Unsupervised Anomaly Detection for LHC Collision Data Using a Transformer-Based Autoencoder.

ASRITH KRISHNA RADHAKRISHNAN
UNIVERSITY OF BOLOGNA, INFN, ITALY

SUPERVISORS:

- *PROF. MAXIMILIANO SIOLI,*
- *PROF. LORENZO RINALDI*

EUROPEAN COALITION FOR AI IN FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2026

Motivation

PIPELINE: Unsupervised Anomaly detection to identify candidates for new physics

- ❑ Dedicated searches remain the most sensitive for targeted BSM signatures, but cannot provide complete coverage of possible new-physics scenarios.
- ❑ This motivates a complementary, fully data-driven and model-independent approach to identify unexpected event topologies directly from data.
- ❑ We develop a fully unsupervised anomaly detection pipeline trained exclusively on ATLAS Run-2 open data

MAIN MOTIVATION OF MY STUDY IS WHAT IF WE LOSE THE REAL BSM DUE TO SELECTION CUTS?

- ❑ **Trigger selection? Quality cuts? Specific event selection cuts?**
- ❑ **Minimal selection, No final state hypothesis, No kinematic cut , UNSUPERVISED, DATA DRIVEN**
- ❑ **Which makes the model completely different from conventional model.**
- ❑ **Mixture topology, mixture decay channels,**

WHATS NEW?

- ❑ **A transformer autoencoder trained only on DATA, with a novel unsupervised anomaly score**
- ❑ **A clusterization based multi-engine pipeline that mines the distinct classes of BSM signatures.**

A NEW AE?

A SET transformer Auto encoder?

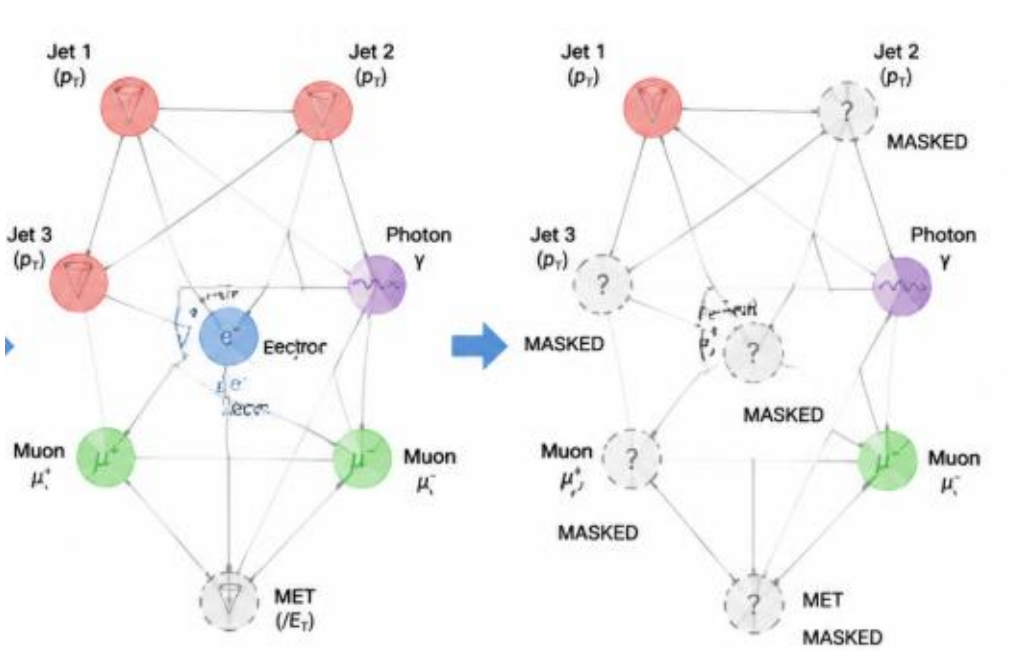
- **Hybrid representation:** Object-level representation combined with tabular event-level variables.

OBJECTS?

- **pT -ordered objects:** Jets, photons, leptons, large-R jets, and MET areordered by pT. The architecture remains permutation invariant.
- **Event-level variables:** Njets , Nleptons , Nγ , HT , and event density.

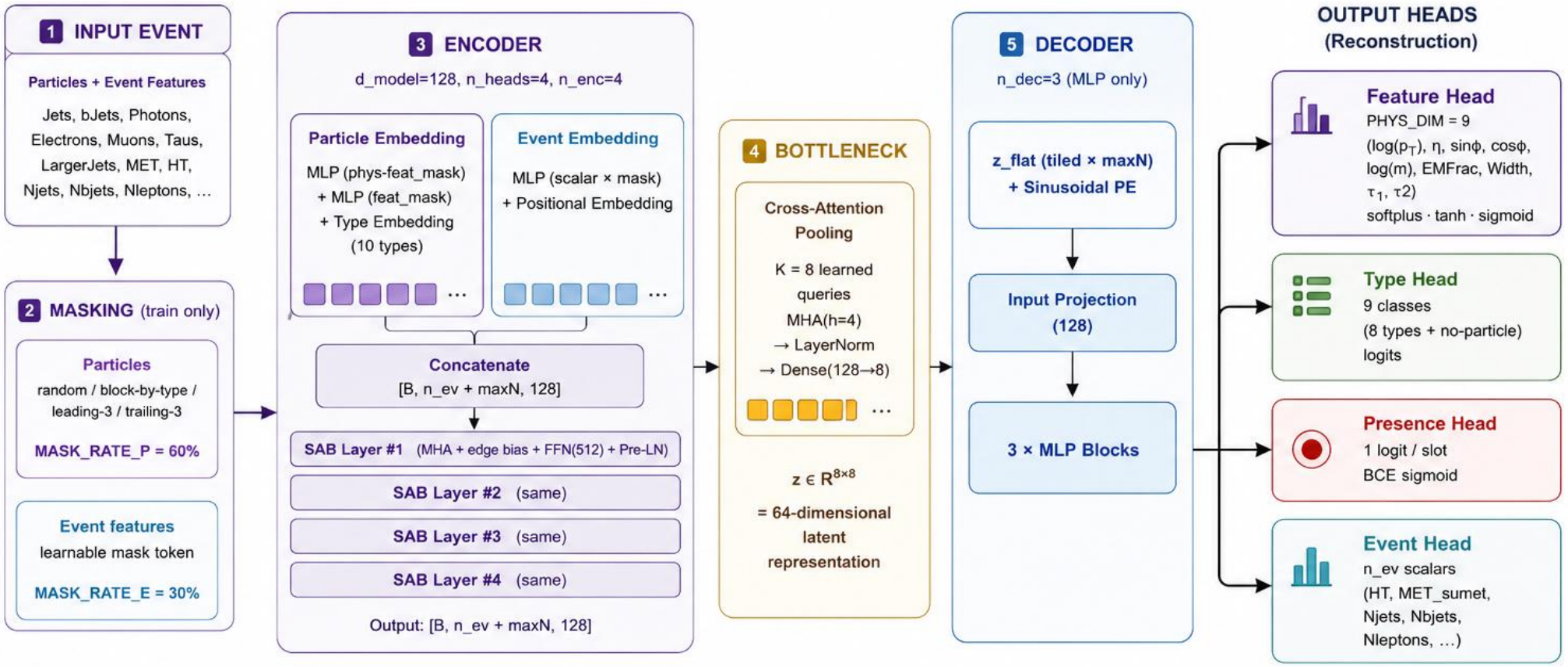
PATTERN AND CORRELATION LEARNING?

- **Random object masking: OBJECT LEVEL MASKING .6, EVENT LEVEL .3.**
- **Geometry-aware attention:** Pairwise interaction (Δη, Δϕ, ΔR, inv M)
- **Encoder-focused learning:** An **MLP decoder,** object relationships in the latent representation.
- **Physics-informed symmetry learning:** Azimuthal rotation (ϕ),enforce physics symmetries.



- **RUN 2 data released by ATLAS 2015-2016 Open Data for Research from the ATLAS experiment DOI:10.7483/OPENDATA.ATLAS.9HK7.P5SI**

- Quality control: Removed events with no candidates.
- Not remove soft particle detail, detector defect



- Tested on RUN2 data for 33 fb-1 of data.
- Training loss is hubber loss, comparing with MAE, MSE, the hubber gave better reconstruction accuracy to trained background data
- Bottleneck: pooling-by-multihead-attention (PMA) with K=8 learned query slots, projecting the variable-length object set down toa fixed 8-dimensional latent code z per event

Anomaly Score

NEW ANOMALY SCORE?

- Needed a anomaly score combines complementary local and global event information.
- Reconstruction residuals quantify deviations from the learned background representation.
- The calibrated score is used for unbiased event selection across different event topologies.
- Each score is first calibrated using an exact-N empirical probability integral transform (PIT) to obtain multiplicity-independent p-values. Score also decorrelated again HT as well to make sure model is not over sensitive to scales.
- The calibrated p-values are combined using Fisher's method and recalibrated with a second exact- PIT, producing the final anomaly score with a uniform false-positive rate across multiplicities.

$$S_{\text{recon}} = \frac{\sum_i w_i\, m_i\, (x_i - \hat{x}_i)^2}{\sum_i w_i\, m_i},$$

$$\mathbf{u} = \Lambda^{-1/2} P^T (\mathbf{x} - \boldsymbol{\mu}),$$

$$D^2_{\text{global}} = \sum_{i=1}^{d} \lambda_i^{-(1+\beta)} u_i^2,$$

$$p_i = 1 - F_{\text{bkg}}(D_i), \qquad T_i = -2\ln(p_i),$$

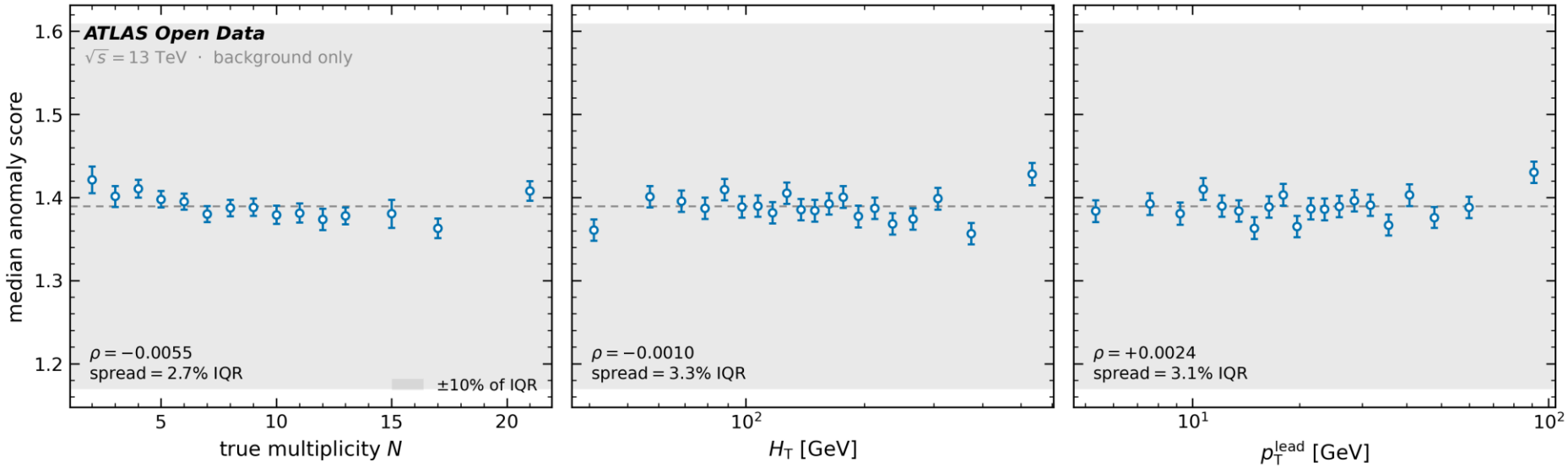
$$S_{\text{hybrid}} = \max(T_{\text{recon}}, T_{\text{global}}) + \beta\,(T_{\text{recon}} + T_{\text{global}})\,.$$

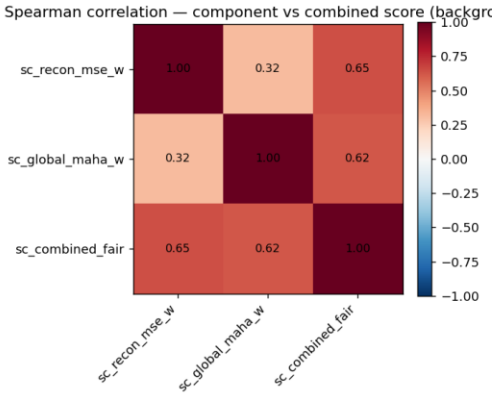
Stage 1: $p_i = \dfrac{\#\{s_{\text{ref}} \geq s_i \,|\, N\} + 0.5}{m_N + 1}$

Stage 2: $S_{\text{Fisher}} = -2(\ln p_{\text{recon}} + \ln p_{\text{maha}})$

Stage 3: $\text{Score} = -2\ln\left(\dfrac{\#\{S_{\text{ref}} \geq S_{\text{Fisher}} \,|\, N\} + 0.5}{m_N + 1}\right)$

Per-N empirical PIT yields an N-independent score; $p_{\min} = \dfrac{0.5}{m_N + 1}$.

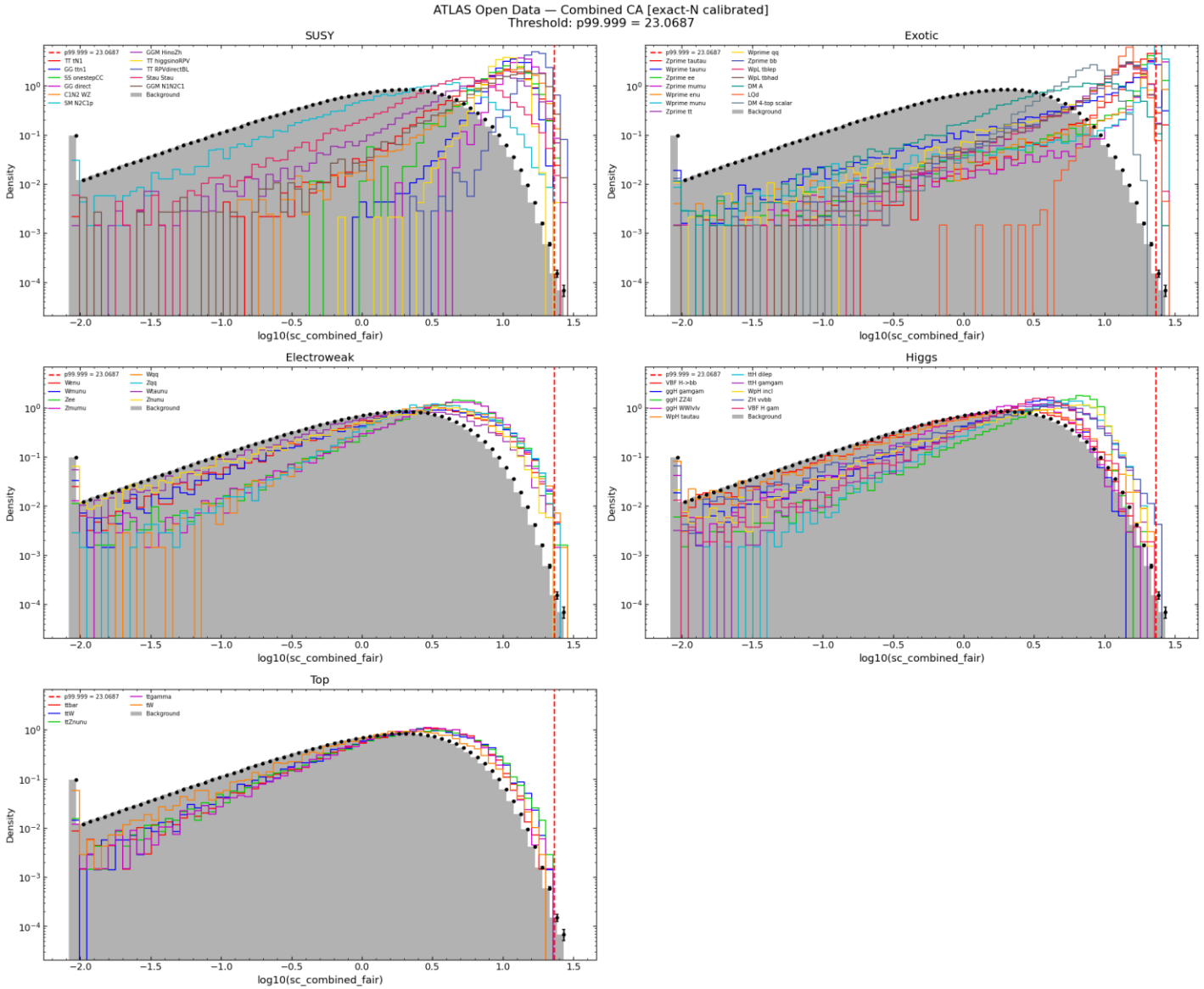




Anomaly score distributions

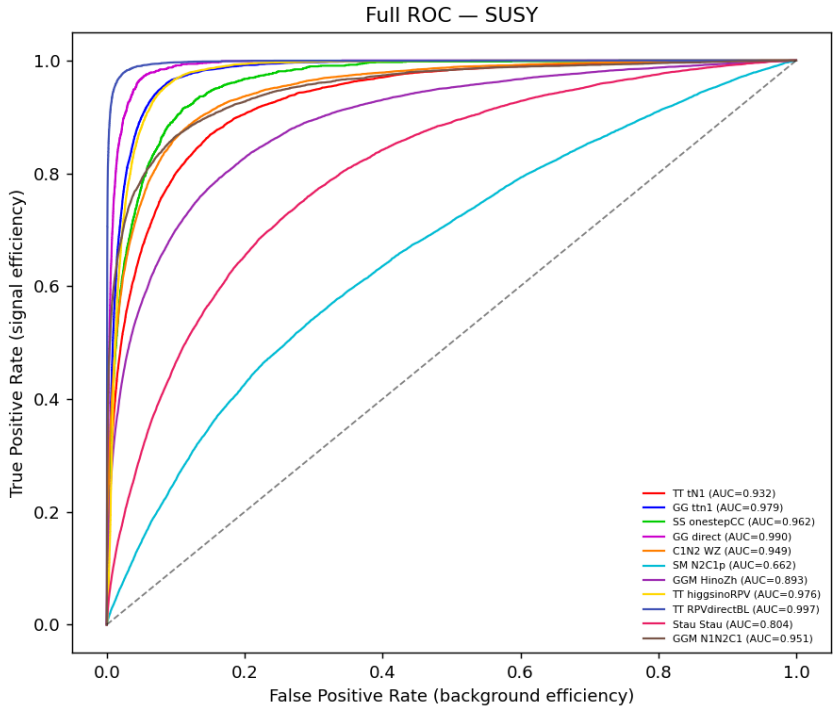
- **Validation:** Benchmarked using injected BSM signals (Z', W', leptoquarks, SUSY, DM, RPV) with **injected events**.

