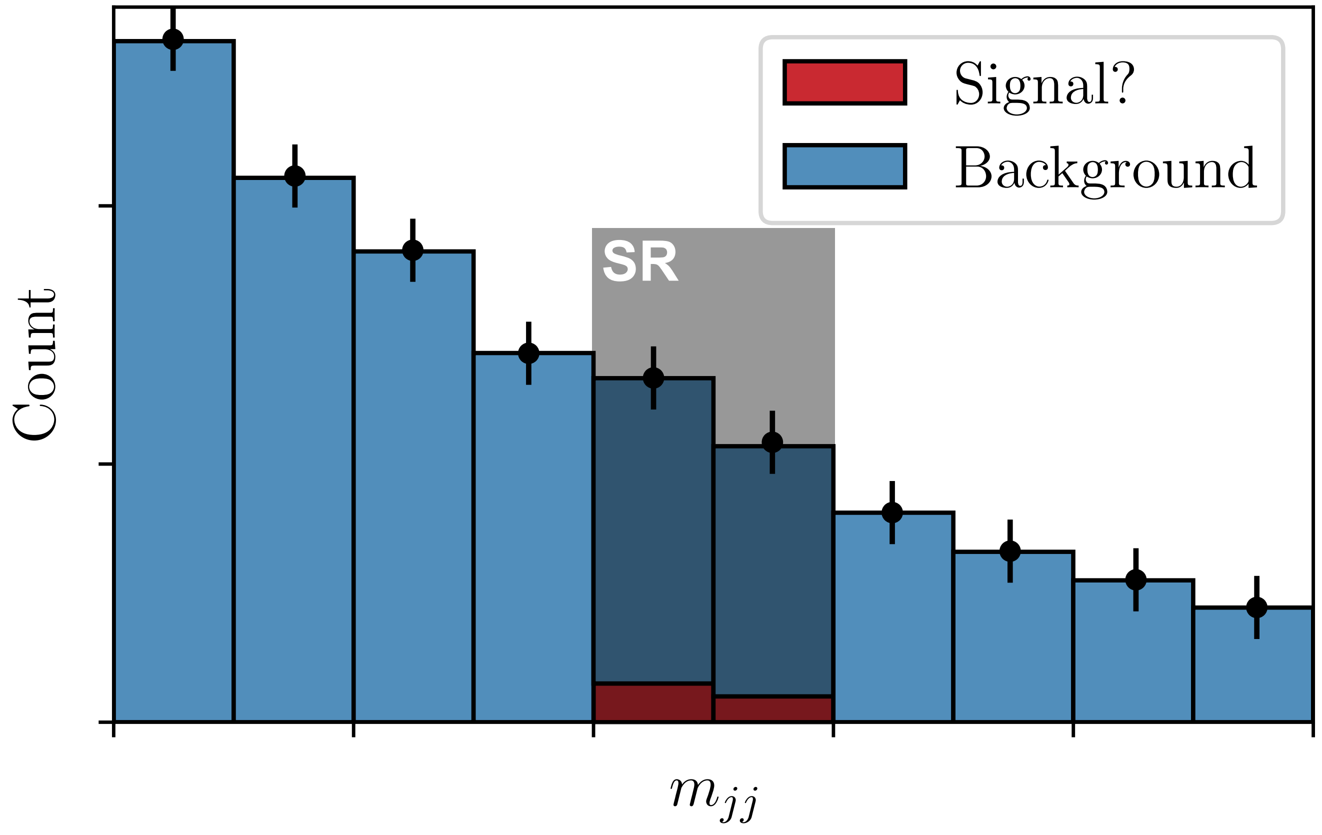
Inference: Uncertainties

- Each bin count has two uncertainties:
  - Poisson of counts
  - Classifier ensemble



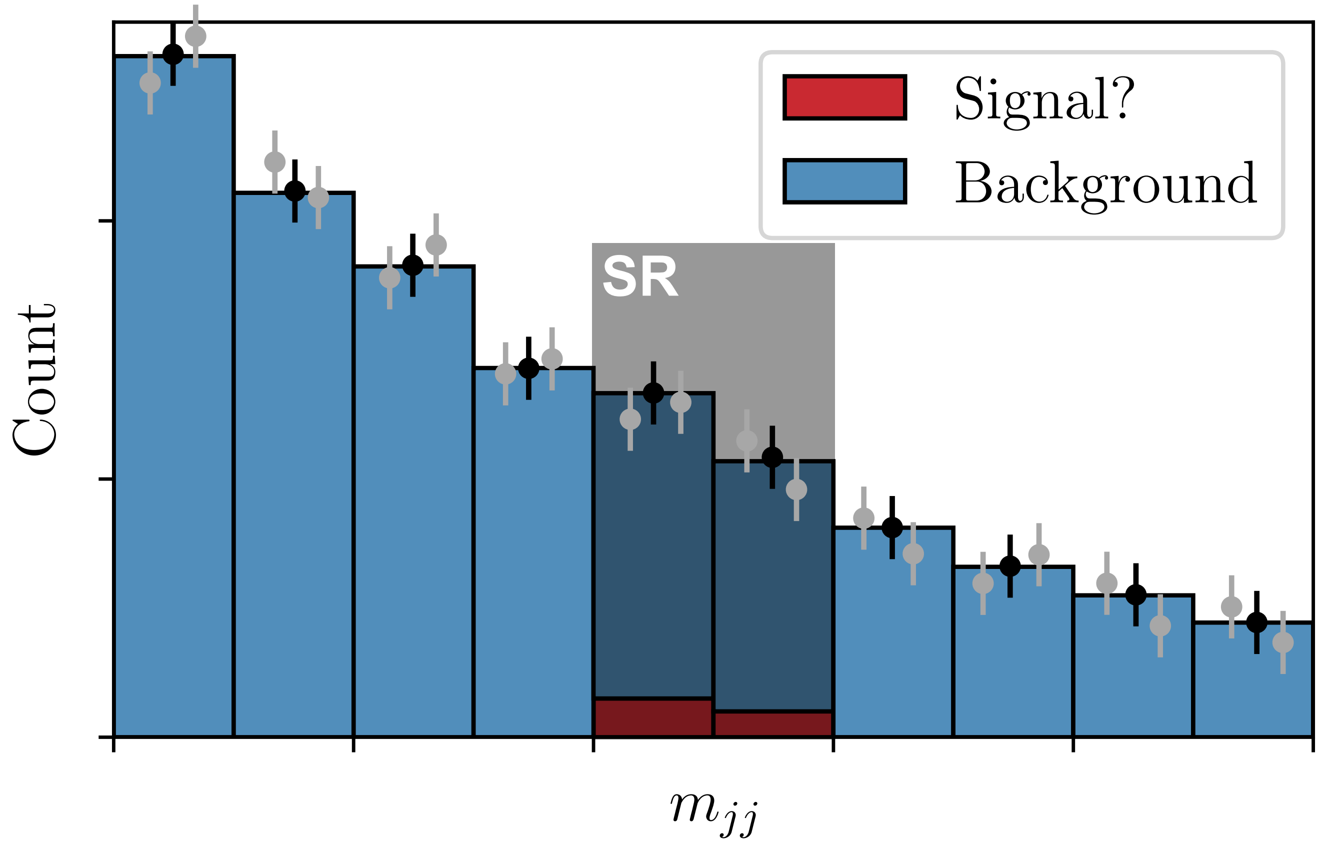
Inference: Uncertainties

- Each bin count has two uncertainties:
  - Poisson of counts
  - Classifier ensemble