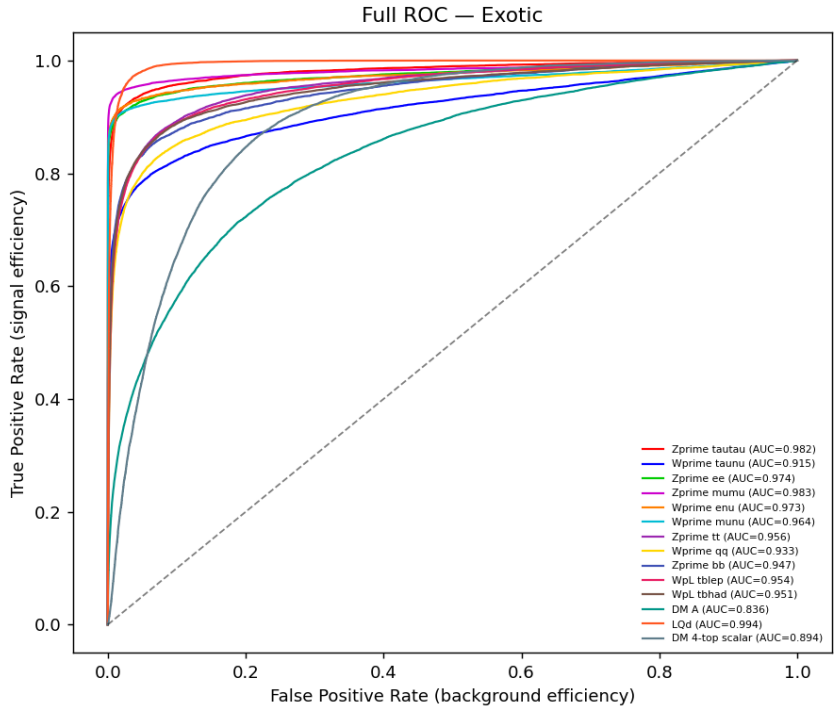
- **Performance:** High-mass resonances are efficiently retained (**Z'→μμ: 56–80%**, **W'→eν/μν: 50–68%**), while compressed/SM-like signals have intentionally low acceptance.

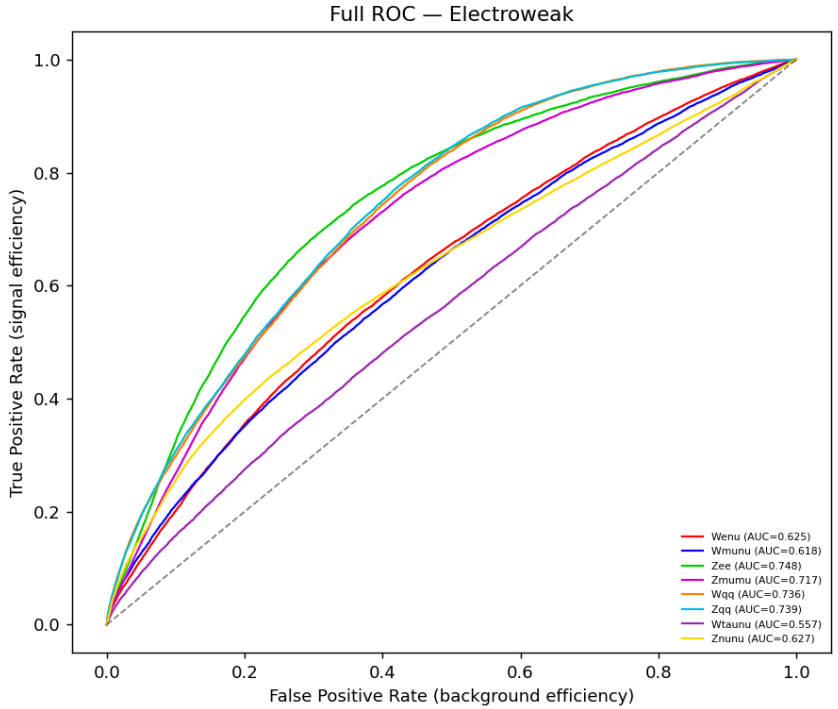


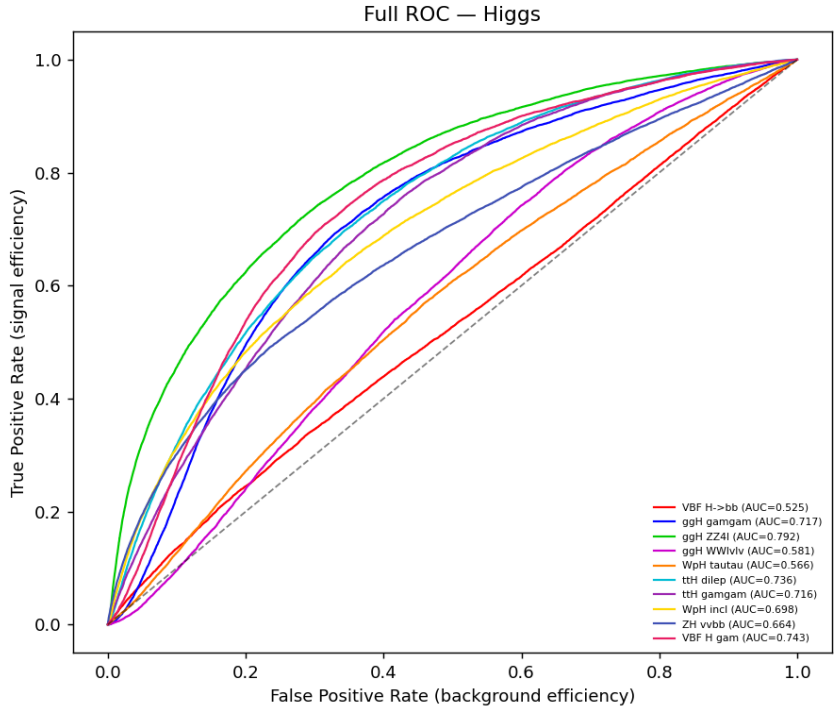
ROC curves

ROC curves computed for full space, 10k signal events against 5 million run2 samples.









Signal injection studies

- Most anomalous 100k events.

Signal	N=10	N=25	N=50	N=100	N=500	N=10
A14N23LO_TT_tN1_1200_200_MS_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	8/500 (1.6%)	15/1000 (
A14N_GG_ttn1_2000_5000_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	1/1000 (
14N23LO_SS_onestepCC_1600_900_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	2/100 (2.0%)	7/500 (1.4%)	8/1000 (
GG_direct_2400_200_0	1/10 (10.0%)	2/25 (8.0%)	3/50 (6.0%)	8/100 (8.0%)	17/500 (3.4%)	30/1000 (
4N23LO_C1N2_WZ_500p0_100p0_3L_2L7_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	2/100 (2.0%)	12/500 (2.4%)	22/1000 (
LO_SM_N2C1p_105_100_2L2MET75_MadSpin_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (
23LO_C1N2N1_GGMHinoZh50_800_noFilter_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	6/500 (1.2%)	12/1000 (
A14N23LO_TT_higgsinoRPV_1075_700_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (
A14N23LO_TT_RPVdirectBL_1400_0	2/10 (20.0%)	3/25 (12.0%)	11/50 (22.0%)	18/100 (18.0%)	77/500 (15.4%)	143/1000 (
StauStauDirect_200p0_1p0_TFilt_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	2/500 (0.4%)	3/1000 (
A14NNPDF23LO_GGM_N1N2C1_950_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	5/100 (5.0%)	34/500 (6.8%)	63/1000 (
Zprime_NoInt_tautau_SSM3000_0	5/10 (50.0%)	11/25 (44.0%)	21/50 (42.0%)	41/100 (41.0%)	224/500 (44.8%)	459/1000 (
Wprime_taunu_SSM3000_0	2/10 (20.0%)	6/25 (24.0%)	12/50 (24.0%)	32/100 (32.0%)	160/500 (32.0%)	322/1000 (
MSTW2008LO_Zprime_NoInt_ee_SSM3000_0	5/10 (50.0%)	12/25 (48.0%)	32/50 (64.0%)	59/100 (59.0%)	315/500 (63.0%)	637/1000 (
STW2008LO_Zprime_NoInt_mumu_SSM3000_0	9/10 (90.0%)	21/25 (84.0%)	37/50 (74.0%)	78/100 (78.0%)	392/500 (78.4%)	792/1000 (
A14NNPDF23LO_Wprime_enu_SSM3000_0	8/10 (80.0%)	20/25 (80.0%)	35/50 (70.0%)	73/100 (73.0%)	378/500 (75.6%)	742/1000 (
14NNPDF23LO_Wprime_munu_SSM3000_0	6/10 (60.0%)	16/25 (64.0%)	36/50 (72.0%)	81/100 (81.0%)	400/500 (80.0%)	804/1000 (
A14NNPDF23LO_zprime3000_tt_0	1/10 (10.0%)	1/25 (4.0%)	1/50 (2.0%)	2/100 (2.0%)	16/500 (3.2%)	28/1000 (
A14NNPDF23LO_Wprime_qq_3000_0	1/10 (10.0%)	2/25 (8.0%)	4/50 (8.0%)	6/100 (6.0%)	33/500 (6.6%)	72/1000 (
A14NNPDF23LO_Zprimebb3000_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	7/100 (7.0%)	32/500 (6.4%)	69/1000 (
WpL_tblep_M3000_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	4/100 (4.0%)	26/500 (5.2%)	56/1000 (
WpL_tbhad_M3000_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	4/100 (4.0%)	17/500 (3.4%)	37/1000 (
LO_DMA_500_700_gq0p25_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	10/500 (2.0%)	18/1000 (
NLO_LQd_gstML_0p3_nonallhad_M1000_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	2/100 (2.0%)	17/500 (3.4%)	31/1000 (
LO_DM_4topscalar_p1000_c1_nonallhad_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (

AE model benchmark comparison study

- Model comparison:https://arxiv.org/html/2409.03065v1 **ADFilter model**
- Working points fixed to the anomaly regions of PRL 132 081801:

  $\varepsilon_B = 8.4\times10^{-3} / 8.4\times10^{-4} / 8.4\times10^{-5}$.
- **Leptonic (3 TeV):**

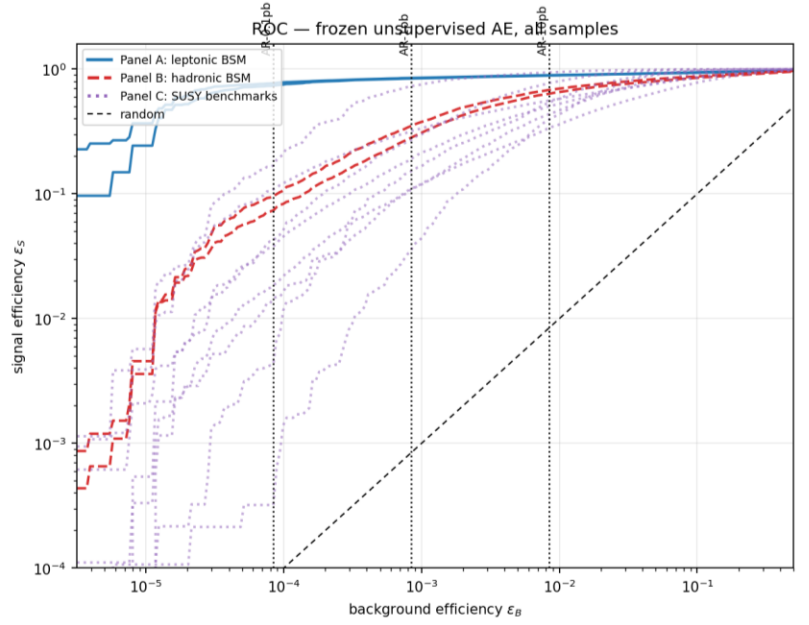
  W'→ev **0.85**, Z'→ee **0.84** at $\varepsilon_B = 8.4\times10^{-4}$ → SIC ≈ 29
- **Hadronic (3 TeV):**

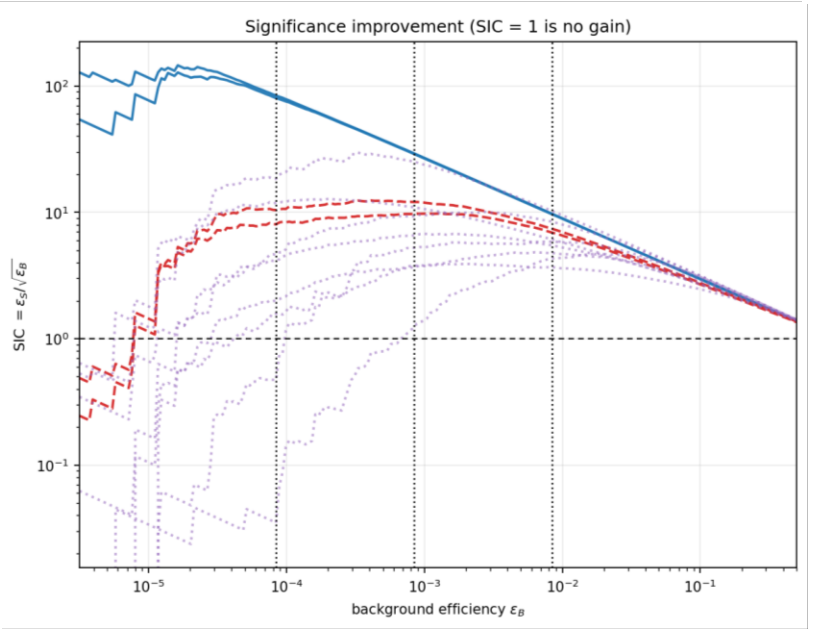
  Z'→$b\bar{b}$ **0.35**, W'→qq **0.28**
- **SUSY:**

  RPV stop 1400 **0.73**, GGM $N_1N_2C_1$ 950 **0.32**,

  gluino direct 2400 **0.27**, C_1N_2→WZ 500 **0.19**,

  gluino→$t\bar{t}N_1$ 2000 **0.15**, stop→tN_1 0.11, higgsino RPV 0.04





CLUSTERIZATION and scan Frame work

Engine	Input	Statistical search
E1 – Compact Clusters	SNN + Leiden + HDBSCAN	Multi-scale HDBSCAN density clustering + deduplication
E3 – Resonance	Physics variables	Sliding-window bump scan + NMS
E5 – Tail Search	Physics variables	Upper/lower tail scan + MMD validation
Endpoint	Kinematic observables	Endpoint/edge scan

Stage	Method	Result
Validation	Sequential MMD permutation test + SM-null calibration	Candidate significance; resolution-limited flag
Multiple testing	BH-FDR	Corrected *p* per engine
Cluster significance	SM-null calibrated cluster significance	Significance calibrated against SM null
Characterization	Cohen's *d*, neff, top-3 fraction, held-out labels	Candidate characterization
Physics	Interpretation + artifact flags	Physics / artifact assessment
Stability (E1)	10× bootstrap Jaccard	≈0.6

Stage 5: Statistical Validation

$$\mathrm{MMD}^2(X,Y) = \frac{1}{n^2}\sum_{i,j} k(x_i,x_j) + \frac{1}{m^2}\sum_{i,j} k(y_i,y_j) - \frac{2}{nm}\sum_{i,j} k(x_i,y_j)$$

$$p_{\mathrm{local}} = \frac{1 + \#(\mathrm{MMD}_{\mathrm{null}} \geq \mathrm{MMD}_{\mathrm{obs}})}{B+1}$$

$$p_{\mathrm{corr}} = 1 - (1 - p_{\mathrm{local}})^{N_{\mathrm{eff}}}$$

$$d = \frac{\mu_{\mathrm{candidate}} - \mu_{\mathrm{background}}}{\sigma_{\mathrm{background}}}$$

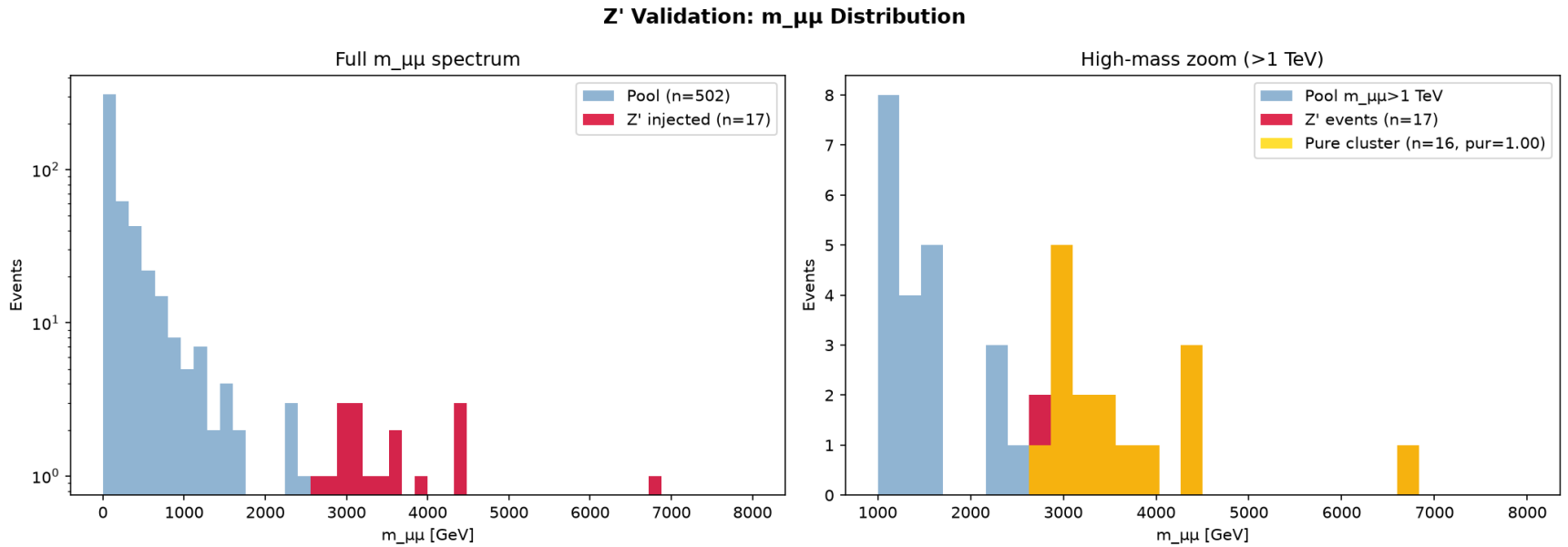
$$n_{\mathrm{eff}} = \exp\left(-\sum_i p_i \log p_i\right), \qquad p_i = \frac{d_i^2}{\sum_j d_j^2}$$

Results

Does the framework is working?