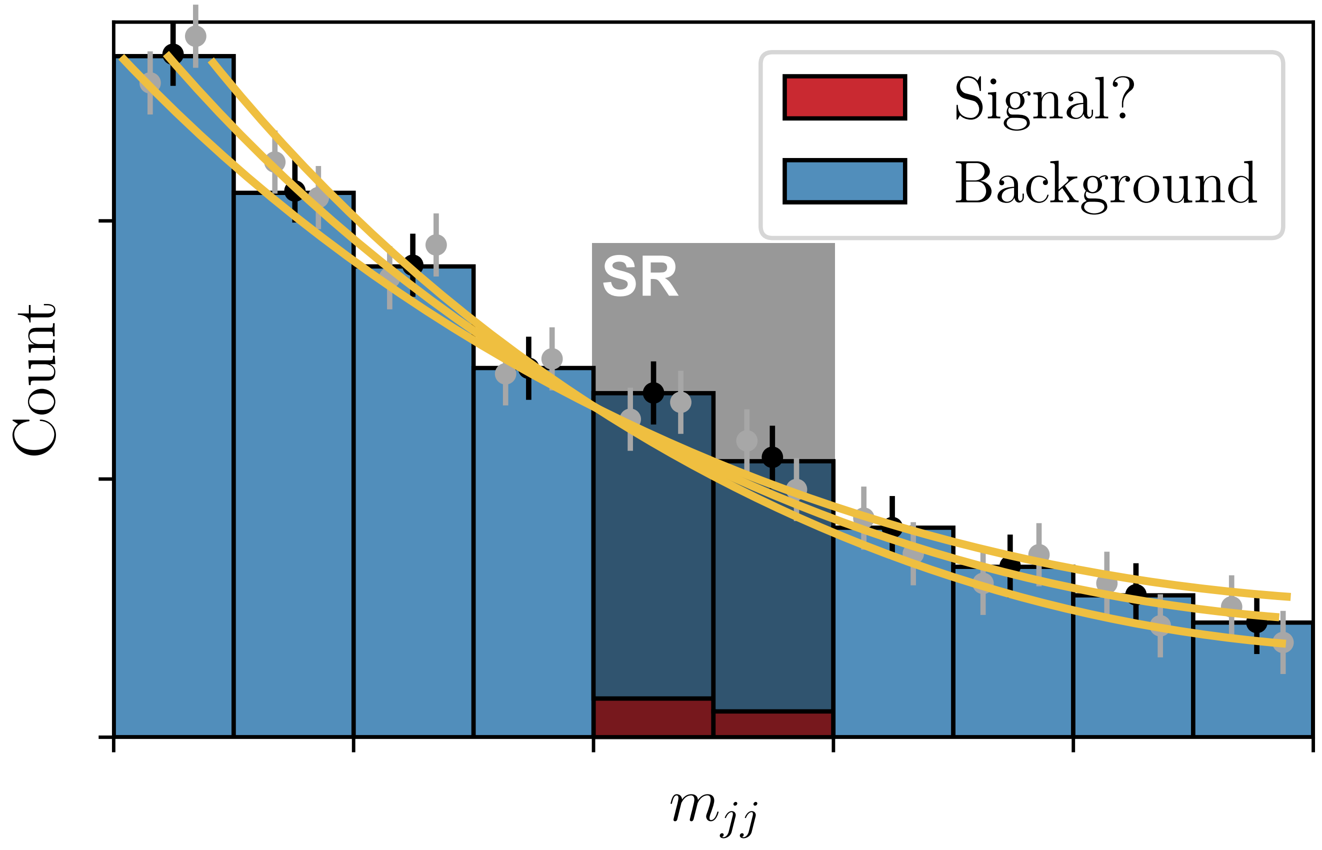
Inference: Uncertainties

- Each bin count has two uncertainties:
  - Poisson of counts
  - Classifier ensemble



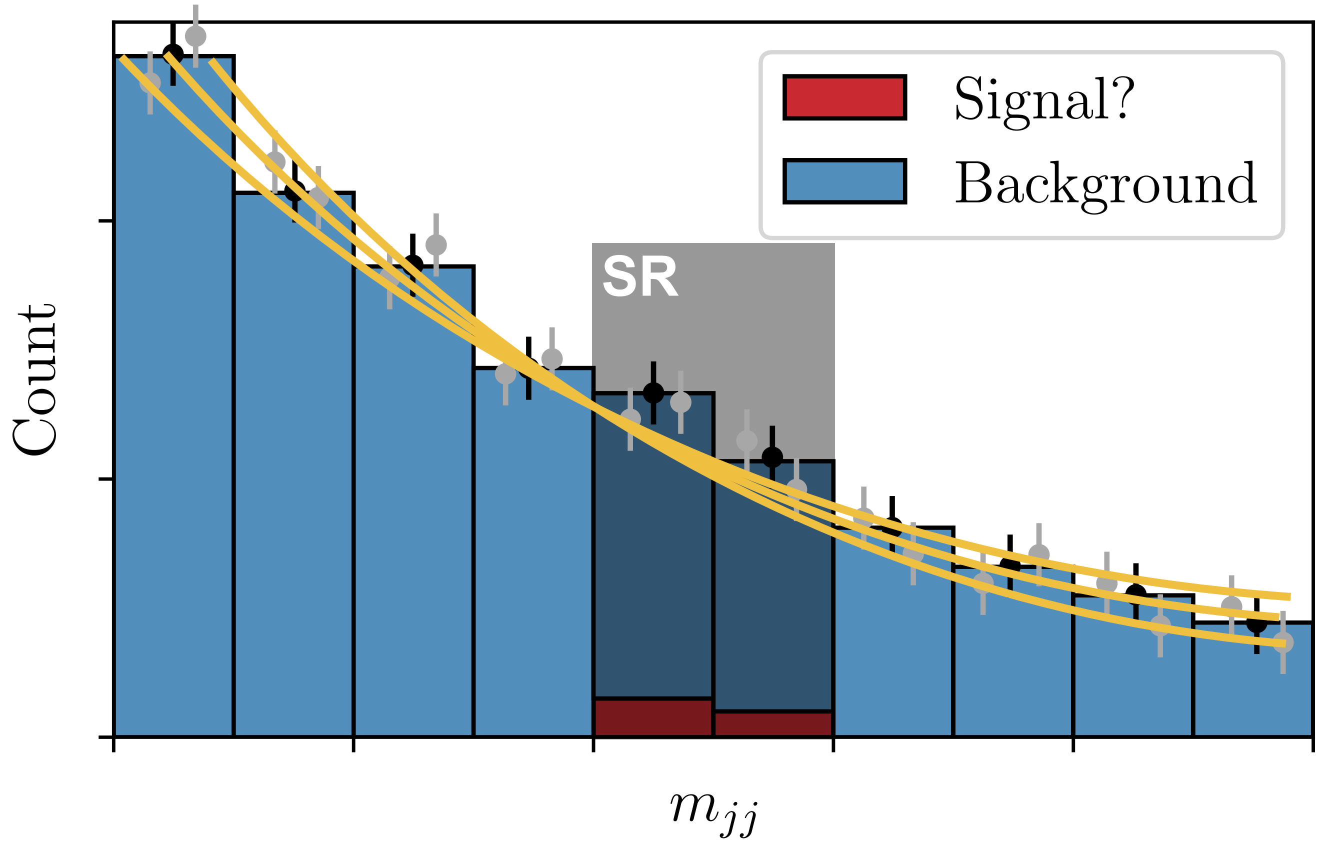
Inference: Uncertainties

- Each bin count has two uncertainties:
  - Poisson of counts
  - Classifier ensemble



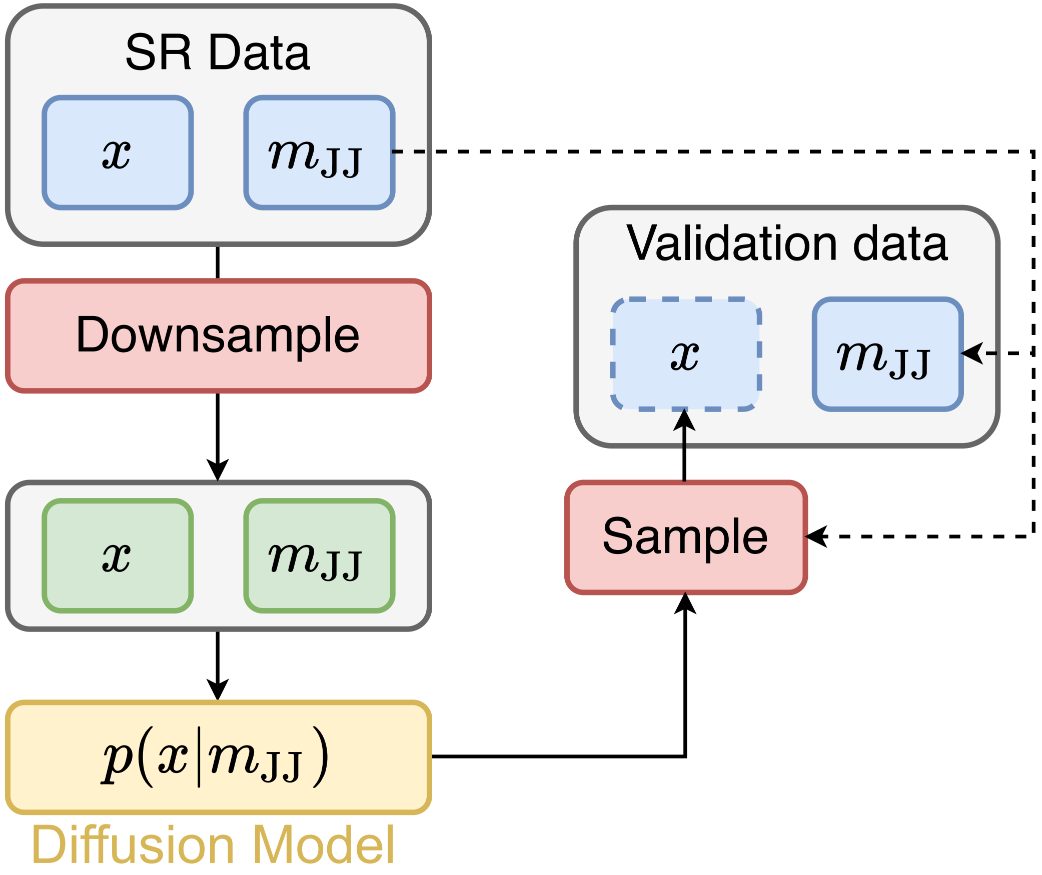
Inference: Uncertainties

- Each bin count has two uncertainties:
  - Poisson of counts
  - Classifier ensemble
- Largest uncertainties in SR:
  1. Bkg est: Fit
  2. Data: Poisson
  3. Data: Ensemble
- (Limits without signal uncertainties)