Signal injection studies

Injected Signal	Injected Events	Best Recall	Best Purity	Best Engine	Status
Z' mumu	30	24/25 (96%)	12/25 in 12 events (1.000)	Endpoint (recall), E5 (purity)	✓ Recovered
LQd	4	4/4 (100%)	3/4 in 22 events (0.136)	Endpoint (recall), E3 (purity)	✓ Recovered
Wprime_enu	20	17/20*	(1.000)	E3 (recall), E5 (purity)	✓ Recovered
C1N2_WZ	17	10/17	10/17 in 30 events	Endpoint (recall)	✓ Recovered



Conclusions

- **A fully unsupervised, model-independent pipeline** identifies anomalous event topologies with minimal selection bias.
- **Sequential MMD testing with SM-null calibration and multiple-testing correction** provides statistically controlled candidate selection.
- **Multiple complementary engines** probe compact clusters, resonances, tails, endpoints, and multiplicities.
- **Signal-injection studies recover representative BSM signatures**, demonstrating sensitivity across different anomaly topologies.
- The model give a path to isolate the anomalies along with recommending a signal region for search studies.



thanks

PHYSICS CHARECTERIZTION (supervised)

- **Accept statistically significant candidates** from Stage 5 after multiple-testing correction.
- **Identify the most discriminating physics variables** using Cohen's d to understand what drives each anomaly.
- **Characterize the candidate topology** using compactness, purity, effective dimensionality and cluster properties.
- **Assign possible BSM signatures** by comparing the observed variable patterns and topology with known physics scenarios (Z', SUSY, LLP, etc.).
- **Generate a complete candidate summary**, including significance, dominant variables, event composition and engine of origin.
- **Prioritize candidates into Tier-1, Tier-2 and Tier-3**, providing a final shortlist for detailed physics analysis.

Statistically validated anomaly candidates.

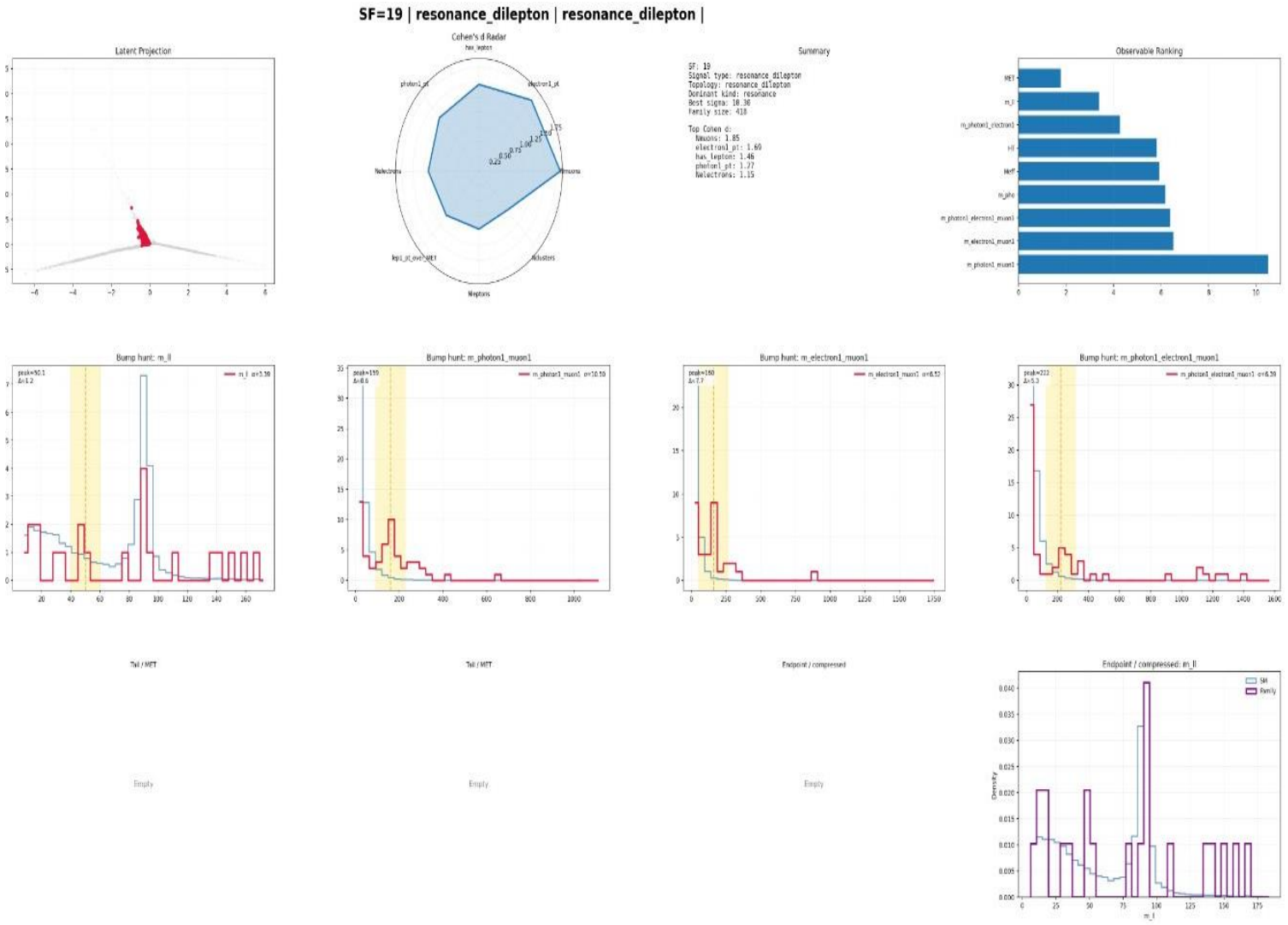
Physics-driven signal region definition.

Standard HEP background estimation (CR/Sidebands/MC).

Final discovery via profile-likelihood inference including systematic uncertainties.

Z to ll removed model

- To compare the model performance to recover SM events, I have conducted a veto study.
- Removed z to mumu samples from the training data and reperformed the analysis.
- AUC went up from .717 to .758.



ANALYSIS STRATEGY

- **RUN 2 data released by ATLAS 2015-2016 Open Data for Research from the ATLAS experiment DOI:10.7483/OPENDATA.ATLAS.9HK7.P5SI**

❑Preprocessing:

- Quality control: removed events with no candidates

❑ Data representation:

- low-level features(p_T, E, η, ϕ, q) :Jets(6 leading), Leptons (6 leading)
- Photons (6 leading), tau, Large R jets (6) , MET , Energy Density, HT,
- Jet width, EM frac, Tau Weta variables

❑ Data Processing Strategy (PhysLite Dataset)

- ❑ **Environment**: Executed on **CERN SWAN**, using ROOT with **Distributed RDataFrame (RDF)** and **Dask** .
- ❑ **Dataset**: Based on **PhysLite format**, totaling ~75 TB of ATLAS data across ROOT files.
- ❑ **Event Selection**: 5 **million events** used for full-scale training and validation.
- ❑ **Preprocessing**: Distributed RDF used to variable derivation in parallel processing.
  Physics variables of interest were extracted and stored in new ROOT files.

AUC curves

- AUC computed 10k signal against 5 million background samples

#	Category	Signal	ROC AUC	pAUC @FPR<0.01	pAUC @FPR<0.001
1	SUSY	TT RPVdirectBL	0.997	0.933	0.779
2	Exotic	LQd	0.994	0.860	0.639
3	SUSY	GG direct	0.990	0.766	0.581
4	Exotic	Zprime mumu	0.983	0.960	0.937
5	Exotic	Zprime tautau	0.982	0.917	0.843
6	SUSY	GG ttn1	0.979	0.659	0.521
7	SUSY	TT higgsinoRPV	0.976	0.595	0.503
8	Exotic	Zprime ee	0.974	0.934	0.904
9	Exotic	Wprime enu	0.973	0.941	0.914
10	Exotic	Wprime munu	0.964	0.939	0.914
11	SUSY	SS onestepCC	0.962	0.641	0.533
12	Exotic	Zprime tt	0.956	0.752	0.596
13	Exotic	WpL tblep	0.954	0.760	0.615
14	SUSY	GGM N1N2C1	0.951	0.743	0.626
15	Exotic	WpL tbhad	0.951	0.767	0.608
16	SUSY	C1N2 WZ	0.949	0.656	0.549
17	Exotic	Zprime bb	0.947	0.784	0.637
18	Exotic	Wprime qq	0.933	0.748	0.601
19	SUSY	TT tN1	0.932	0.622	0.533
20	Exotic	Wprime taunu	0.915	0.817	0.744
21	Exotic	DM 4-top scalar	0.894	0.521	0.501
22	SUSY	GGM HinoZh	0.893	0.607	0.534
23	Exotic	DM A	0.836	0.590	0.532
24	SUSY	Stau Stau	0.804	0.529	0.506

SM models

25	Higgs	ggH ZZ4l	0.792	0.517	0.500
26	Electroweak	Zee	0.748	0.510	0.502
27	Higgs	VBF H gam	0.743	0.501	0.500
28	Electroweak	Zqq	0.739	0.516	0.504
29	Higgs	ttH dilep	0.736	0.504	0.500
30	Electroweak	Wqq	0.736	0.516	0.504
31	Electroweak	Zmumu	0.717	0.509	0.502
32	Higgs	ggH gamgam	0.717	0.499	0.500
33	Higgs	ttH gamgam	0.716	0.506	0.501
34	Higgs	WpH incl	0.698	0.508	0.500
35	Higgs	ZH vvbb	0.664	0.514	0.502
36	SUSY	SM N2C1p	0.662	0.508	0.501
37	Electroweak	Znunu	0.627	0.511	0.502
38	Electroweak	Wenu	0.625	0.509	0.502
39	Electroweak	Wmunu	0.618	0.510	0.502
40	Top	ttZnunu	0.606	0.503	0.500
41	Top	ttbar	0.600	0.500	0.500
42	Top	ttgamma	0.600	0.500	0.500
43	Top	ttW	0.597	0.502	0.500
44	Higgs	ggH WWlvlv	0.581	0.498	0.500
45	Higgs	WpH tautau	0.566	0.500	0.500
46	Electroweak	Wtaunu	0.557	0.505	0.501
47	Top	tW	0.532	0.499	0.500
48	Higgs	VBF H→bb	0.525	0.501	0.500

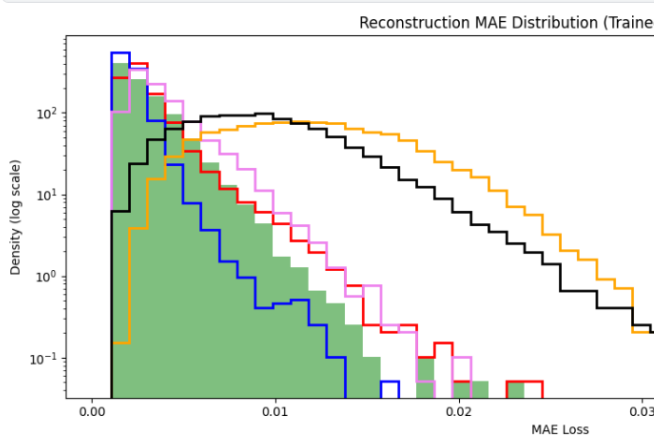
GITHUB repository

Folder / File	Contents
configs/	pipeline_config.json, README.md
docs/	OPEN_DATA.md, PIPELINE.md, assets/
examples/outputs/	atlas_quality_cut_summary.csv, stage5_shortlist.csv, stage5_significant.csv, stage6_inventory.csv, stage6_summary.csv, stage6_tier1.csv
notebooks/01_data_preparation/	rdataframe_to_root_background.ipynb, rdataframe_to_root_signal.ipynb
notebooks/02_autoencoder/	transformer_autoencoder_training.ipynb
notebooks/03_scoring/	FINAL_empirical_lambda2.ipynb, empirical_pipeline_with_sizestudy.ipynb
scripts/	run_stage3_scoring.sh, run_stage4_discovery.sh
src/hep_bsm/autoencoder/	model_definition.py
src/hep_bsm/discovery/	bridge.py, characterization.py, e1_pipeline.py, encoder_tokens.py, engine_names.py, engines.py, pipeline_config.py, significance.py
src/hep_bsm/scoring/	scan_combined_fair_tdf.py, scan_combined_fair_uproot.py
tests/	test_pipeline_invariants.py
root files	README.md, LICENSE, CITATION.cff, pyproject.toml, environment.yml, .gitignore

BSM search model comparison



- arXiv:2409.03065, the ADFilter. https://arxiv.org/abs/2502.14036 and other anomaly detection model uses signal selection method.
- ADFilter is really close to the unsupervised model, but for the RMM matrix representation it is selecting lepton trigger and selecting lepton pt > 60.
- I tried this method last summer, because of this lepton trigger on this, I used tried to create RMM for my method, but issue is while it is good for lepton triggered data, for full data the RMM is mostly empty and the performance was really bad,
- Used 20*20/10*10 RMM matrix, tried different method, tried Convolute layers, flattened layer to reduce computation complexity, none gave the better reconstruction for background itself,



RMM

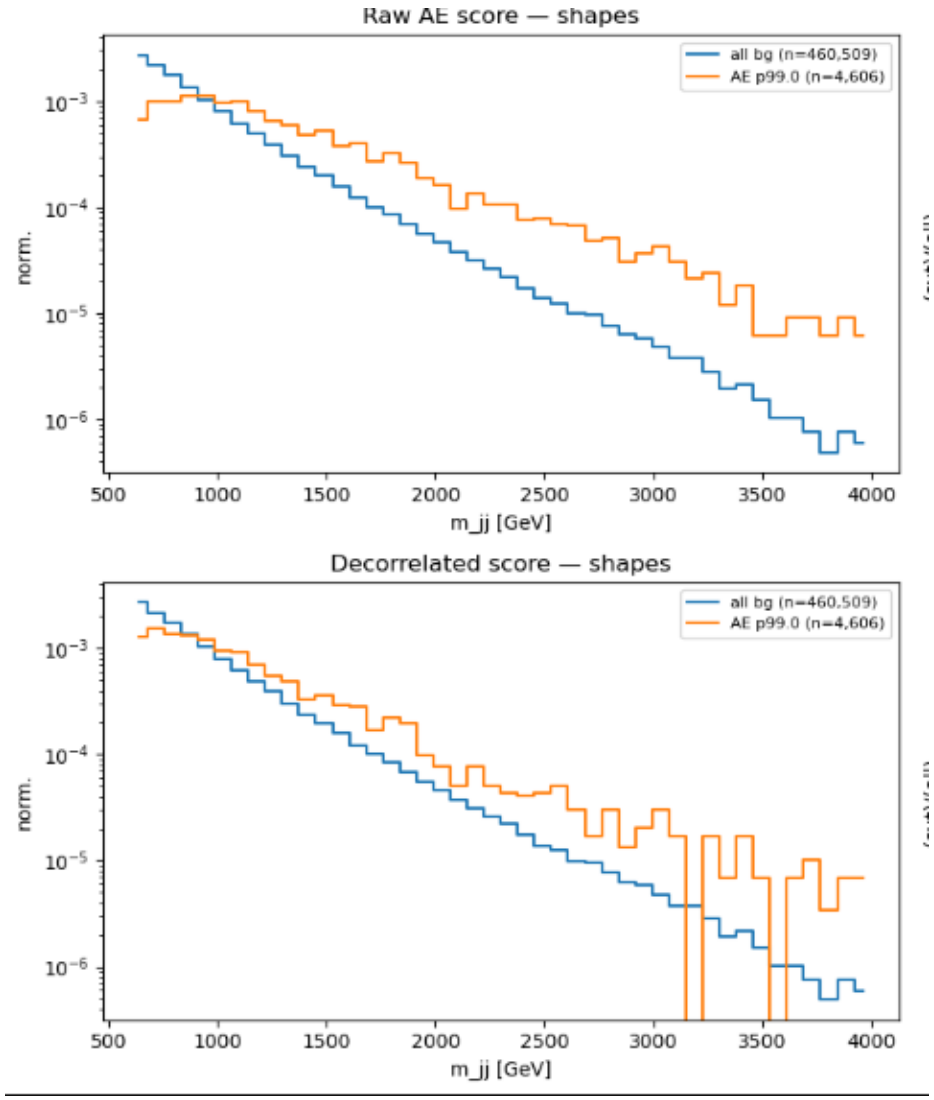
$$\begin{pmatrix}
e_T^{miss} & m_T(j_1) & m_T(j_2) & \dots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \dots m_T(\mu_N) \\
h_L(j_1) & e_T(j_1) & m(j_1,j_2) & \dots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \dots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta e_T(j_2) & \dots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \dots m(j_2,\mu_N) \\
\dots & \dots & \dots & \dots, & \dots & \dots & \dots \\
h_L(j_N) & h(j_1,j_N) & \dots & \dots \delta e_T(j_N) & m(j_N,\mu_1) & m(j_N,\mu_2) & \dots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \dots h(\mu_1,j_N) & e_T(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \dots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta e_T(\mu_2) & m(\mu_2,\mu_N) \\
\dots & \dots & \dots & \dots & \dots & \dots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \dots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta e_T(\mu_N)
\end{pmatrix}$$

$$\delta e_T(i_n) = \frac{E_T(i_{n-1}) - E_T(i_n)}{E_T(i_{n-1}) + E_T(i_n)}, \quad n = 2, \dots, N,$$

$$h_L(i_n) = C(\cosh(y) - 1)$$

https://arxiv.org/pdf/1810.06669

- AE anomaly scores use to fins anomalous events and reconstruct the invariant mass mjj.
- A mass-sculpting check verified that the AE selection preserved the background mjj shape.
- The background was estimated by fitting the mjj sidebands with a smooth polynomial.
- The fit was extrapolated into the signal region, compared with event yield.
- The bump hunt was repeated for different AE score thresholds.
- The optimal AE threshold was chosen by maximizing the local significance.
- The workflow uses methods in the **ATLAS ADFilter strategy**, replacing the original anomaly detector with the proposed autoencoder to check the AE performance.