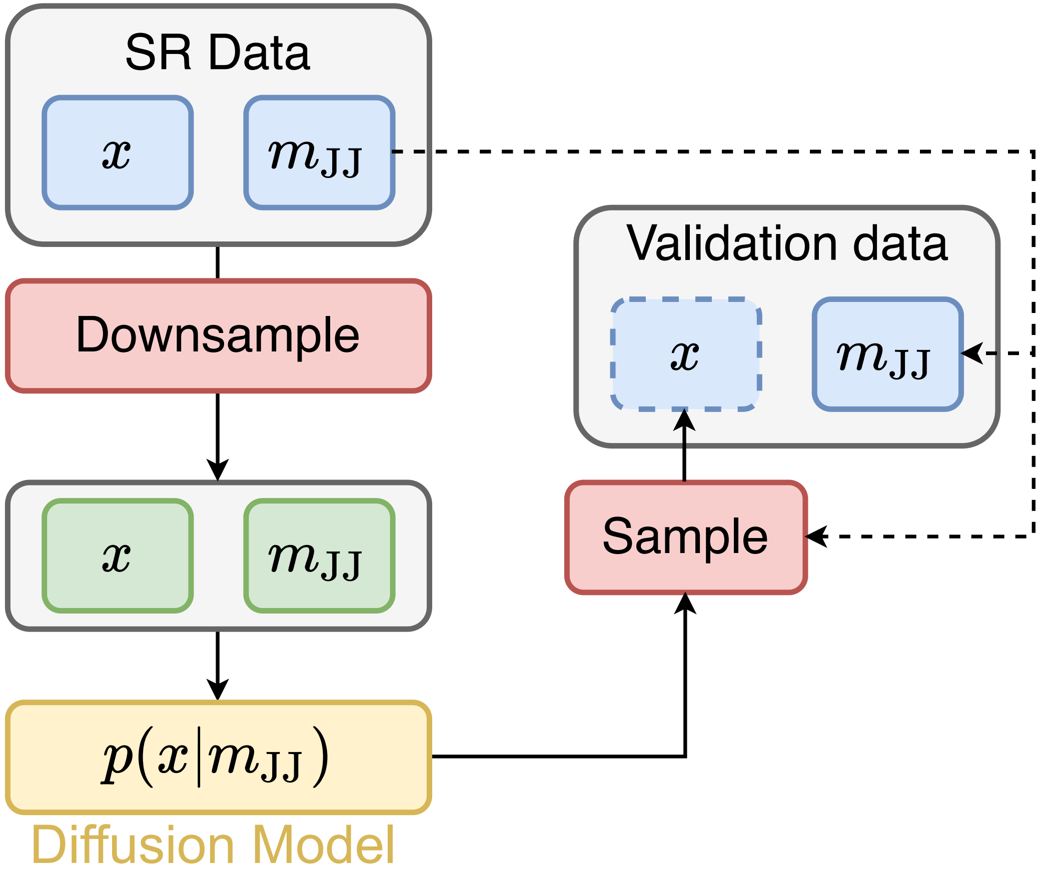
Three Test Datasets for Validation

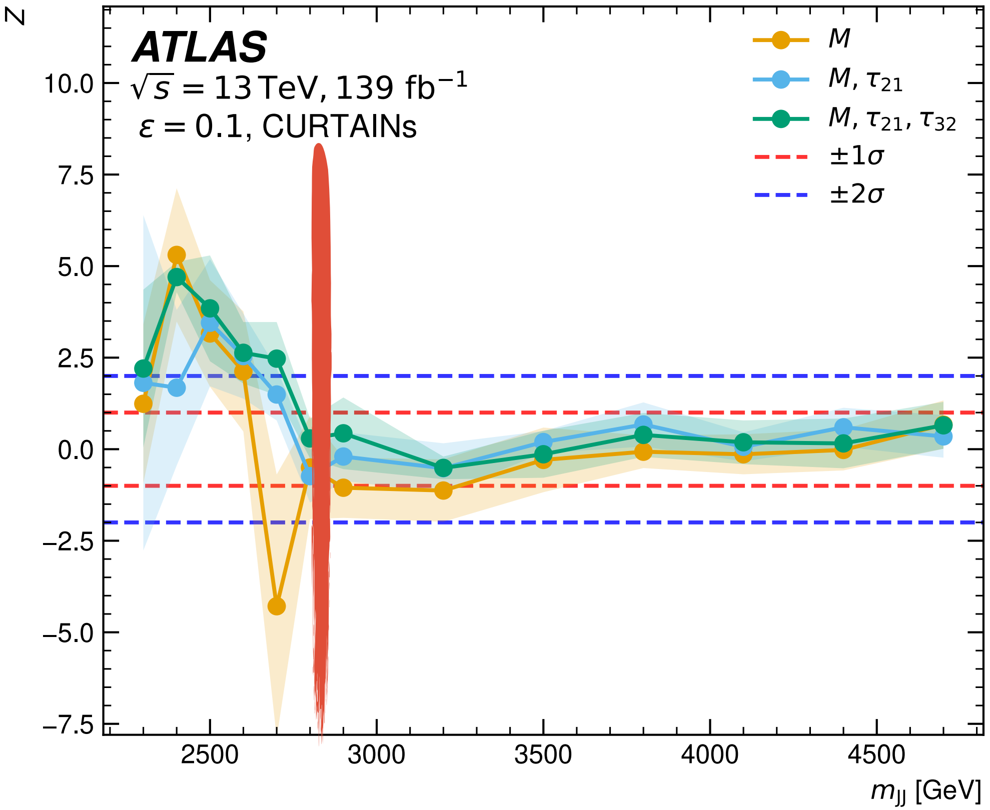
- Performance of anomaly detection analysis depends on unknown signal → **Validation challenging**
- Use three different test datasets without signal: Large ΔY, MC, **Up-Down-Sampling (below)**



Three Test Datasets for Validation

- Performance of anomaly detection analysis depends on unknown signal → **Validation challenging**
- Use three different test datasets without signal: Large ΔY, MC, **Up-Down-Sampling (below)**
- **Valid for m_{JJ} > 2900 GeV (more challenging background estimation for small m_{JJ})**





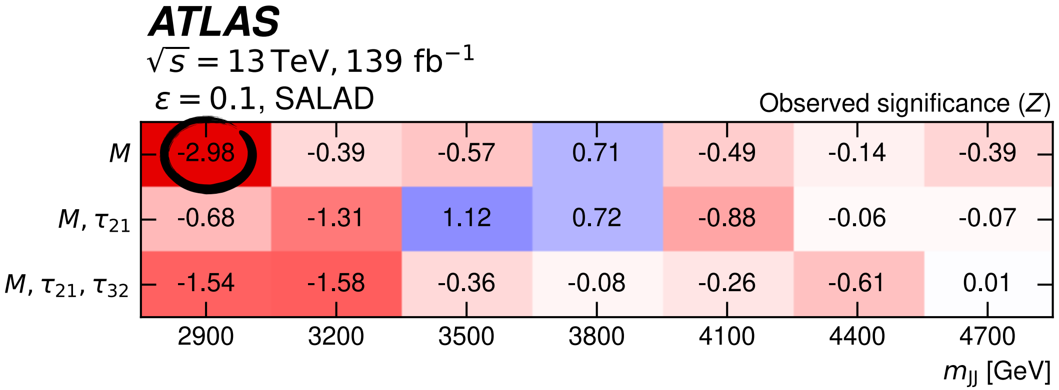
Results: Signal Agnostic Significances

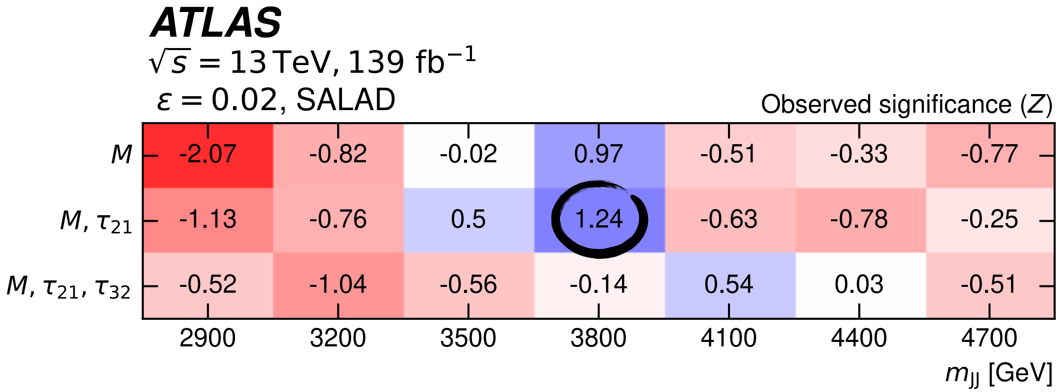
- Observed significance of: 2 methods, 2 efficiencies, 7 m_{JJ} regions, 3 features sets
- Largest significance is 1.24σ (1.26σ) & local deficit of -2.98σ (-2.54σ) for SALAD (CURTAINs)

ε = 10%

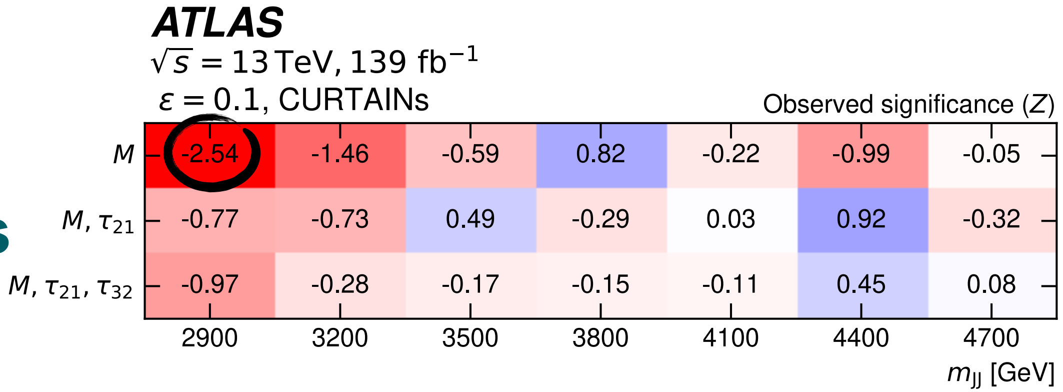
ε = 2%

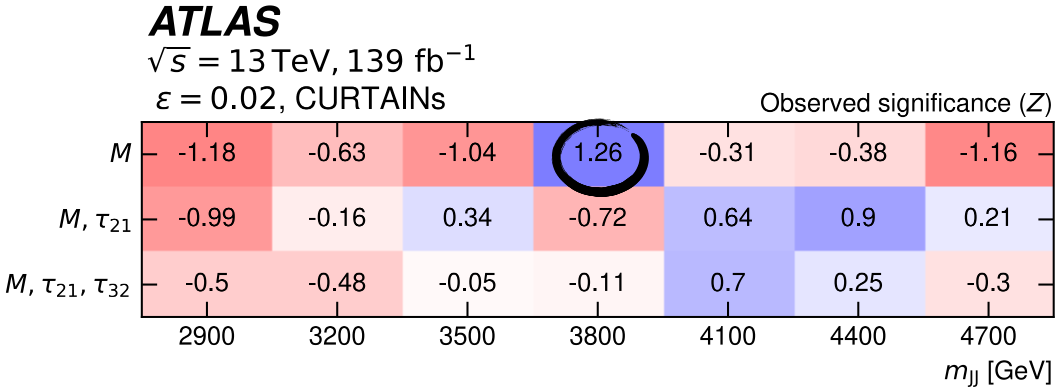
SALAD





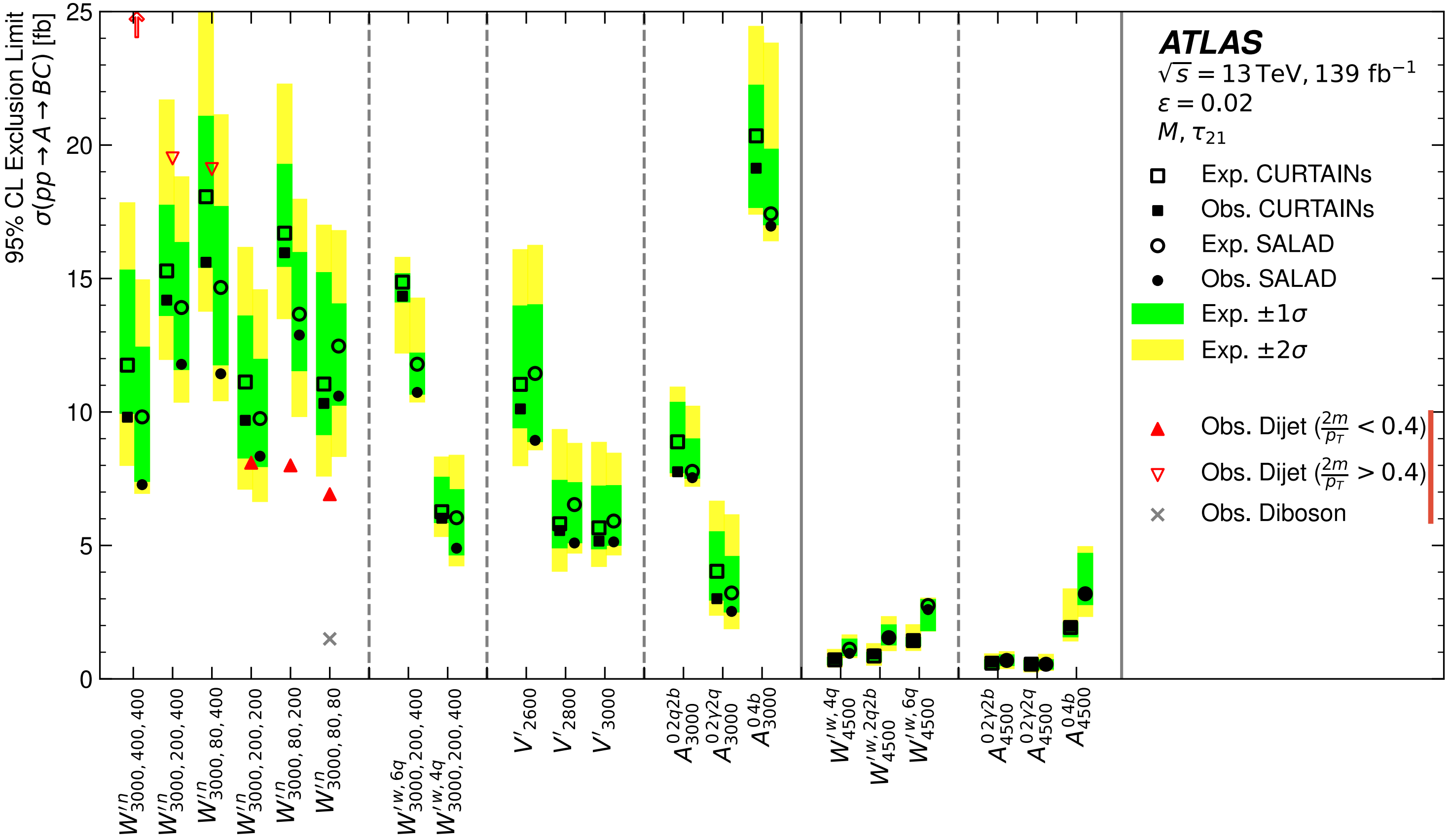
CURTAINs





Results: Signal Specific Limits

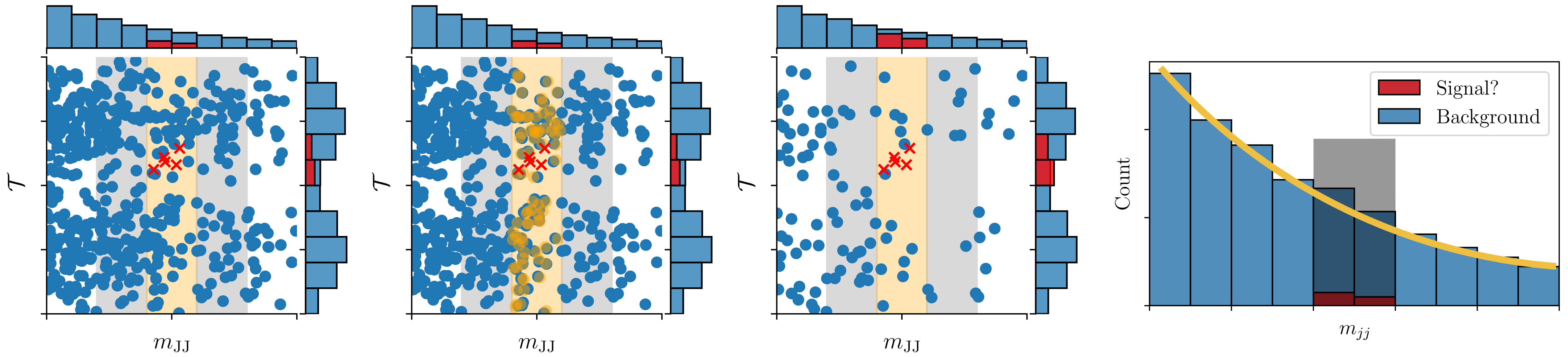
- Set limits @ 95% CLs to **20 investigated signal models**
- Analysis has a broad performance on many different models (better and/or more general than…)
- Similar performance for SALAD and CURTAINs



past ATLAS analyses

Conclusion

- New Anomaly Detection Analysis on ATLAS Run 2 Data
- Targeting dijet events with large radius jets
- Using four-step analysis strategy (see below)
- Results:
  - No significant excess observed
  - Set limits on many different signals