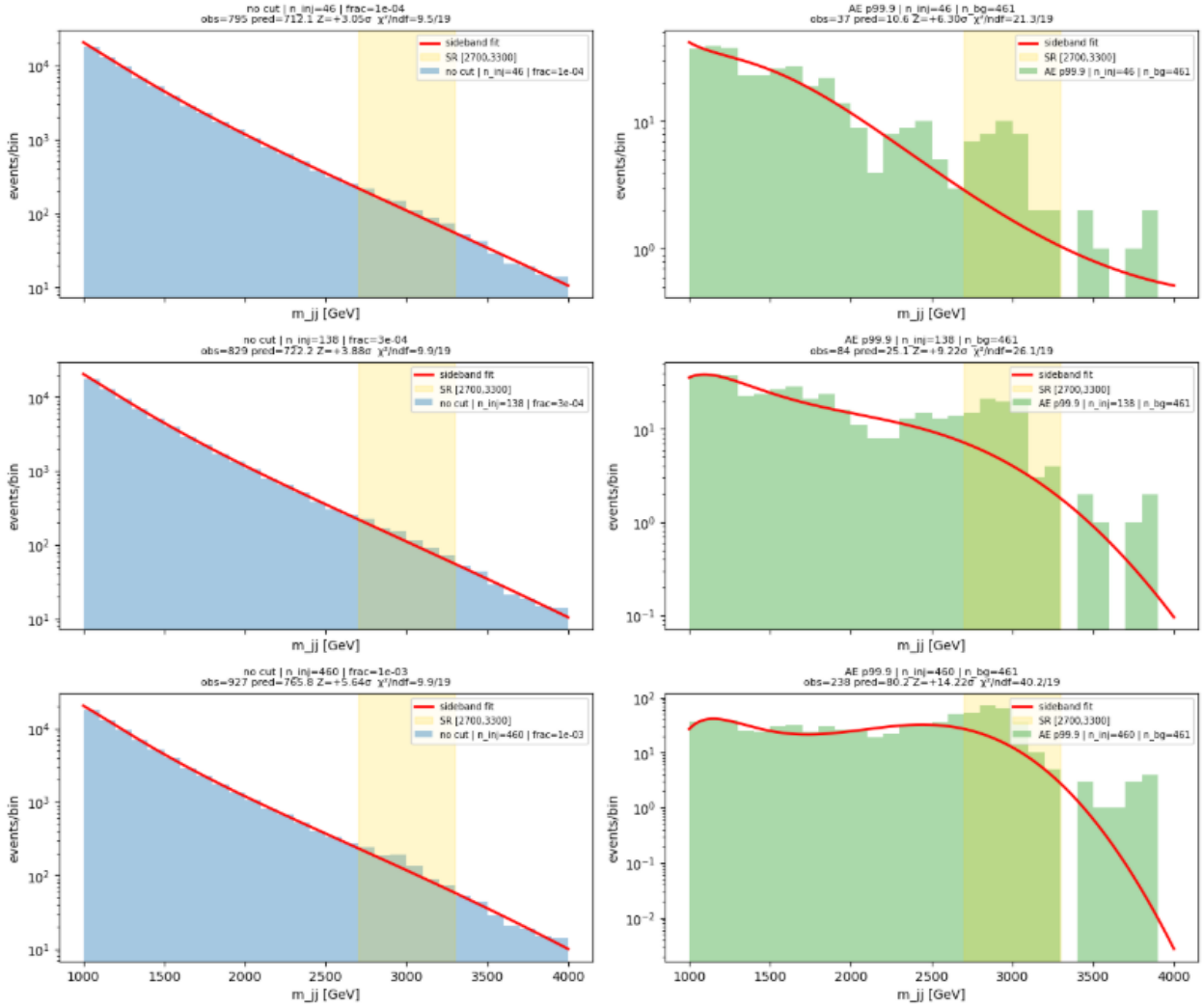
- Decorrelated mjj distribution.
- Significance table is given with different signal injection fraction

```
--- A14NNPDF23LO_Wprime_qq_3000_0 ---
     frac   n_inj   Z_nocut      p90.0      p95.0      p99.0      p99.5      p99.9
  1.0e-04      46      3.26       3.69       3.33       2.71       3.22       8.21
  3.0e-04     138      4.03       4.29       6.47       9.84       9.51       9.73
  1.0e-03     460      6.61      12.82      12.75      10.90      13.66      16.90
  3.0e-03    1381     12.51      20.33      24.14      21.11      23.56      17.53

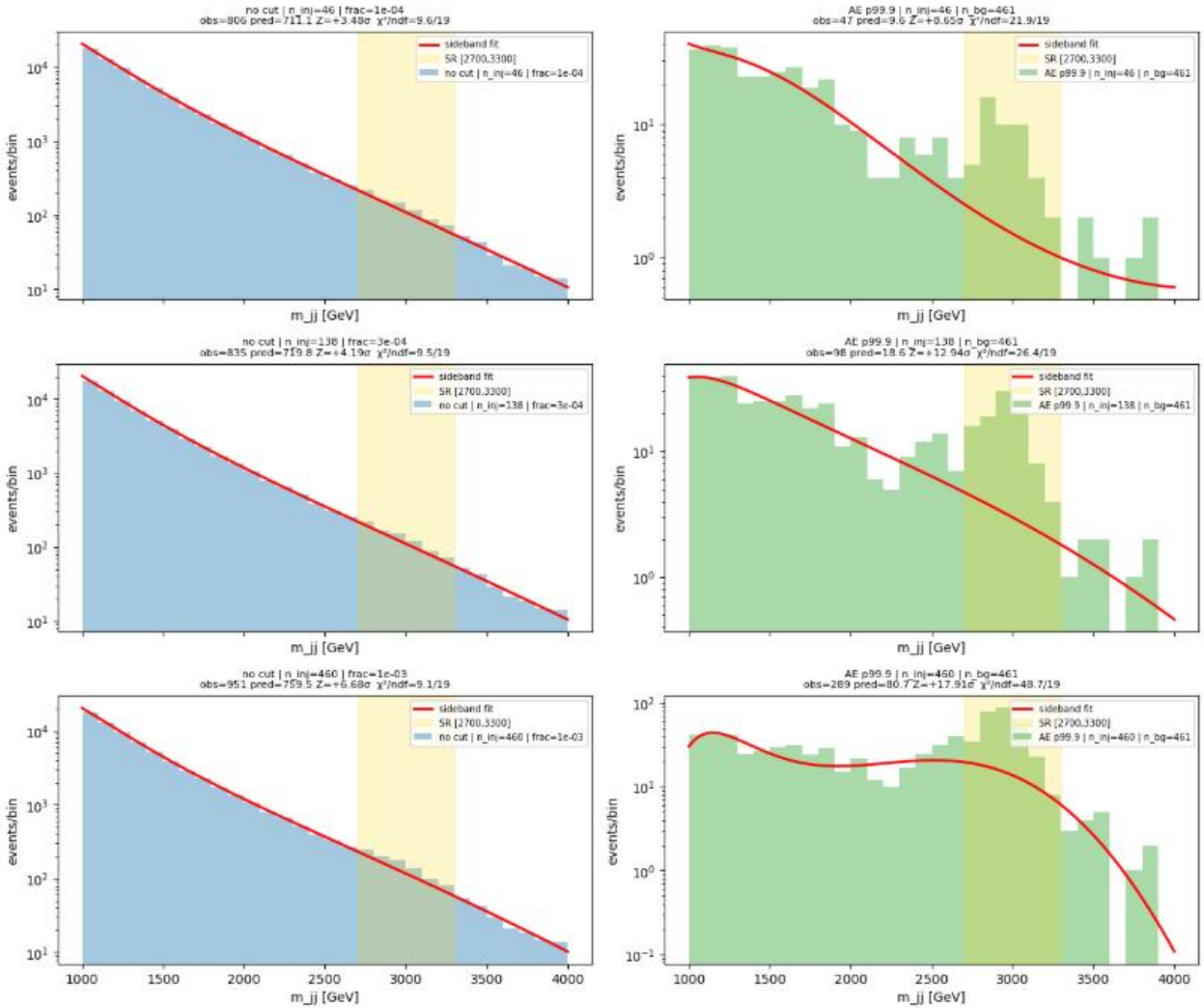
--- A14NNPDF23LO_Zprimebb3000_0 ---
     frac   n_inj   Z_nocut      p90.0      p95.0      p99.0      p99.5      p99.9
  1.0e-04      46      3.33       3.02       2.64       2.83       1.99       3.51
  3.0e-04     138      3.73       3.86       4.48       5.25       7.38       7.95
  1.0e-03     460      6.00       8.14       8.81      11.58       9.51       8.29
  3.0e-03    1381      9.68      13.07      16.45      13.75      17.24      17.06
```

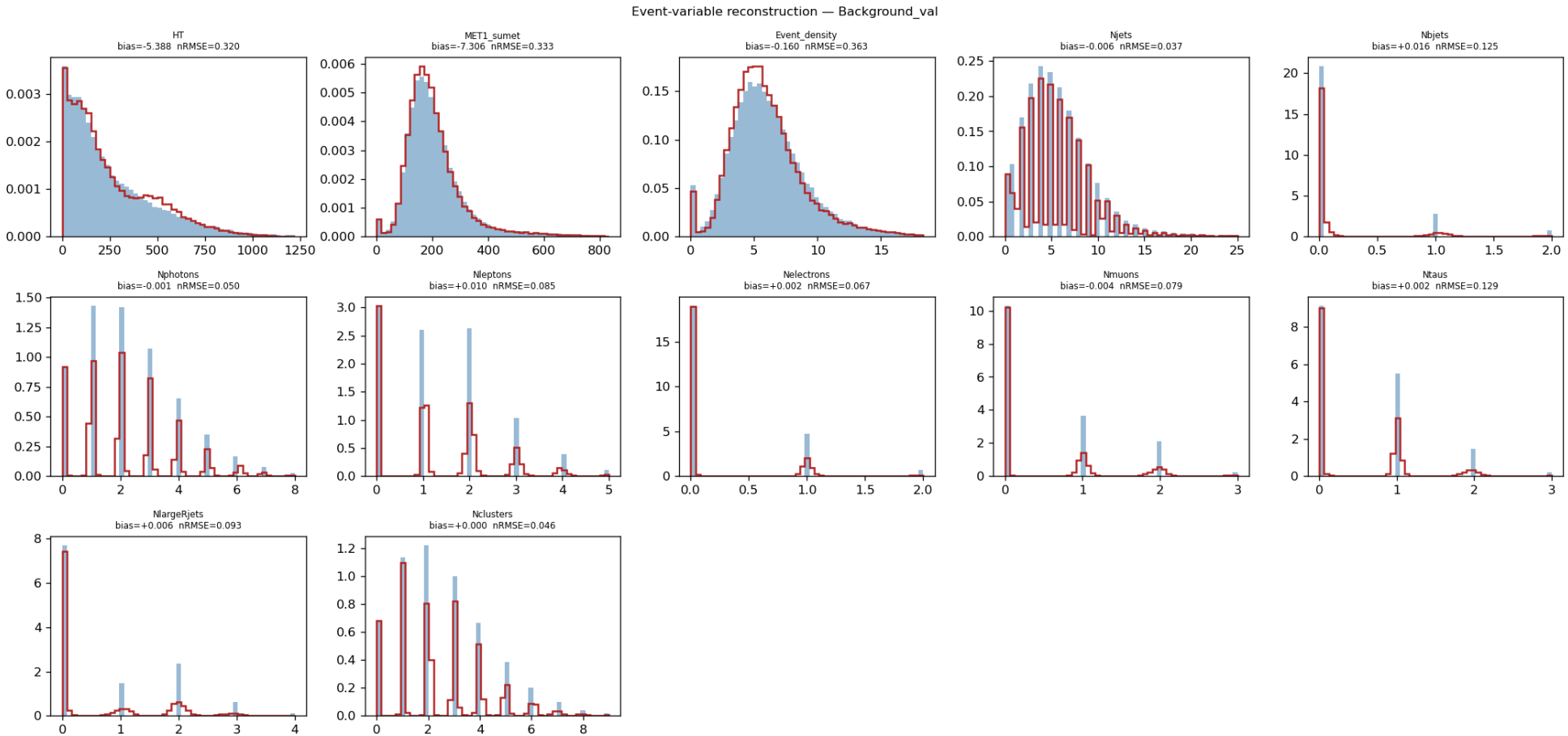
signal	frac	Z_nocut	Z_bestAEcut_net	best_AE_pct	improvement
A14NNPDF23LO_Wprime_qq_3000_0	0.000	3.263	8.213	99.900	2.517
A14NNPDF23LO_Wprime_qq_3000_0	0.000	4.027	9.728	99.900	2.416
A14NNPDF23LO_Wprime_qq_3000_0	0.001	6.607	16.904	99.900	2.559
A14NNPDF23LO_Wprime_qq_3000_0	0.003	12.511	17.532	99.900	1.401
A14NNPDF23LO_Zprimebb3000_0	0.000	3.327	3.509	99.900	1.055
A14NNPDF23LO_Zprimebb3000_0	0.000	3.727	7.947	99.900	2.132
A14NNPDF23LO_Zprimebb3000_0	0.001	6.004	8.285	99.900	1.380
A14NNPDF23LO_Zprimebb3000_0	0.003	9.684	17.062	99.900	1.762

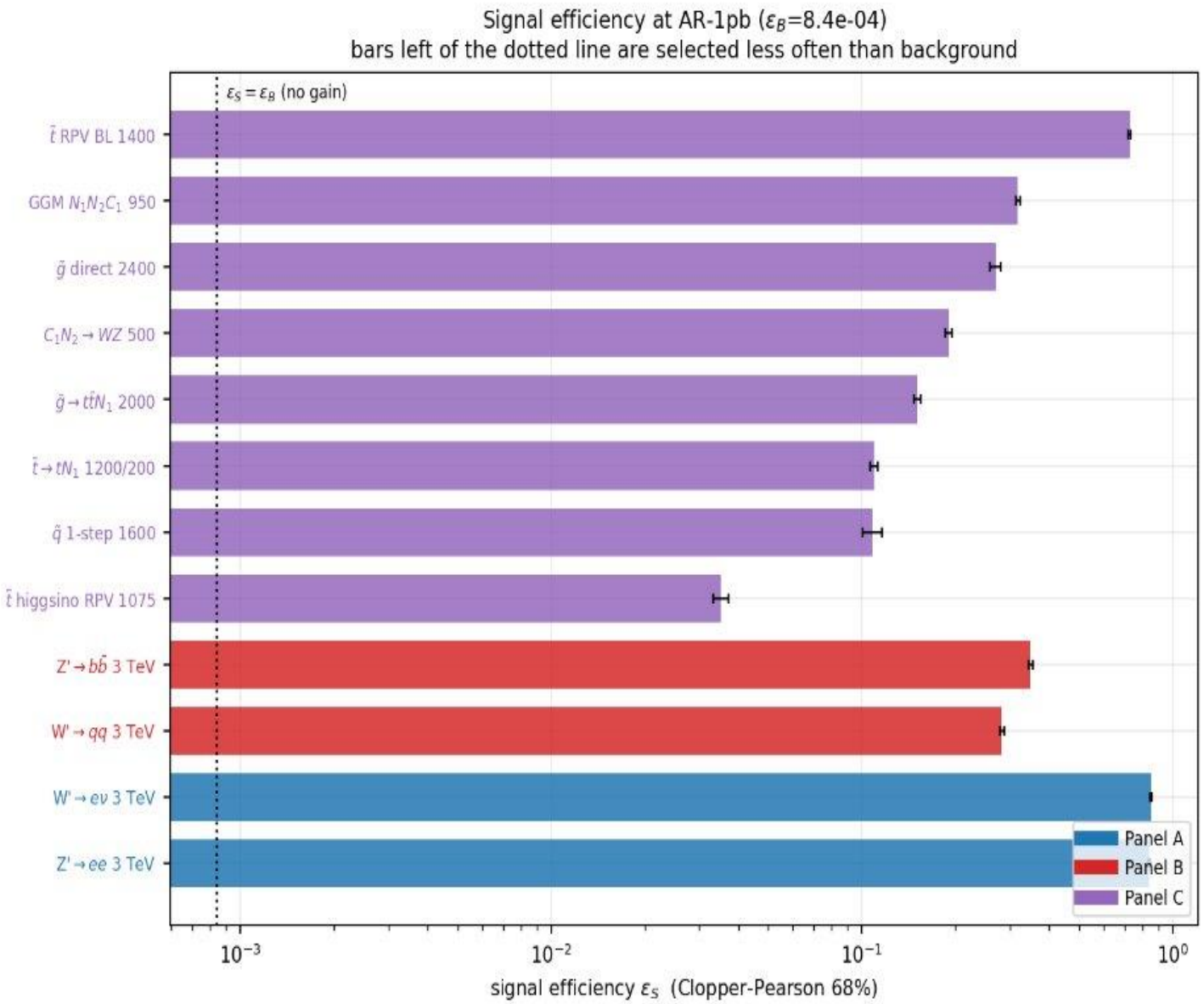
Zprime to bb



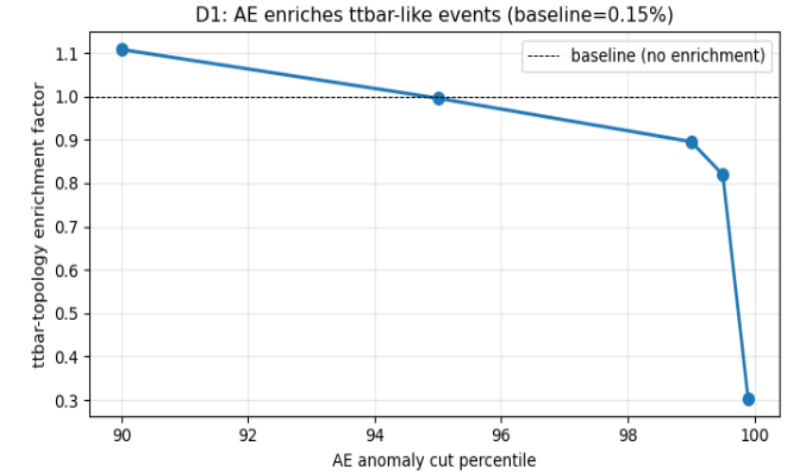
Wprime to qq

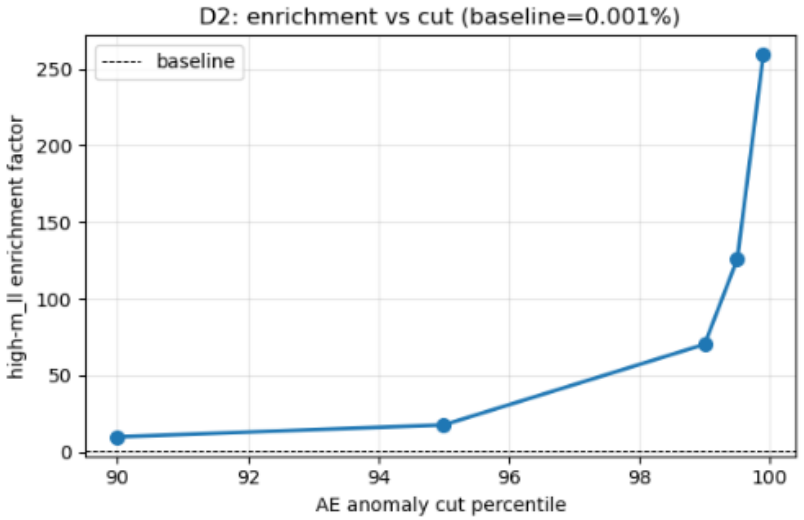



Event-variable reconstruction — Background_val
HT
bias=-5.388 nRMSE=0.320
MET1_sumet
bias=-7.306 nRMSE=0.333
Event_density
bias=-0.160 nRMSE=0.363
Njets
bias=-0.006 nRMSE=0.037
Nbjets
bias=+0.016 nRMSE=0.125
Nphotons
bias=-0.001 nRMSE=0.050
Nleptons
bias=+0.010 nRMSE=0.085
Nelectrons
bias=+0.002 nRMSE=0.067
Nmuons
bias=-0.004 nRMSE=0.079
Ntaus
bias=+0.002 nRMSE=0.129
NlargeRjets
bias=+0.006 nRMSE=0.093
Nclusters
bias=+0.000 nRMSE=0.046

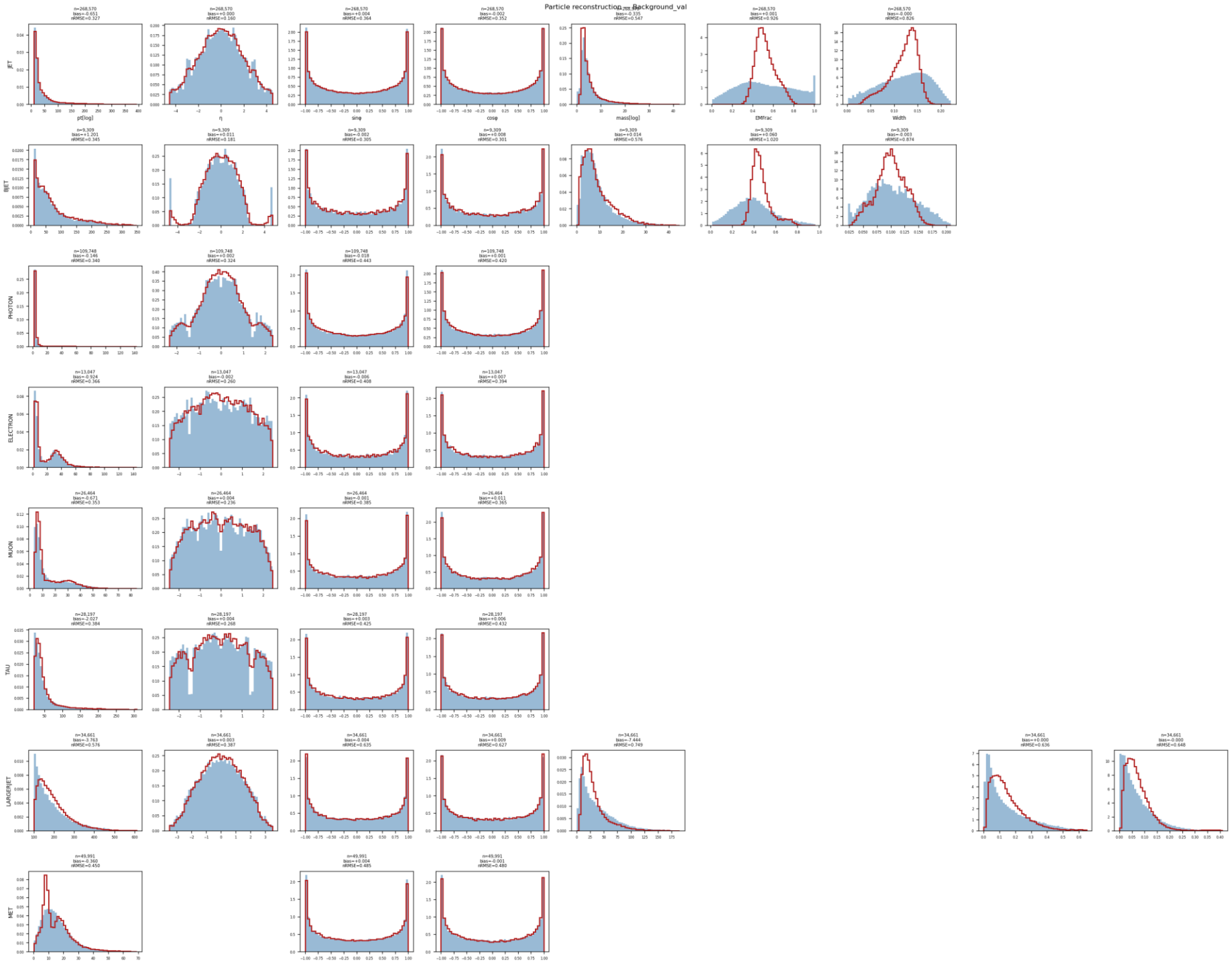


- **$t\bar{t}$, high mass DrellYan process (> 500)**, and **Higgs** candidates were reconstructed using specific event selections.
- Enrichment factors and local significances were computed.





Reconstruction plots



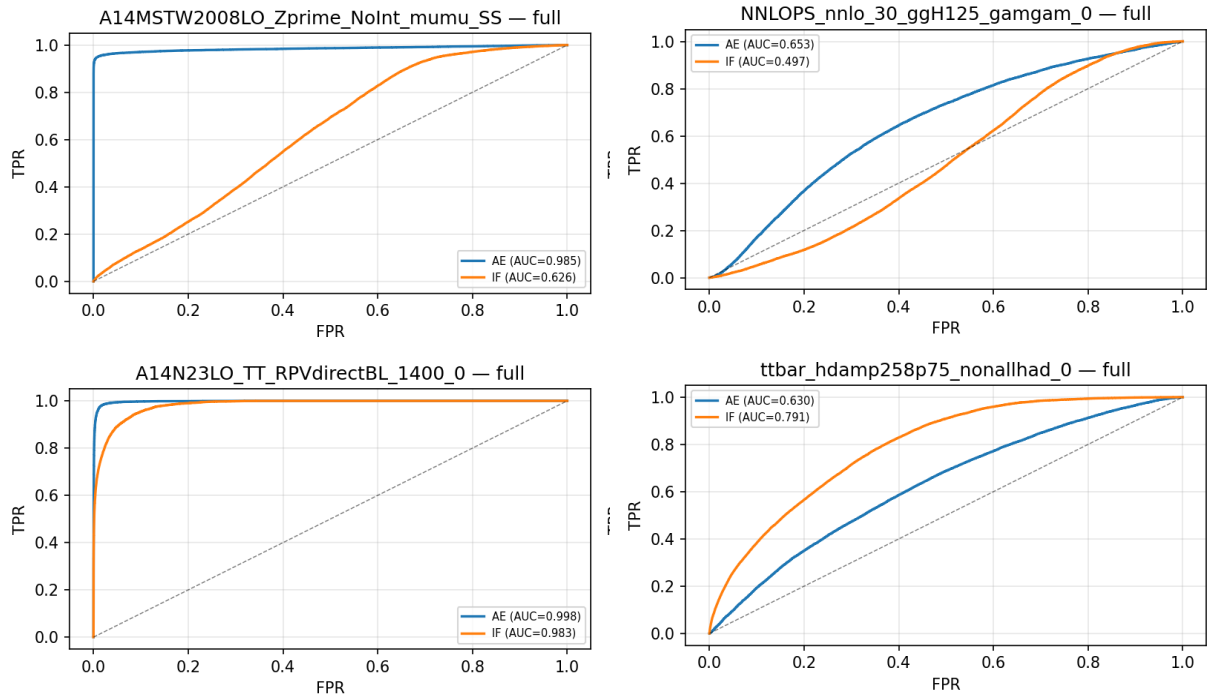
Isolation-forest model comparison

- Built with **500 isolation trees**, each using **4096 random samples**.
- Implemented using **GPU-accelerated cuML Isolation Forest**.

AUC comparison

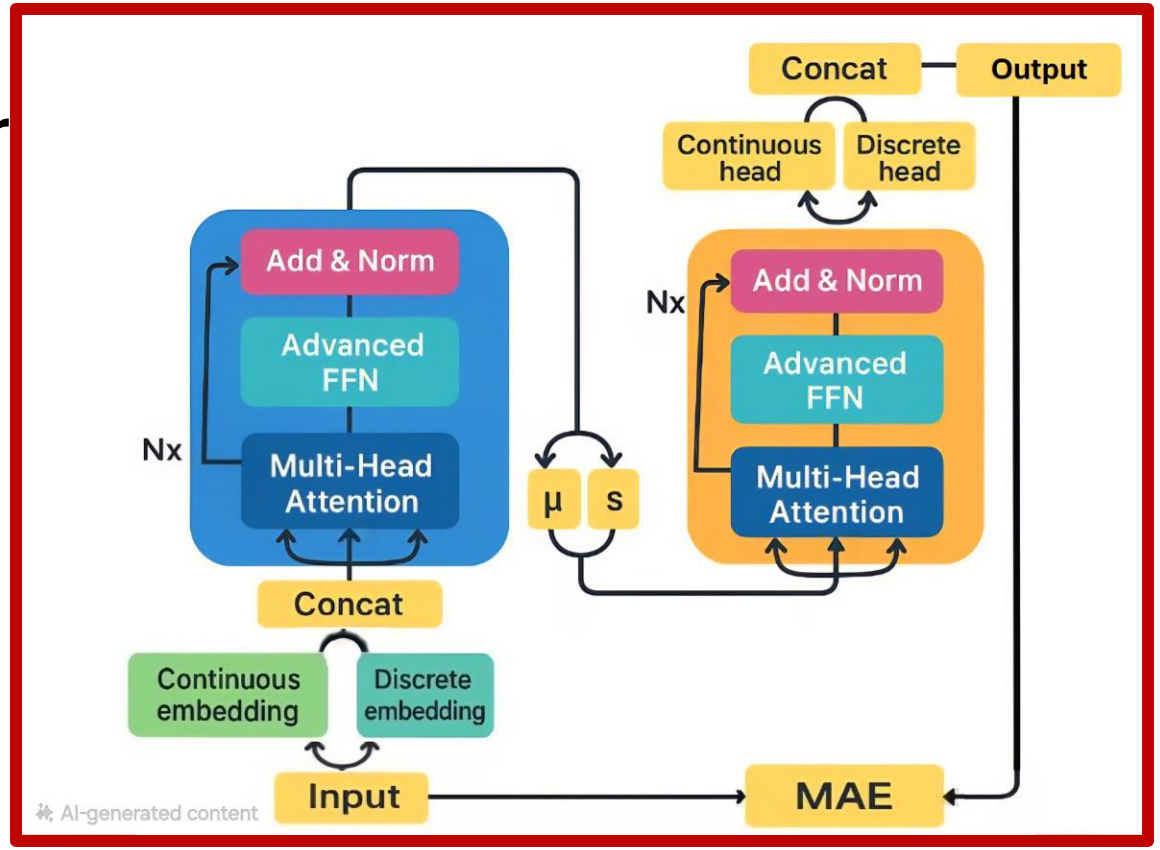
signal	AE_full	IF_full	AE_t1e3	IF_t1e3	AE_t1e4	IF_t1e4
A14MSTW2008LO_Zprime_NoInt_ee_SSM3000_0	0.974	0.861	0.903	0.511	0.797	0.502
A14MSTW2008LO_Zprime_NoInt_mumu_SSM3000_0	0.985	0.633	0.946	0.501	0.868	0.500
A14N23LO_C1N2N1_GGMHinoZh50_800_noFilter_0	0.907	0.758	0.540	0.505	0.508	0.500
A14N23LO_C1N2_WZ_500p0_100p0_3L_2L7_0	0.965	0.724	0.595	0.504	0.519	0.500
A14N23LO_SM_N2C1p_105_100_2L2MET75_MadSpin	0.683	0.443	0.502	0.500	0.500	0.500
A14N23LO_SS_onestepCC_1600_900_200_0	0.972	0.953	0.549	0.542	0.505	0.504
A14N23LO_TT_RPVdirectBL_1400_0	0.998	0.983	0.767	0.686	0.555	0.559
A14N23LO_TT_higgsinoRPV_1075_700_0	0.988	1.000	0.510	0.937	0.500	0.744
A14N23LO_TT_tN1_1200_200_MS_0	0.946	0.923	0.545	0.530	0.506	0.504
A14NNPDF23LO_GGM_N1N2C1_950_0	0.955	0.792	0.616	0.505	0.531	0.501
A14NNPDF23LO_Wprime_enu_SSM3000_0	0.973	0.658	0.915	0.500	0.847	0.500
A14NNPDF23LO_Wprime_munu_SSM3000_0	0.966	0.457	0.927	0.500	0.868	0.500
A14NNPDF23LO_Wprime_qq_3000_0	0.939	0.829	0.646	0.511	0.532	0.502
A14NNPDF23LO_Zprimebb3000_0	0.955	0.858	0.665	0.519	0.538	0.503
A14NNPDF23LO_zprime3000_tt_0	0.962	0.927	0.627	0.541	0.521	0.508
A14NNPDF23_NNPDF30ME_ttH125_gamgam_0	0.712	0.967	0.501	0.566	0.500	0.507
A14N_GG_ttn1_2000_5000_200_0	0.988	0.997	0.540	0.817	0.501	0.627
GG_direct_2400_200_0	0.995	0.956	0.640	0.539	0.519	0.504
LO_DMA_500_700_gq0p25_0	0.849	0.547	0.538	0.500	0.509	0.500
LO_DM_4topscalar_p1000_c1_nonallhad_0	0.927	0.997	0.502	0.767	0.500	0.579
NLO_LQd_gstML_0p3_nonallhad_M1000_0	0.997	0.988	0.645	0.726	0.514	0.574
NNLOPS_NN30_ggH125_WWlvlv_EF_15_5_0	0.607	0.433	0.500	0.500	0.500	0.500
NNLOPS_nnlo_30_ggH125_ZZ4l_0	0.813	0.678	0.500	0.501	0.500	0.500
NNLOPS_nnlo_30_ggH125_gamgam_0	0.653	0.473	0.500	0.500	0.500	0.500
NNPDF30ME_ttH125_dilep_0	0.781	0.930	0.501	0.548	0.500	0.506