Backup

About me

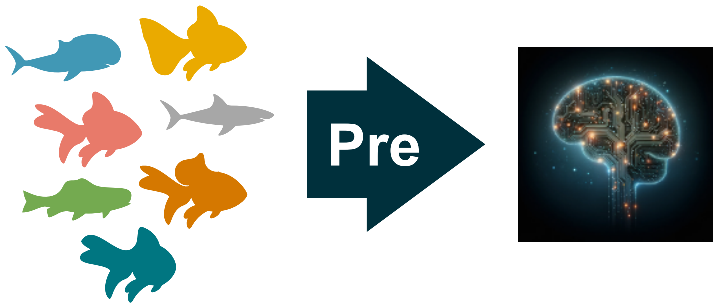
- Second year postdoc at Berkeley Lab
- Data analysis for CERN experiments since > 7 years

- **Physics**: Higgs, Anomaly Detection
- **Deep Learning**: Supervised, Unsupervised, Reinforcement
- **Computing**: Fast O(TB) Data Processing & Computing Pipelines

Anomaly Detection Strategies

Semi Supervised

- **Data with some labels**
- Use potential signal models as prior



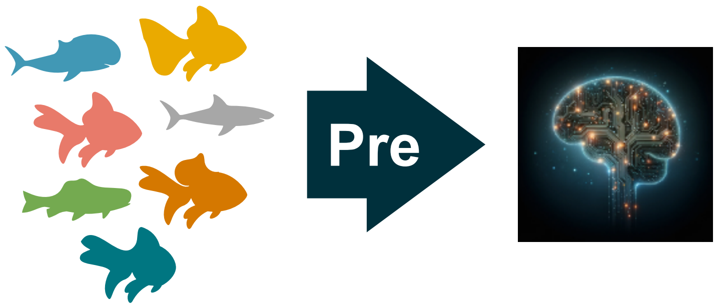
- **Challenge**: Need to choose (potential) signal models

Choose your degree of supervision

Anomaly Detection Strategies

Semi Supervised

- **Data with some labels**
- Use potential signal models as prior



- **Challenge**: Need to choose (potential) signal models

Weakly Supervised

- **Data with noisy labels**
- Use domain knowledge to place weak assumptions on the data

$$R(x) = \frac{p_{data}}{p_{bg}}$$

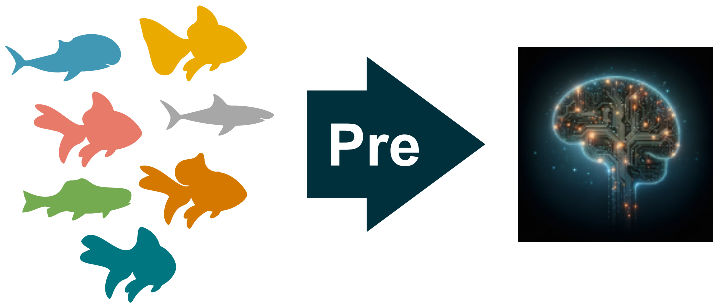
- **Challenge**: Weak assumptions need to be good

Choose your degree of supervision

Anomaly Detection Strategies

Semi Supervised

- **Data with some labels**
- Use potential signal models as prior



- **Challenge**: Need to choose (potential) signal models

Weakly Supervised

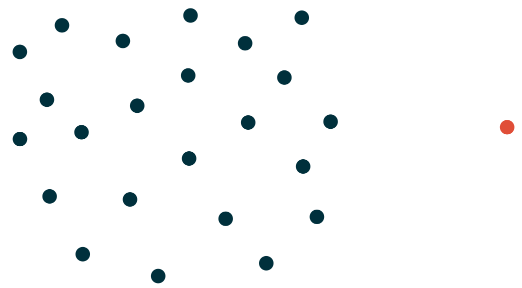
- **Data with noisy labels**
- Use domain knowledge to place weak assumptions on the data

$$R(x) = \frac{p_{data}}{p_{bg}}$$

- **Challenge**: Weak assumptions need to be good

Unsupervised

- **Data without labels**
- Often uses a "clever trick", like e.g. compressing SM-like events (Autoencoder)



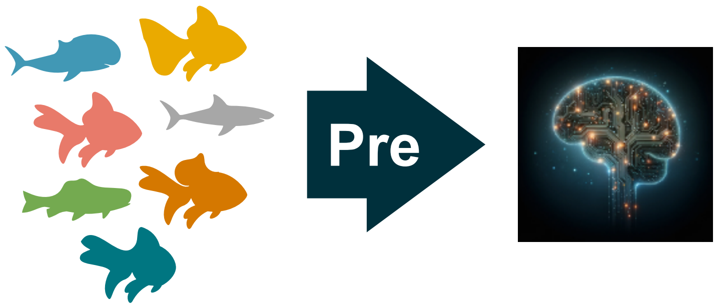
- **Challenge**: Get anomalous events but often no background estimate

Choose your degree of supervision

Anomaly Detection Strategies

Semi Supervised

- **Data with some labels**
- Use potential signal models as prior



- **Challenge**: Need to choose (potential) signal models

Weakly Supervised

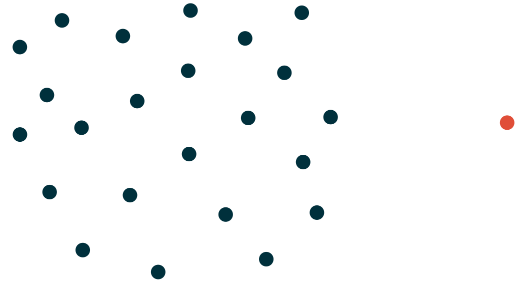
- **Data with noisy labels**
- Use domain knowledge to place weak assumptions on the data

$$R(x) = \frac{p_{data}}{p_{bg}}$$

- **Challenge**: Weak assumptions need to be good

Unsupervised

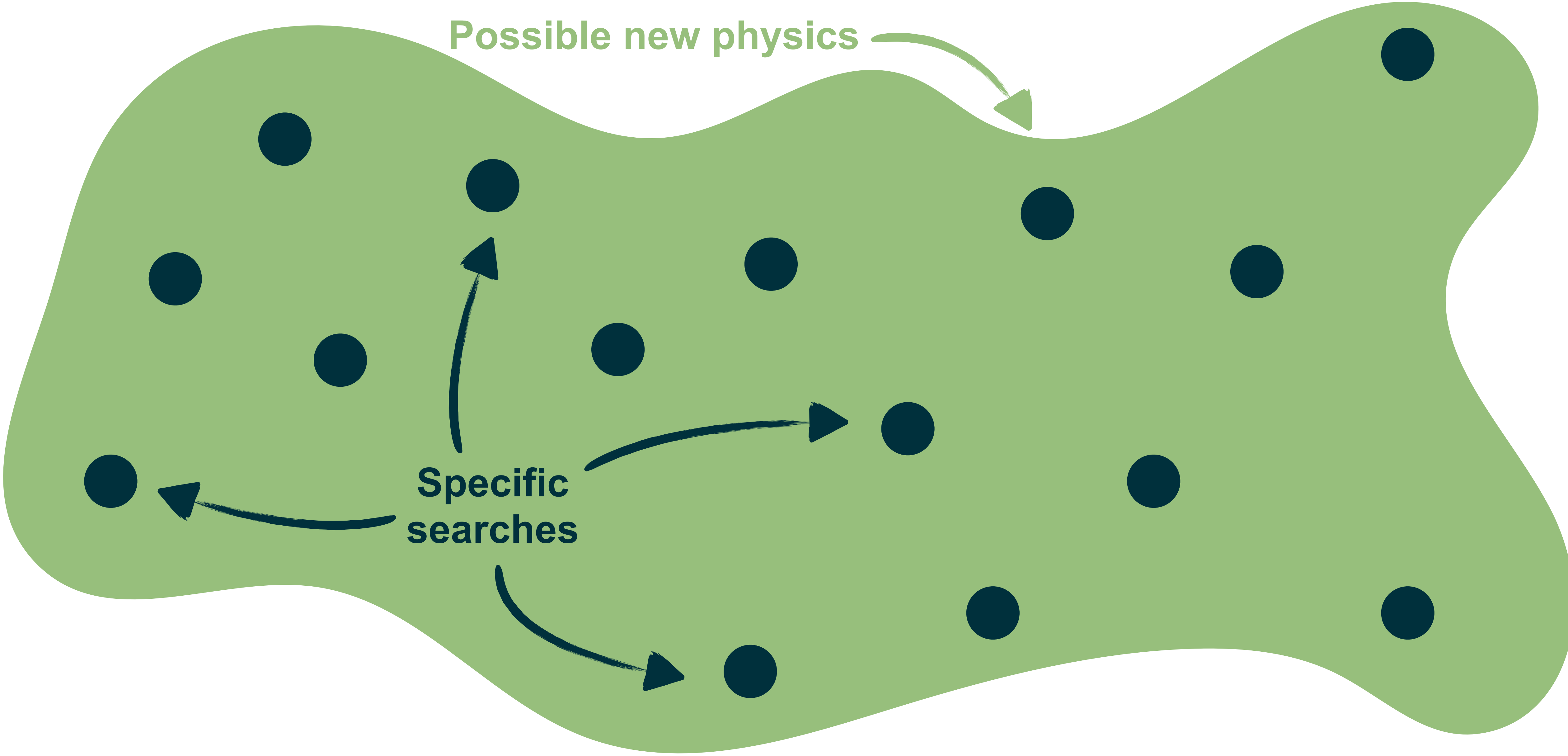
- **Data without labels**
- Often uses a "clever trick", like e.g. compressing SM-like events (Autoencoder)