ROC curves



Past work: TAB - Transformer autoencoder

- ❑ Trained a masked tabular transformer,
- ❑ 1 million run 2 data Is used to train the model.
- ❑ Masked random variables up to .3 fraction.
- ❑ Forced correlation learning.
- ❑ Variables grouped based on the shape and applied embedding .
- ❑ Developed a hybrid loss based on KNN based mahalanobis distance. Covering local and global deviation of events from SM manifold
- ❑ Tested 200M run2 events from atlas open data for research



Anomaly score

- **Hybrid anomaly score combines complementary global and local measures of event anomalousness.**
- **Empirical background calibration converts both components into a single robust significance score.**

https://pos.sissa.it/485/663

$$\tilde{\mathbf{r}} = \Sigma_{\mathrm{LW}}^{-1/2}(\mathbf{r} - \boldsymbol{\mu}),$$

$$D_{\mathrm{global}} = \|\tilde{\mathbf{r}}\|_2 ,$$

$$D_{\mathrm{kNN}} = \frac{1}{k}\sum_{i=1}^{k} \|\tilde{\mathbf{r}} - \tilde{\mathbf{r}}_{(i)}\|_2 ,$$

$$p_i = 1 - F_{\mathrm{bkg}}(D_i),$$

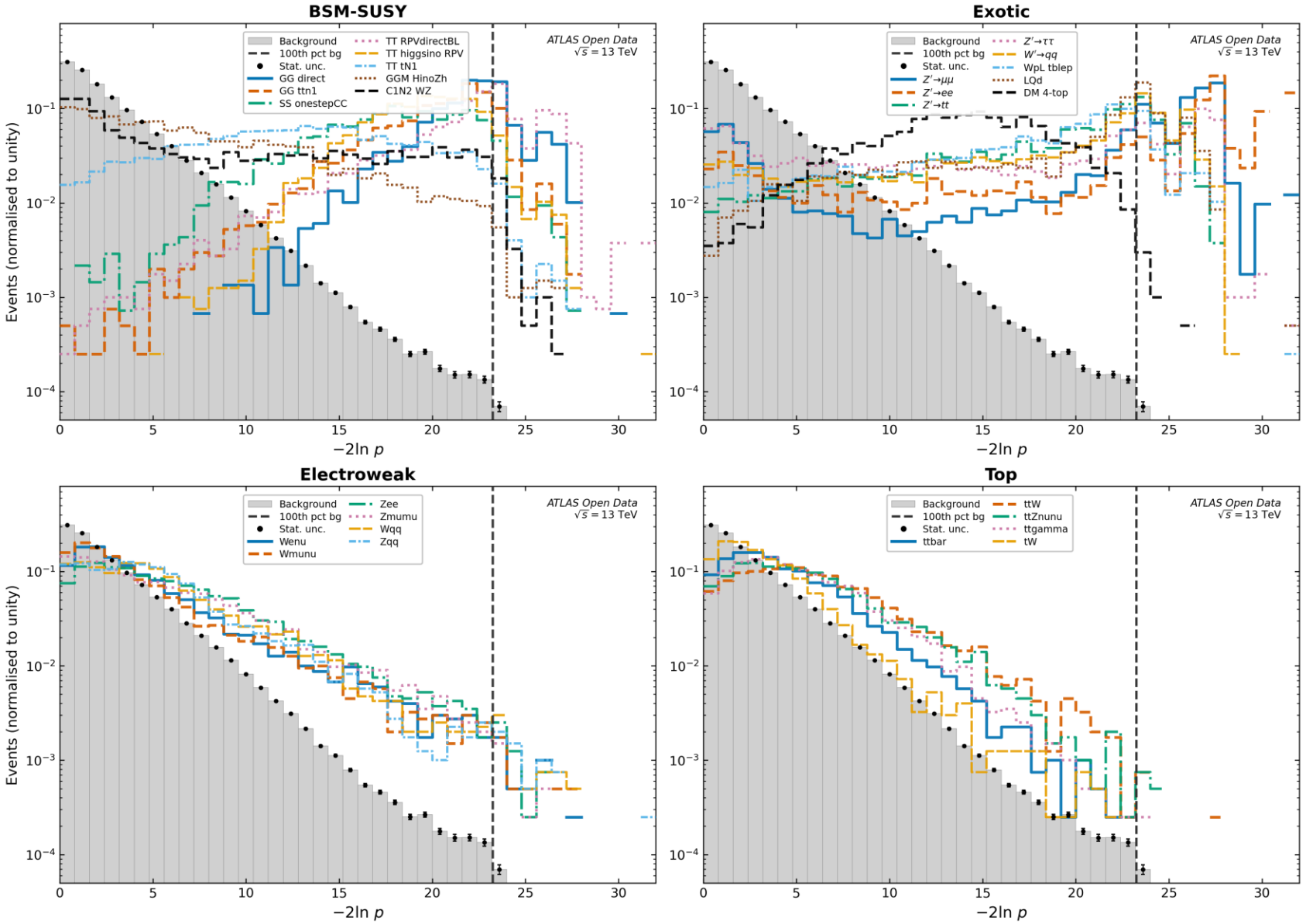
$$T_i = -2\ln(p_i),$$

$$S_{\mathrm{hybrid}} = \max(T_{\mathrm{global}}, T_{\mathrm{kNN}}) + \beta\,(T_{\mathrm{global}} + T_{\mathrm{kNN}}) .$$

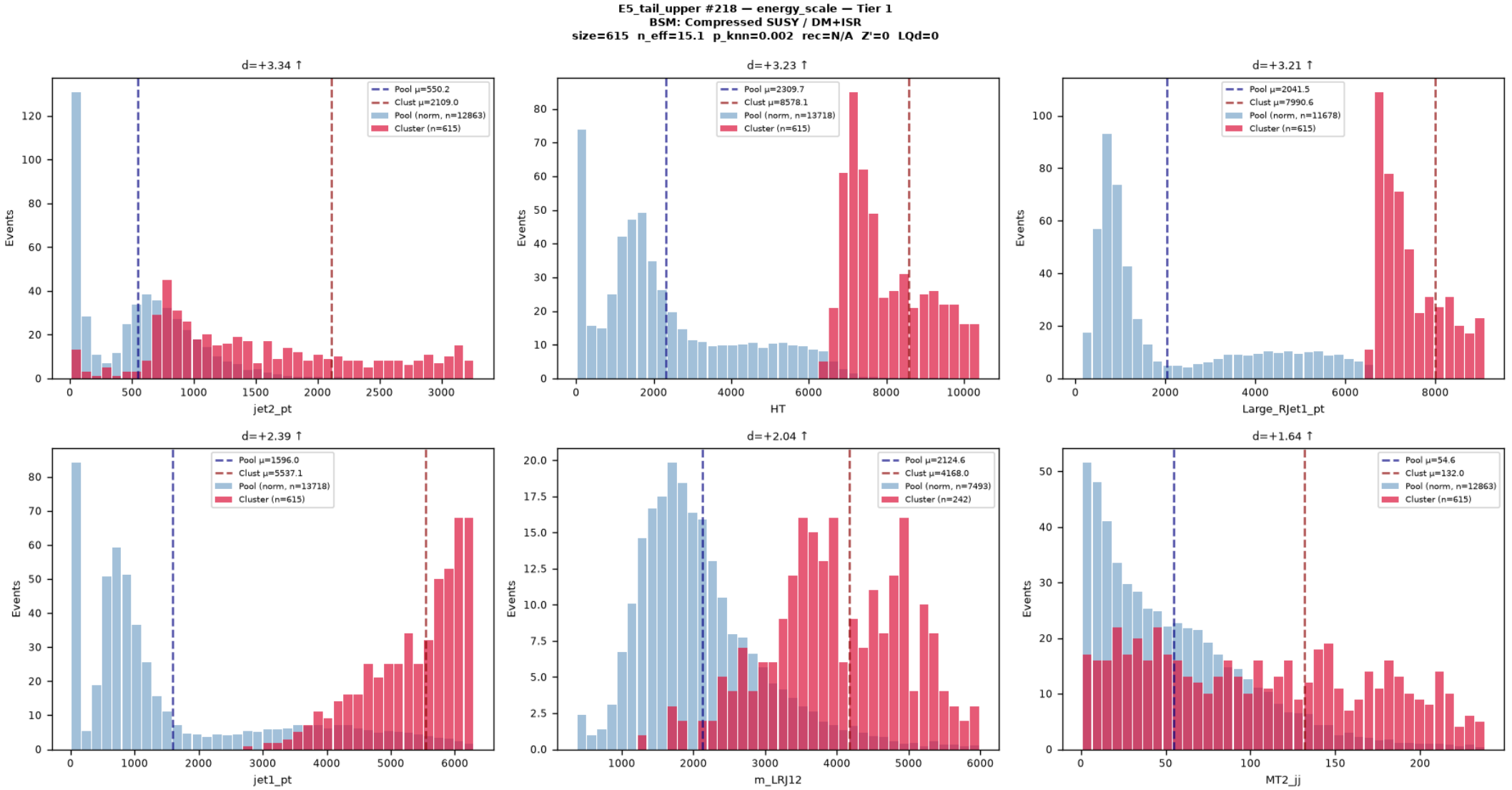
Past work: TAB - Transformer autoencoder

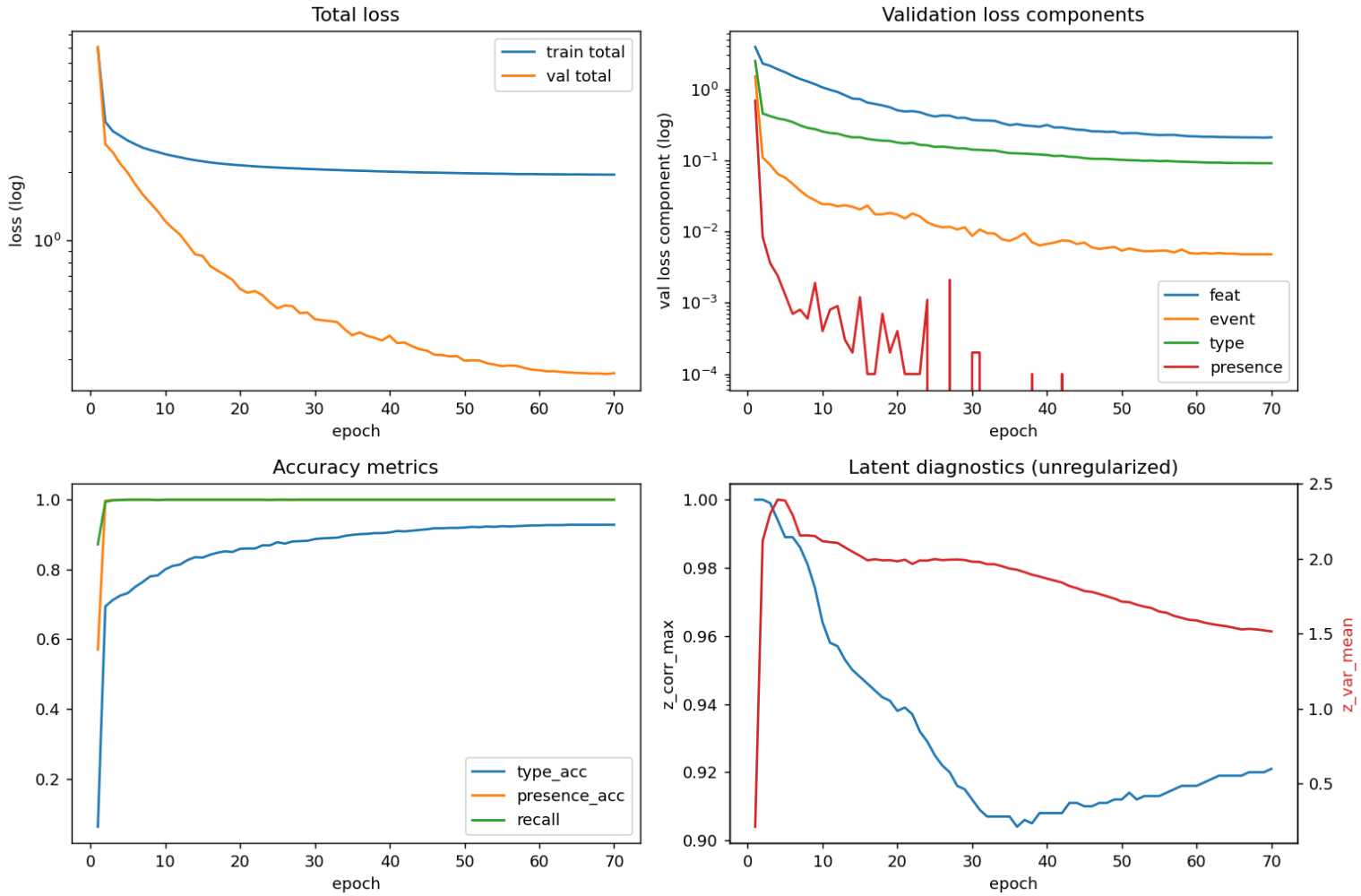
Cat.	Signal	AUC	Tail
BSM-Exotic	Zprime ee	0.944	0.751
	Zprime mumu	0.885	0.612
	Zprime tt	0.977	0.732
	Wprime qq	0.954	0.734
	WpL tblep	0.961	0.681
	LQd	0.969	0.482
	DM 4-top	0.975	0.339

Cat.	Signal	AUC	Tail
BSM-SUSY	TT tN1	0.953	0.376
	GG ttn1	0.998	0.871
	SS onestepCC	0.996	0.818
	GG direct	1.000	0.973
	TT higg. RPV	0.999	0.881
	TT RPVdirBL	0.997	0.893
	GGM N1N2C1	0.878	0.260
	C1N2 WZ	0.726	0.030



High HT values vs clusterization




Total loss
train total
val total
loss (log)
epoch
Validation loss components
val loss component (log)
feat
event
type
presence
epoch
Accuracy metrics
type_acc
presence_acc
recall
epoch
Latent diagnostics (unregularized)
z_corr_max
z_var_mean
epoch

Signal injection studies

- Operating threshold used for the production scan: p99.999 on sc_combined_fair corres ponding to a numeric cut of 22.241 on the fair score.
- Validated by injecting benchmark BSM signals (Z', W', leptoquarks, SUSY, dark-matter models, RPV) into the Run-2 background sample at varying injected yields(N = 10 – 1000) and measuring the fraction recovered above threshold.

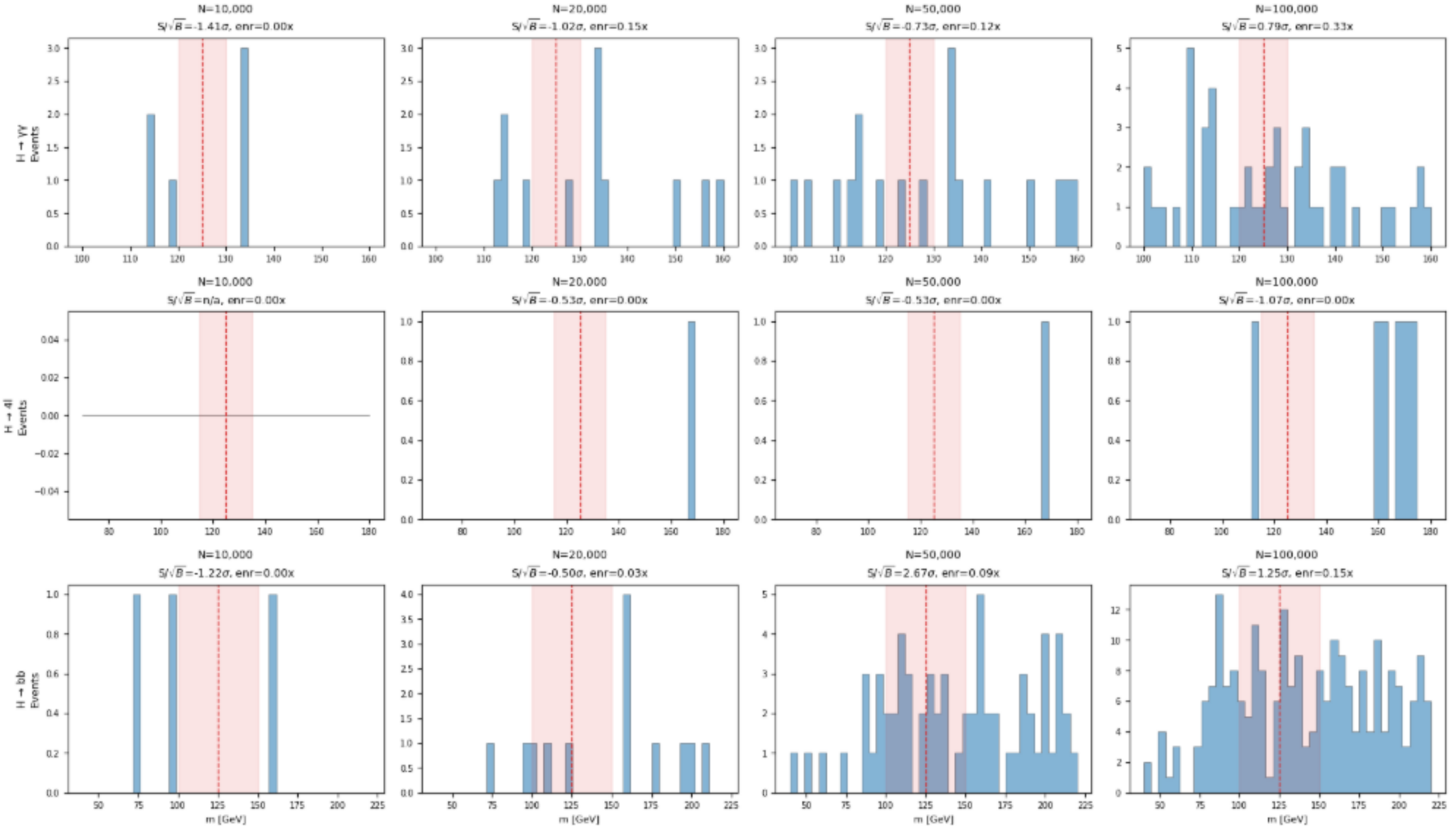
Signal	N=10	N=25	N=50	N=100	N=500	N=1
A14N23LO_TT_tN1_1200_200_MS_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	7/500 (1.4%)	12/1000
A14N_GG_ttn1_2000_5000_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000
4N23LO_SS_onestepCC_1600_900_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	3/500 (0.6%)	3/1000
GG_direct_2400_200_0	1/10 (10.0%)	2/25 (8.0%)	2/50 (4.0%)	6/100 (6.0%)	11/500 (2.2%)	20/1000
N23LO_C1N2_WZ_500p0_100p0_3L_2L7_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	8/500 (1.6%)	13/1000
LO_SM_N2C1p_105_100_2L2MET75_MadSpin_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000
23LO_C1N2N1_GGMHinoZh50_800_noFilter_0	0/10 (0.0%)	1/25 (4.0%)	1/50 (2.0%)	2/100 (2.0%)	4/500 (0.8%)	8/1000
14N23LO_TT_higgsinoRPV_1075_700_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000
A14N23LO_TT_RPVdirectBL_1400_0	1/10 (10.0%)	2/25 (8.0%)	5/50 (10.0%)	10/100 (10.0%)	41/500 (8.2%)	67/1000
StauStauDirect_200p0_1p0_TFilt_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	2/500 (0.4%)	3/1000
A14NNPDF23LO_GGM_N1N2C1_950_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	4/100 (4.0%)	23/500 (4.6%)	42/1000
Zprime_NoInt_tautau_SSM3000_0	3/10 (30.0%)	9/25 (36.0%)	17/50 (34.0%)	35/100 (35.0%)	188/500 (37.6%)	399/1000
Wprime_taunu_SSM3000_0	3/10 (30.0%)	5/25 (20.0%)	13/50 (26.0%)	31/100 (31.0%)	148/500 (29.6%)	299/1000
MSTW2008LO_Zprime_NoInt_ee_SSM3000_0	4/10 (40.0%)	11/25 (44.0%)	31/50 (62.0%)	59/100 (59.0%)	309/500 (61.8%)	620/1000
STW2008LO_Zprime_NoInt_mumu_SSM3000_0	10/10 (100.0%)	21/25 (84.0%)	36/50 (72.0%)	74/100 (74.0%)	387/500 (77.4%)	771/1000
14NNPDF23LO_Wprime_enu_SSM3000_0	7/10 (70.0%)	20/25 (80.0%)	34/50 (68.0%)	69/100 (69.0%)	366/500 (73.2%)	722/1000
4NNPDF23LO_Wprime_munu_SSM3000_0	6/10 (60.0%)	15/25 (60.0%)	33/50 (66.0%)	73/100 (73.0%)	371/500 (74.2%)	750/1000
A14NNPDF23LO_zprime3000_tt_0	1/10 (10.0%)	1/25 (4.0%)	2/50 (4.0%)	2/100 (2.0%)	12/500 (2.4%)	21/1000
A14NNPDF23LO_Wprime_qq_3000_0	1/10 (10.0%)	3/25 (12.0%)	5/50 (10.0%)	7/100 (7.0%)	28/500 (5.6%)	55/1000
A14NNPDF23LO_Zprimebb3000_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	6/100 (6.0%)	30/500 (6.0%)	67/1000
WpL_tblep_M3000_0	0/10 (0.0%)	1/25 (4.0%)	3/50 (6.0%)	4/100 (4.0%)	21/500 (4.2%)	43/1000
WpL_tbhad_M3000_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	3/100 (3.0%)	12/500 (2.4%)	29/1000
LO_DMA_500_700_gq0p25_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	7/500 (1.4%)	10/1000
LO_LQd_gstML_0p3_nonallhad_M1000_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	1/100 (1.0%)	10/500 (2.0%)	15/1000
O_DM_4topscalar_p1000_c1_nonallhad_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000

Signal injection studies

- 2/3 of the run2 2016 data.
- Operating threshold used for the production scan: p99.999 on sc_combined_fair corres ponding to a numeric cut of 22.241 on the fair score.
- Validated by injecting benchmark BSM signals (Z', W', leptoquarks, SUSY, dark-matter models, RPV) into the Run-2 background sample at varying injected yields(N = 10 – 1000) and measuring the fraction recovered above threshold.

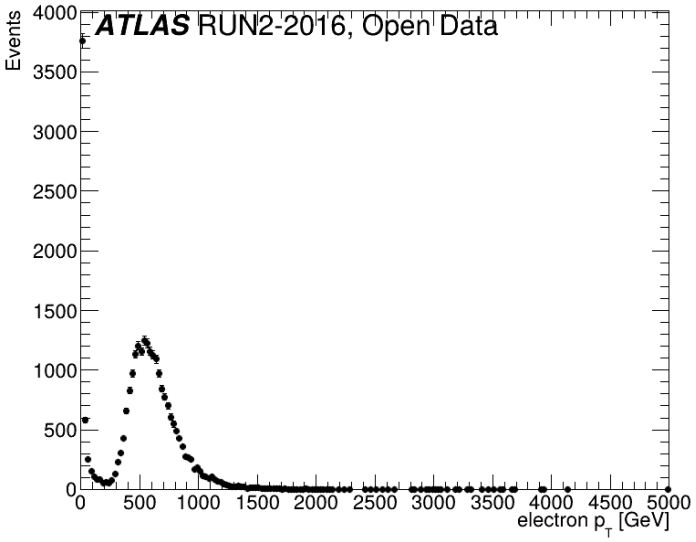
Signal	N=10	N=25	N=50	N=100	N=500	N=10
A14N23LO_TT_tN1_1200_200_MS_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	5/500 (1.0%)	9/1000 (
A14N_GG_ttn1_2000_5000_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (
14N23LO_SS_onestepCC_1600_900_200_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	3/500 (0.6%)	3/1000 (
GG_direct_2400_200_0	0/10 (0.0%)	1/25 (4.0%)	2/50 (4.0%)	4/100 (4.0%)	8/500 (1.6%)	10/1000 (
4N23LO_C1N2_WZ_500p0_100p0_3L_2L7_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	7/500 (1.4%)	13/1000 (
LO_SM_N2C1p_105_100_2L2MET75_MadSpin_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (
I23LO_C1N2N1_GGMHinoZh50_800_noFilter_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	2/500 (0.4%)	6/1000 (
A14N23LO_TT_higgsinoRPV_1075_700_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (
A14N23LO_TT_RPVdirectBL_1400_0	1/10 (10.0%)	2/25 (8.0%)	5/50 (10.0%)	7/100 (7.0%)	28/500 (5.6%)	50/1000 (
StauStauDirect_200p0_1p0_TFilt_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	2/500 (0.4%)	3/1000 (
A14NNPDF23LO_GGM_N1N2C1_950_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	17/500 (3.4%)	29/1000 (
Zprime_NoInt_tautau_SSM3000_0	3/10 (30.0%)	8/25 (32.0%)	16/50 (32.0%)	29/100 (29.0%)	158/500 (31.6%)	337/1000 (
Wprime_taunu_SSM3000_0	2/10 (20.0%)	6/25 (24.0%)	10/50 (20.0%)	25/100 (25.0%)	121/500 (24.2%)	244/1000 (
MSTW2008LO_Zprime_NoInt_ee_SSM3000_0	3/10 (30.0%)	9/25 (36.0%)	25/50 (50.0%)	49/100 (49.0%)	253/500 (50.6%)	511/1000 (
ISTW2008LO_Zprime_NoInt_mumu_SSM3000_0	8/10 (80.0%)	18/25 (72.0%)	31/50 (62.0%)	66/100 (66.0%)	351/500 (70.2%)	713/1000 (
A14NNPDF23LO_Wprime_enu_SSM3000_0	7/10 (70.0%)	19/25 (76.0%)	32/50 (64.0%)	63/100 (63.0%)	339/500 (67.8%)	677/1000 (
14NNPDF23LO_Wprime_munu_SSM3000_0	6/10 (60.0%)	15/25 (60.0%)	33/50 (66.0%)	76/100 (76.0%)	370/500 (74.0%)	739/1000 (
A14NNPDF23LO_zprime3000_tt_0	1/10 (10.0%)	1/25 (4.0%)	1/50 (2.0%)	1/100 (1.0%)	6/500 (1.2%)	10/1000 (
A14NNPDF23LO_Wprime_qq_3000_0	0/10 (0.0%)	0/25 (0.0%)	2/50 (4.0%)	3/100 (3.0%)	16/500 (3.2%)	33/1000 (
A14NNPDF23LO_Zprimebb3000_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	3/100 (3.0%)	15/500 (3.0%)	38/1000 (
WpL_tblep_M3000_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	13/500 (2.6%)	26/1000 (
WpL_tbhad_M3000_0	0/10 (0.0%)	0/25 (0.0%)	1/50 (2.0%)	3/100 (3.0%)	6/500 (1.2%)	17/1000 (
LO_DMA_500_700_gq0p25_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	4/500 (0.8%)	8/1000 (
NLO_LQd_gstML_0p3_nonallhad_M1000_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	1/100 (1.0%)	5/500 (1.0%)	6/1000 (
LO_DM_4topscalar_p1000_c1_nonallhad_0	0/10 (0.0%)	0/25 (0.0%)	0/50 (0.0%)	0/100 (0.0%)	0/500 (0.0%)	0/1000 (

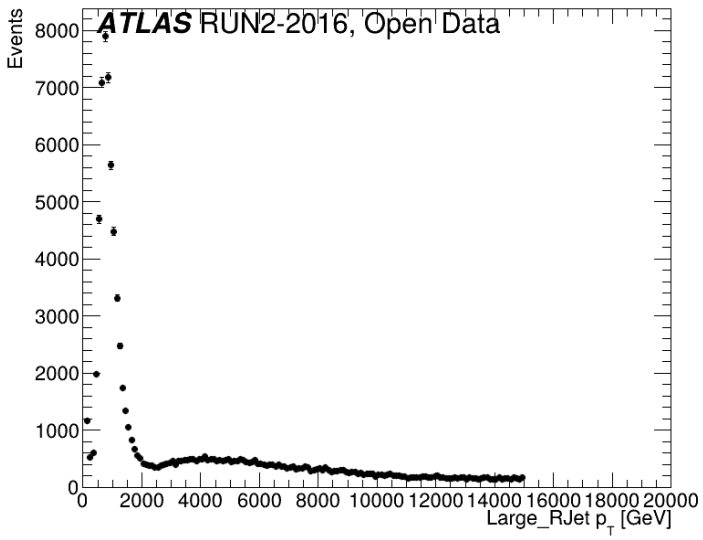
Higgs analysis

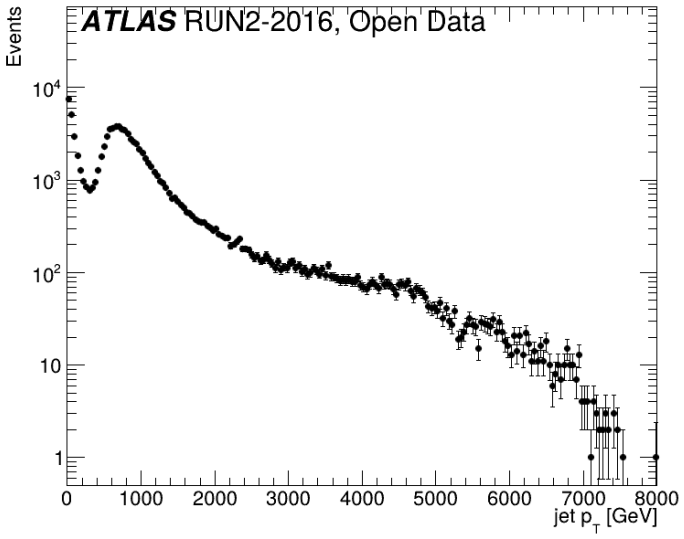


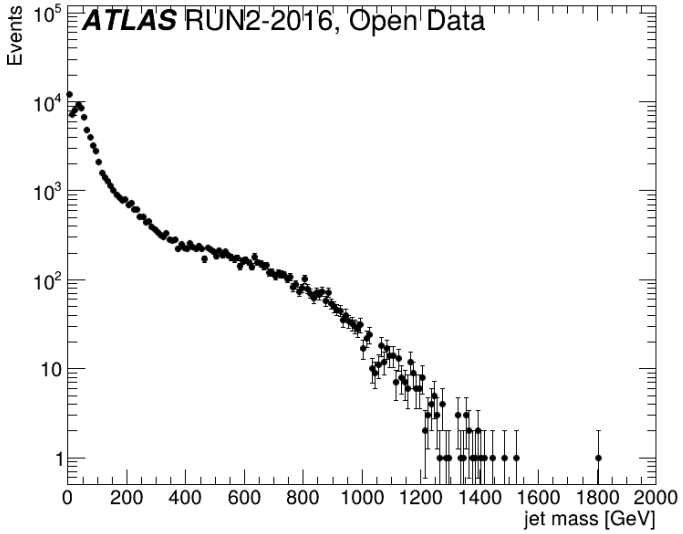
Discussion

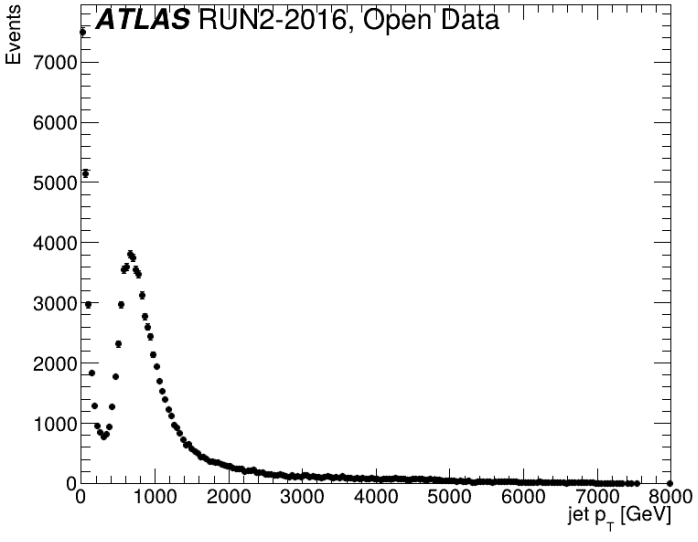
- Even though the model is able to give back injected cluster, it is missing a lot of signals,
- The selected space is having issues.
- Model is sensitive to outliers and selecting data issues and artifacts.
- As I didn't applied any quality cuts, only remove pt > unphysical values.





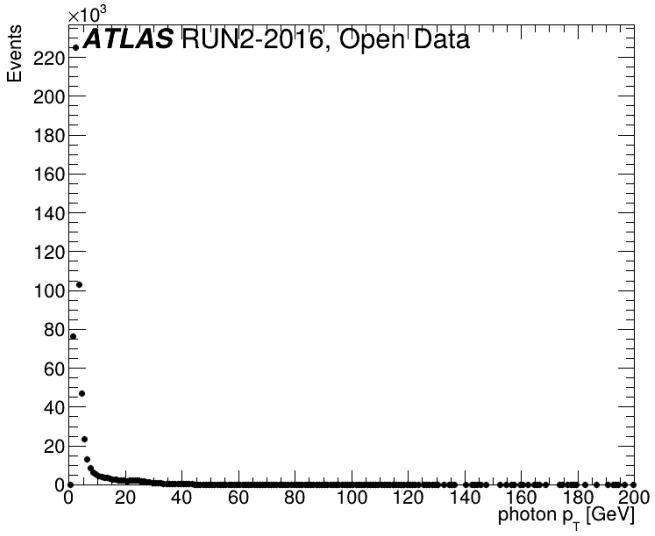


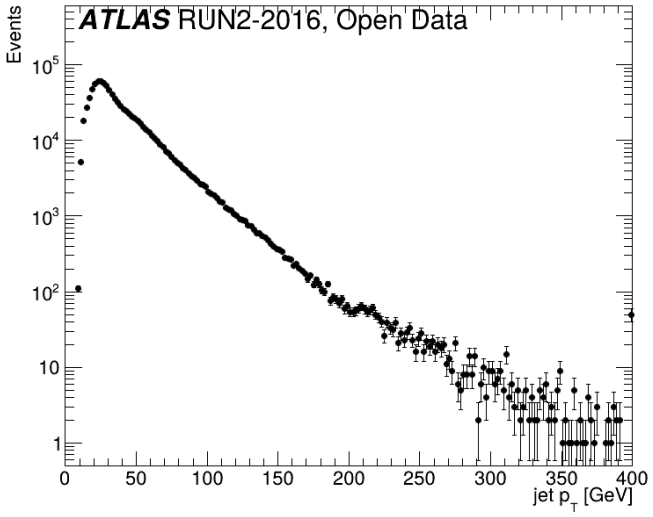


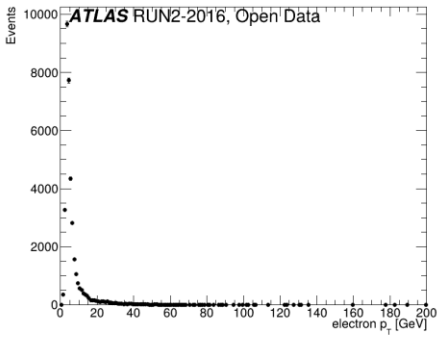


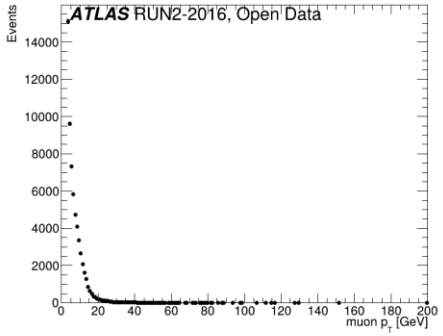
Discussion

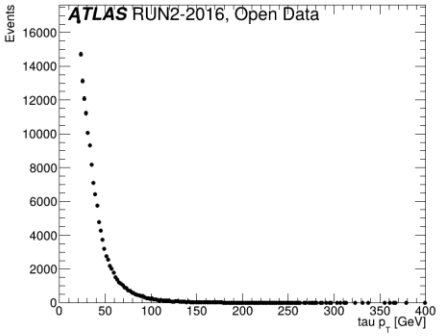
- Around 6k of candidates over 6.5 TeV.
- Because of these model selection artifacts, bump hunting, endpoint scan, clusterization all in noise.
- Random SM events don't have any issues for null computing.