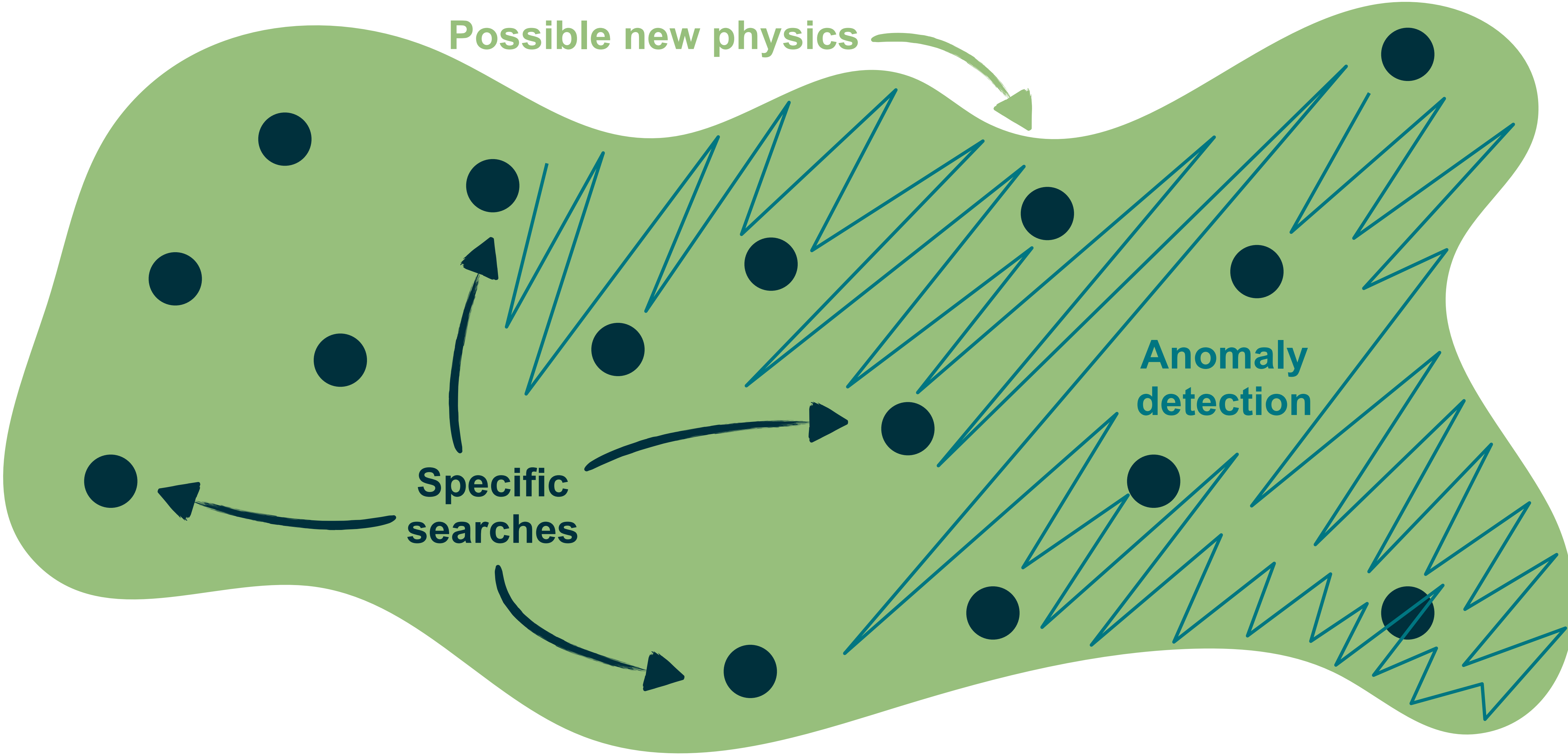
- **Challenge**: Get anomalous events but often no background estimate

Choose your degree of supervision

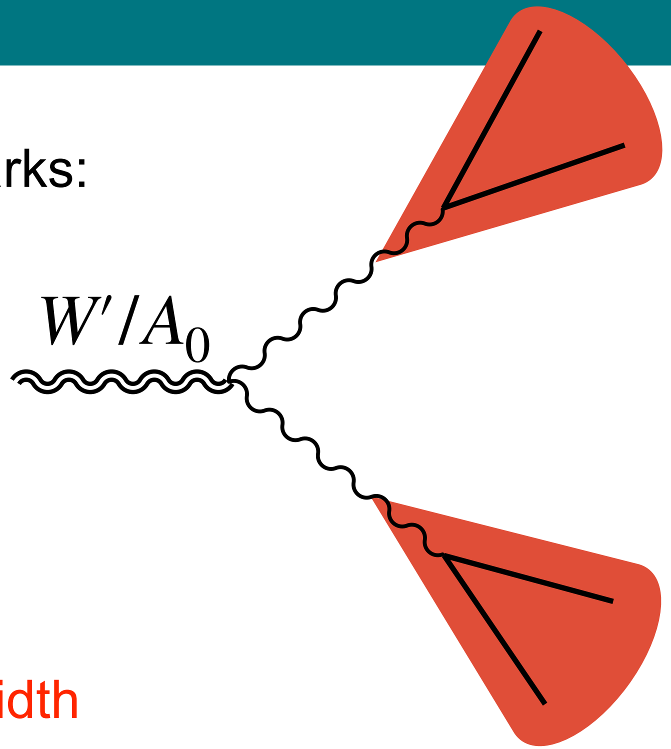
Anomaly Detection in the Bigger Picture



Anomaly Detection in the Bigger Picture



Used signal models



- Analysis is signal agnostic, but 20 signal models used as benchmarks:
  - Vector bosons: $W' \to W''Z'' \to$ qqqq
  - Pseudoscalar bosons: $A_0 \to H''Z' \to$ qqqq
  - Varying masses of parent and children particles