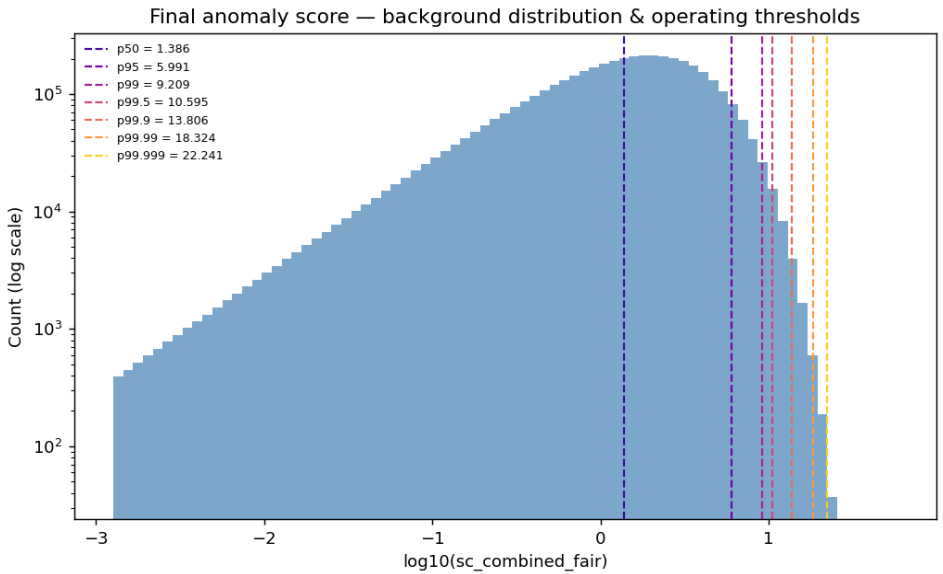






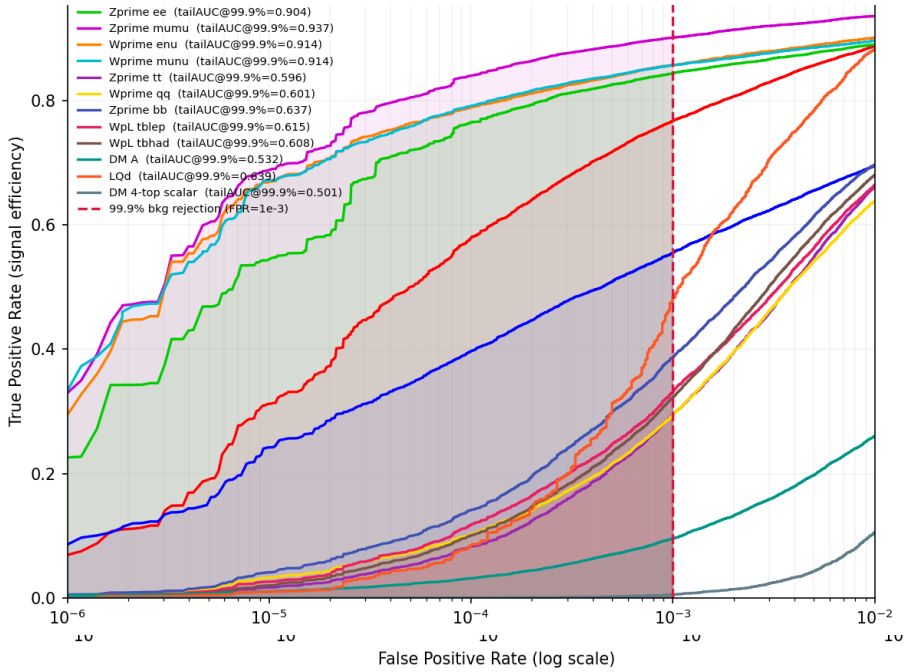


AE signal injection studies



- **Production threshold:** p99.999 on the fair anomaly score (cut = **22.77**),
- **Performance:** High-mass resonances are efficiently retained (**Z'→μμ: 56–80%**, **W'→eν/μν: 50–68%**), while compressed/SM-like signals have intentionally low acceptance.
- **Implication:** The AE acts as a **kinematic outlier filter**, motivating the downstream multi-engine discovery pipeline to recover softer and topology-specific BSM signatures.

Tail ROC curves



PHYSICS CHARECTERIZTION (supervised)

- **Real issue now is to flag the artifact candidates/ remove the quality issue candidates from the final anomalies so that genuine interesting clusters can be observed.**

- **Accept statistically significant candidates** from Stage 5 after multiple-testing correction.
- **Identify the most discriminating physics variables** using Cohen's d to understand what drives each anomaly.
- **Characterize the candidate topology** using compactness, purity, effective dimensionality and cluster properties.
- **Assign possible BSM signatures** by comparing the observed variable patterns and topology with known physics scenarios (Z', SUSY, LLP, etc.).
- **Generate a complete candidate summary**, including significance, dominant variables, event composition and engine of origin.
- **Prioritize candidates into Tier-1, Tier-2 and Tier-3**, providing a final shortlist for detailed physics analysis.

Statistically validated anomaly candidates.

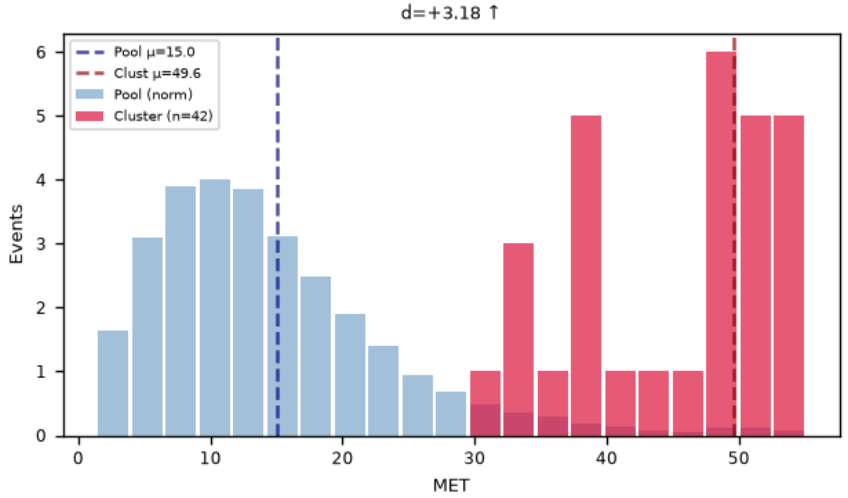
Physics-driven signal region definition.

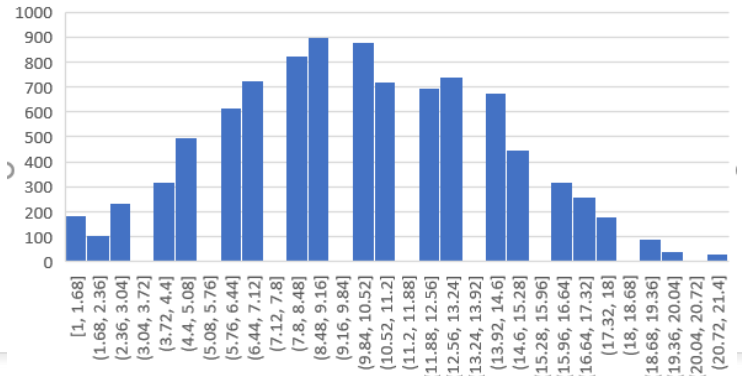
Standard HEP background estimation (CR/Sidebands/MC).

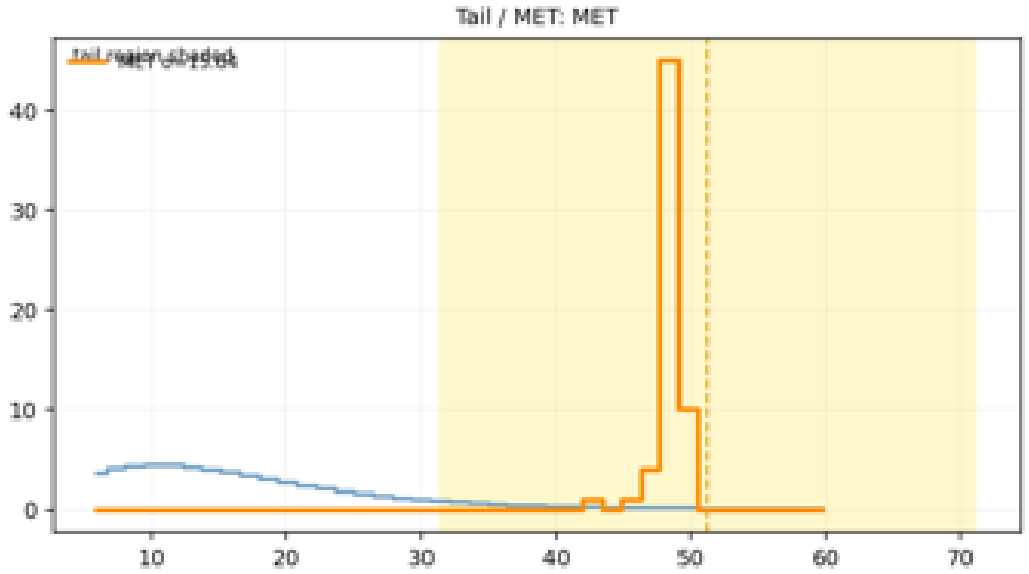
Final discovery via profile-likelihood inference including systematic uncertainties.

MET cluster

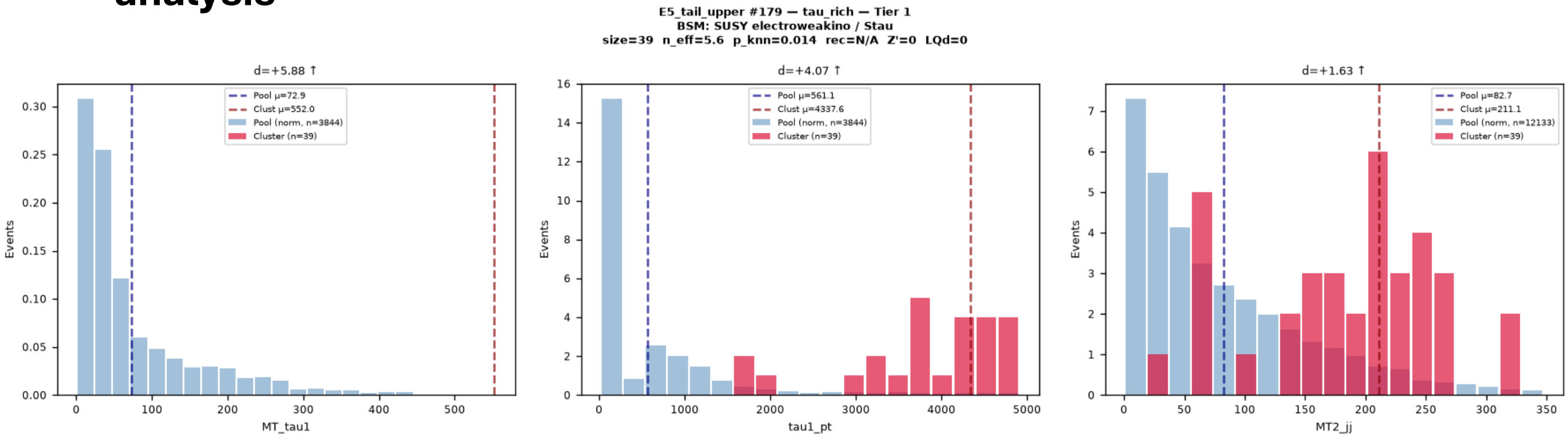
- Found that all events in this cluster are MET only events, the anomalies maintaining structure of the real data partially.
- Contacted atlas open data group.



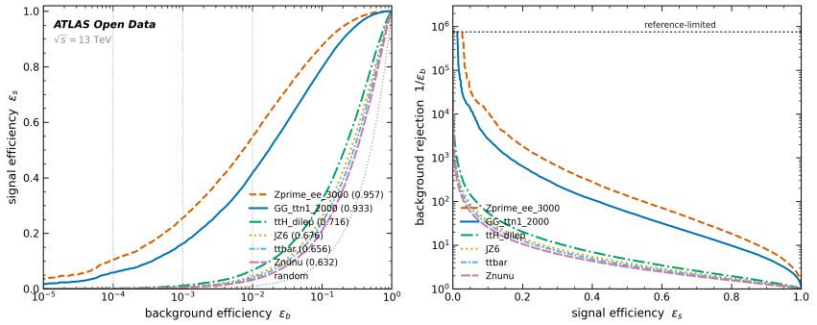


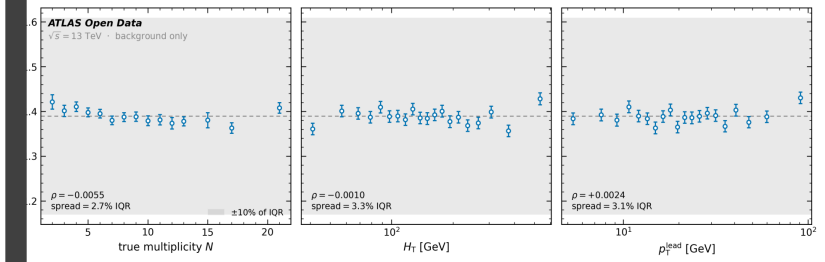


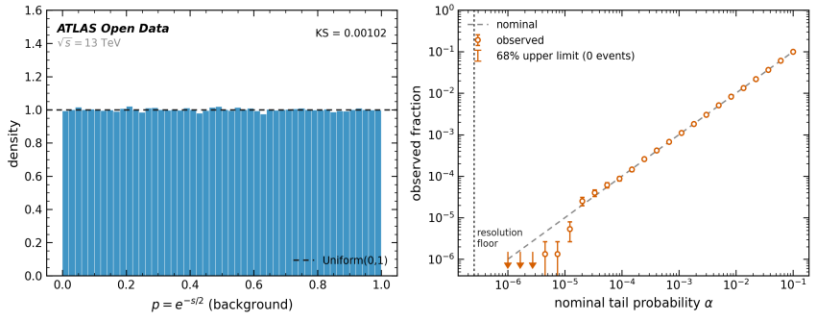
Example clusters survived in the statistical filtering and topology analysis

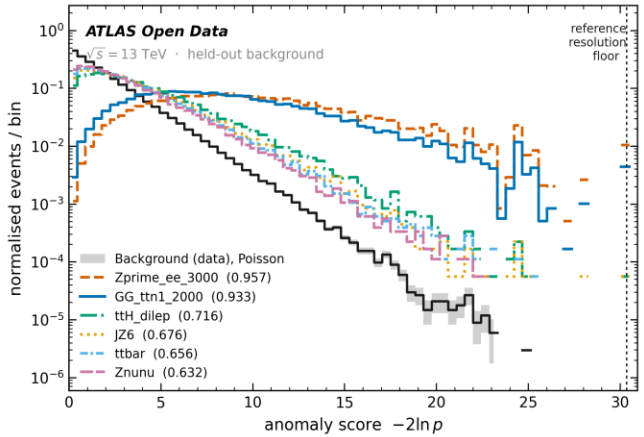


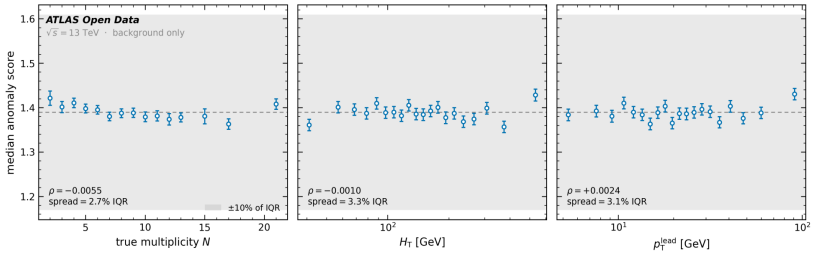
- How to go for a real analysis? Background estimation?
- As we have interesting clusters, first need to confirm the selected cluster have genuine interesting events or not.
- Analyze each event by going through the event id, if it is truly interesting, using the cluster common kinematics, we can define the signal region, and go for background estimation which passes the kinematic and selection cuts.
- Currently working on it

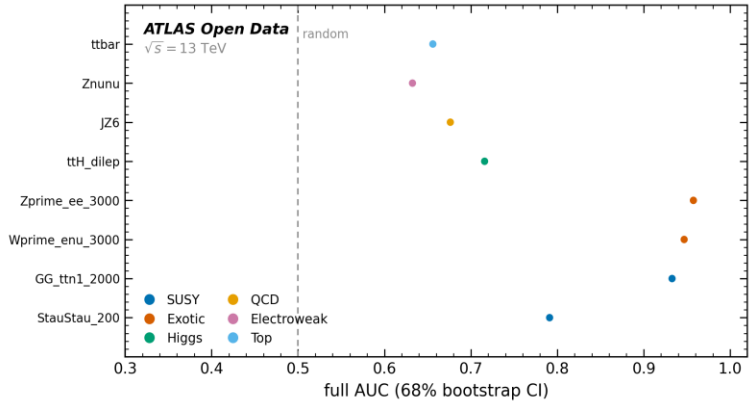












AE test calibration