spectrum width

parent name

decay channel

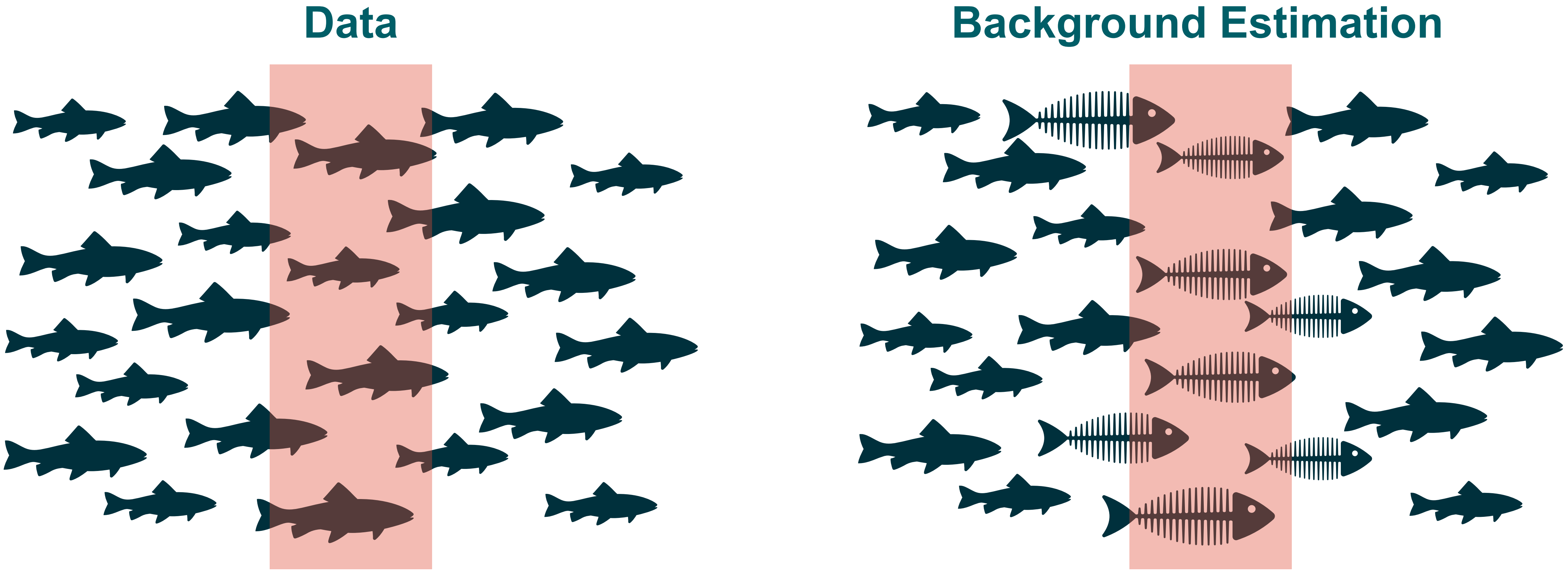
$$P^{s,d}_{p,a,b}$$

children masses

parent mass

Background Sculpting

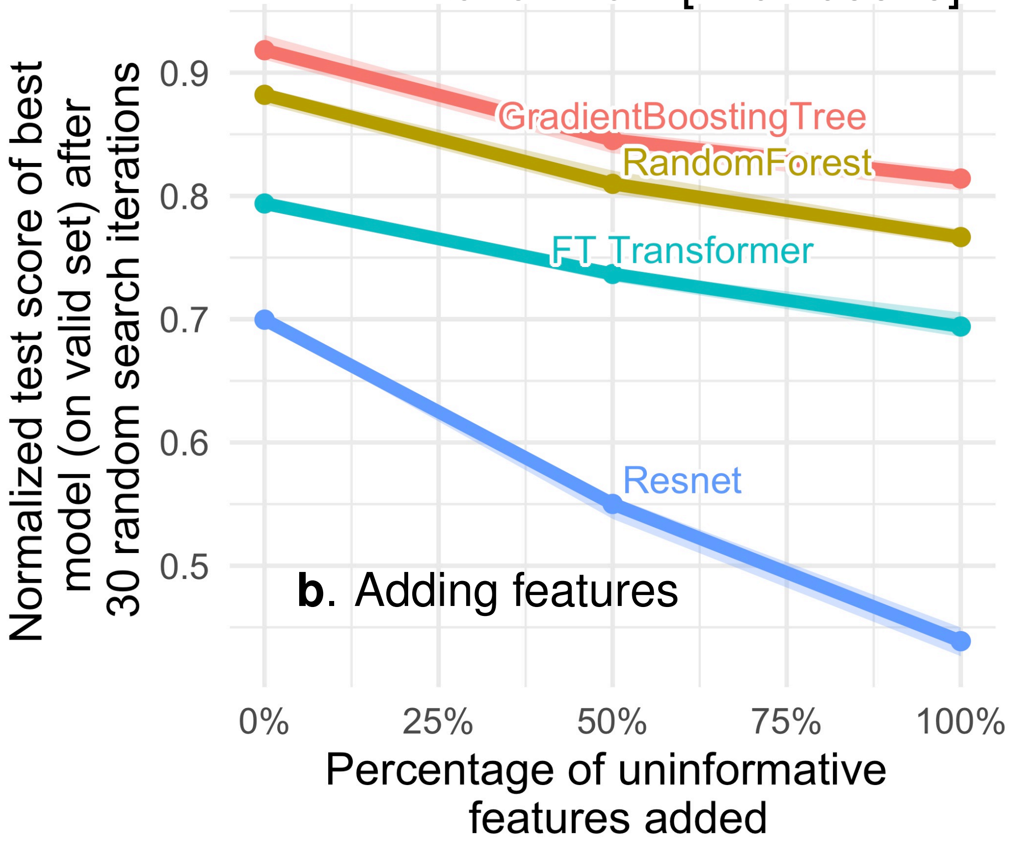
- If background estimation correlated with the regions:
  - Classifier gets sensitive to estimation itself (regions)
  - Sculpts false bump even if there is no signal



A Word on Uninformative Features

- Uninformative features can decrease classifier sensitivity
- This effect is architecture dependent (e.g. tree models vs neutral networks)
- It is not clear a priori which features yield sensitivity

taken from [2207.08815]

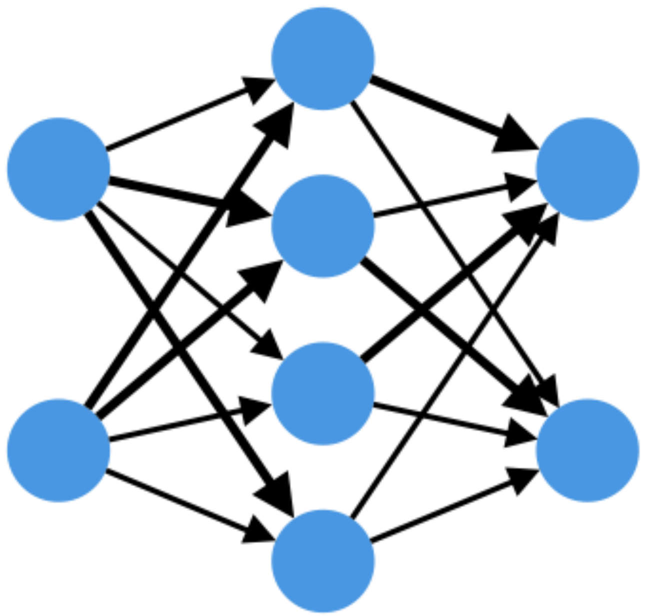


Classification of a Potential Signal (Specifics)

- Tested features of two AK10 jets (j):
  - Jet mass (M_j)
  - Jet substructure (τ_{21j}, τ_{32j})
- Networks: 3 x 64 ReLu, dropout (5%), Adam(lr=0.001), Early stopping (10 epochs)
- Five-fold cross validation to effectively use all data for training

Scan three subsets:

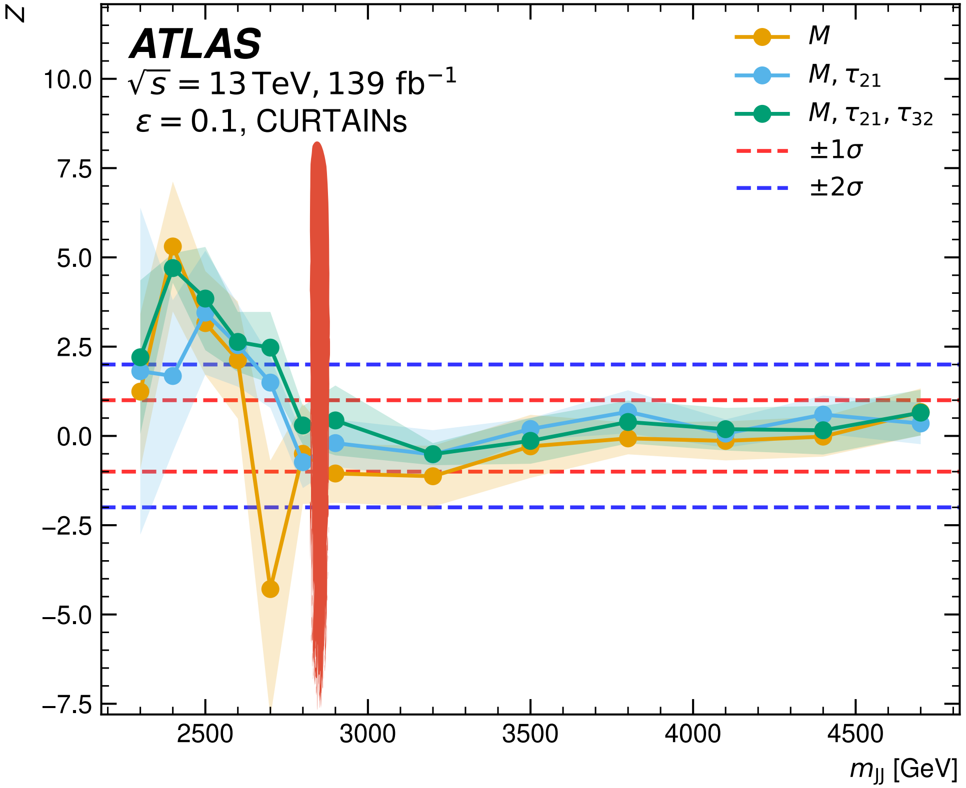
$\{M_j\}$ $\{M_j, \tau_{21j}\}$ $\{M_j, \tau_{21j}, \tau_{32j}\}$



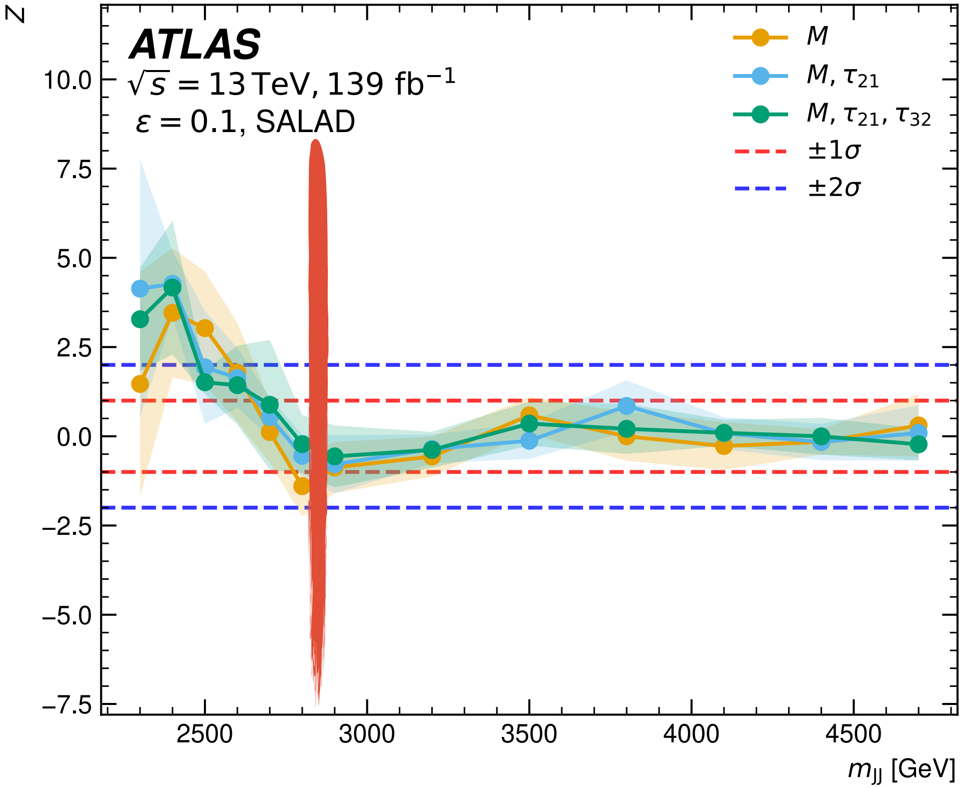
Validation Performed on all Test Datasets

- Each m_{JJ} signal region validated on all test datasets (significance and spread < 1)
- Results: **Successful for m_{JJ} ≥ 2900 GeV**
- Reason: **More challenging background estimation for small m_{JJ}**

Down-Up-Sampling



Down-Up-Sampling



Setting Limits (Signal Injection Procedure)

Standard Analysis

$$N_{sig} = \sigma \times \mathscr{L} \times A \times \epsilon$$

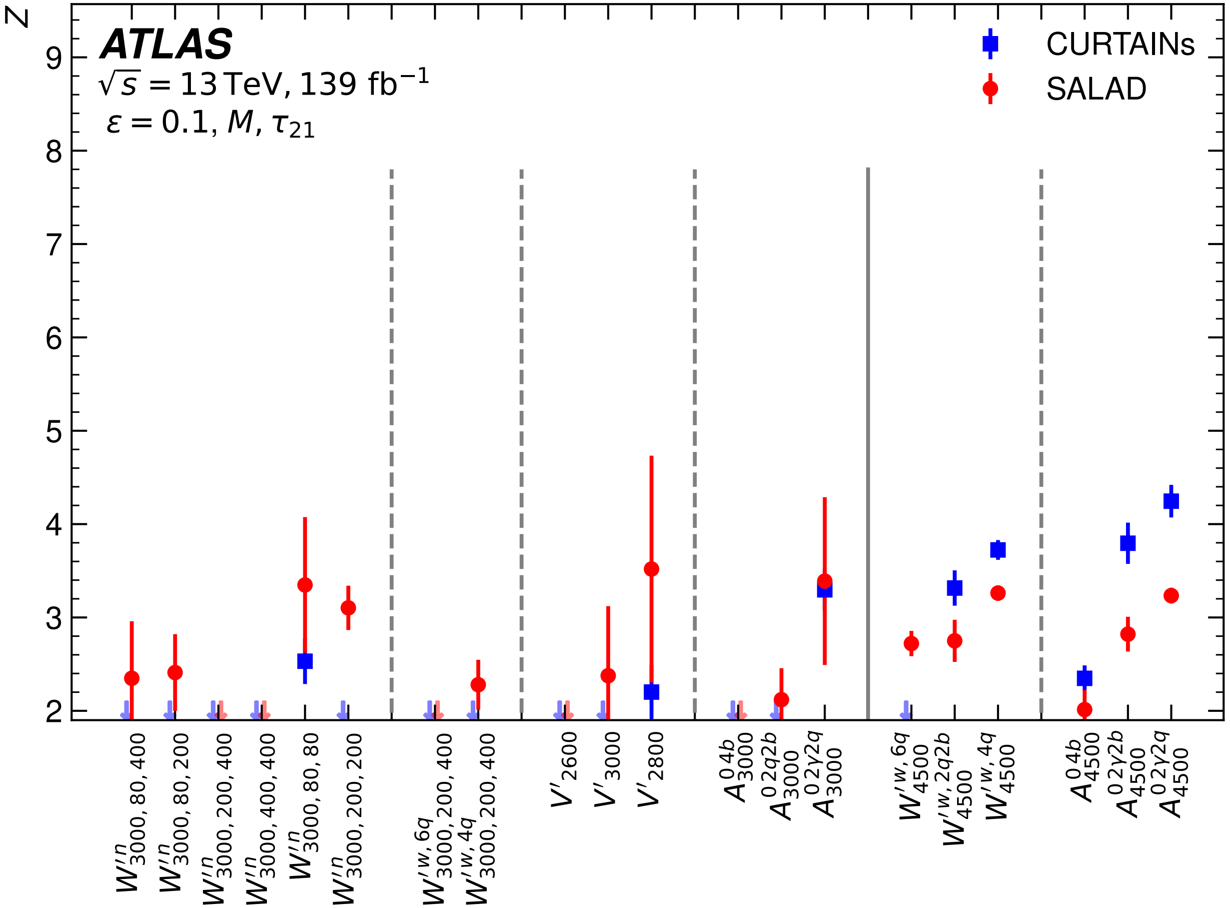
Anomaly Detection

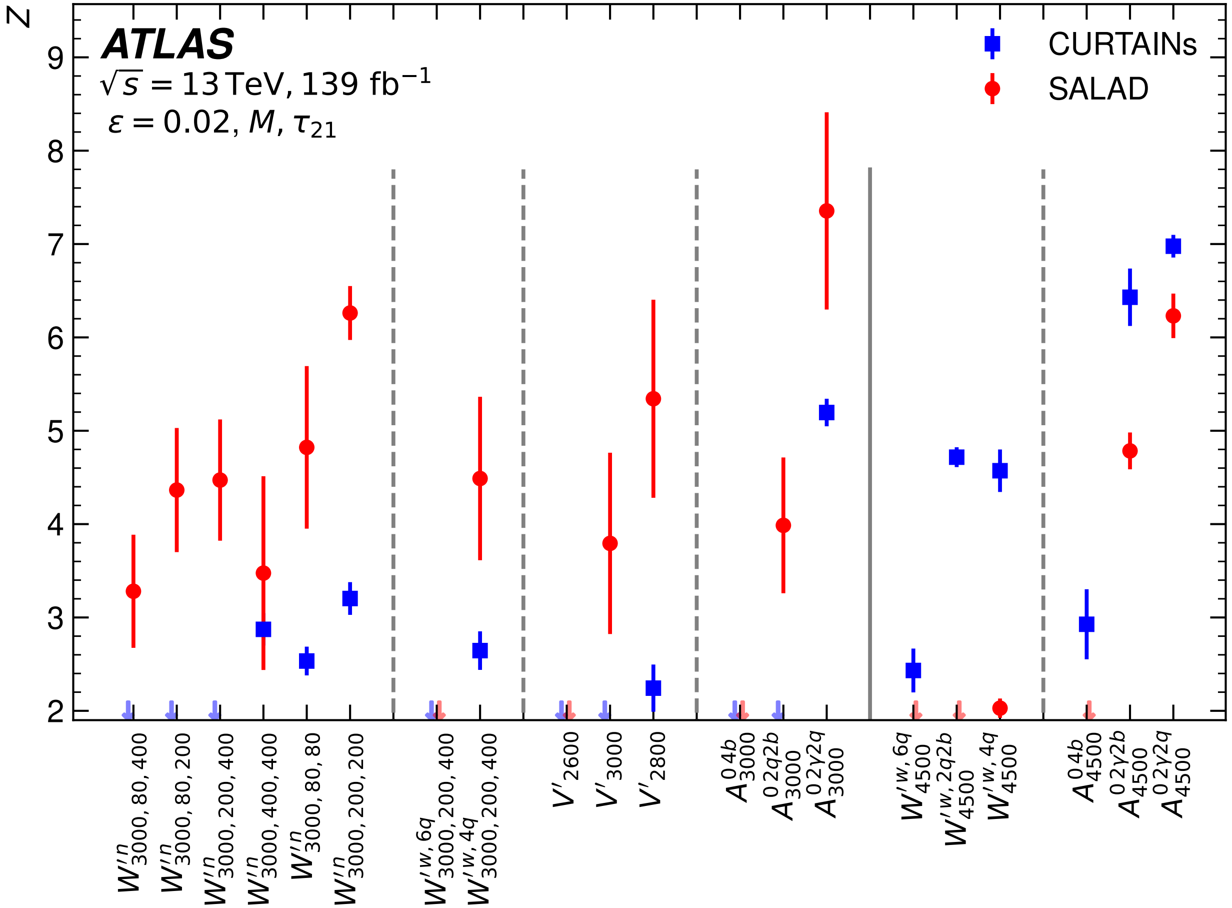
$$N_{sig} = \sigma \times \mathscr{L} \times A \times \epsilon(\sigma)$$

- Efficiency of anomaly detection usually depends on the amount of signal
  - The more signal, the more efficient the analysis (i.e. CWOLA)
- To get upper limit, need to inject signal, and rerun the whole analysis chain until $N_{sig} \propto N_{CL95\%}$

Results: Possible Signal Enhancement

- Inject 3σ signals (inclusive) into signal region & measure significance after analysis
- Sensitivity to many signals is improved by analysis
- SALAD (CURTAINs) better for low-mass (high-mass); Tighter efficiency (2%) more sensitive





Results: Comparison of Features

- Different feature sets have different sensitivity (more not always better)
- Strictest limits are observed (mostly) for the M, τ_{21} feature set with ε = 2%
- Scan over different feature sets is one of the strengths of this analysis