"Flux + Mutability": A Conditional Generative Approach to One-Class Classification and Anomaly Detection

James Giroux (uOttawa, URegina)

Dr. Cristiano Fanelli (MIT, NSF AI Institute)

Dr. Zisis Papandreou (URegina)

June 06, 2022

University of Regina

Outline

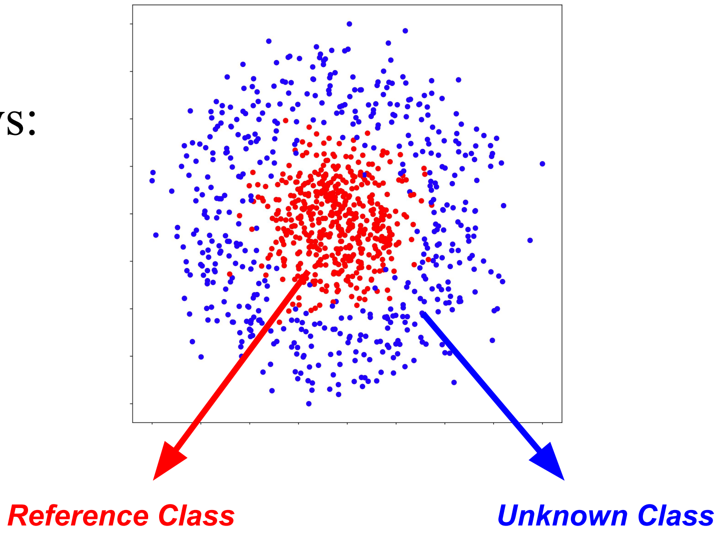
- One-Class Classification (OCC) and Anomaly Detection (AD)
- "Flux + Mutability" (F+M) - A Conditional Generative Approach
- γ/n separation at GlueX - OCC
- Standard Model(SM)/Beyond (BSM) Dijet Separation at LHC - AD
- Summary

C. Fanelli, J. Giroux, Z. Papandreou, "Flux+Mutability": A Conditional Generative Approach to One-Class Classification and Anomaly Detection (2022). https://arxiv.org/abs/2204.08609

University of Regina

OCC and AD

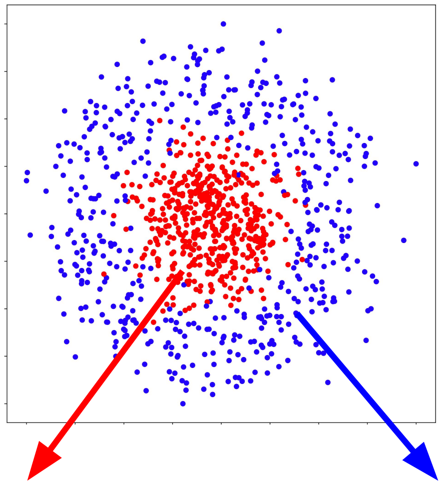
Suppose we have two classes distributed as follows:



OCC and AD

Suppose we have two classes distributed as follows:

1. Can we use deep learning to separate the two more efficiently than standard rectangular cuts?

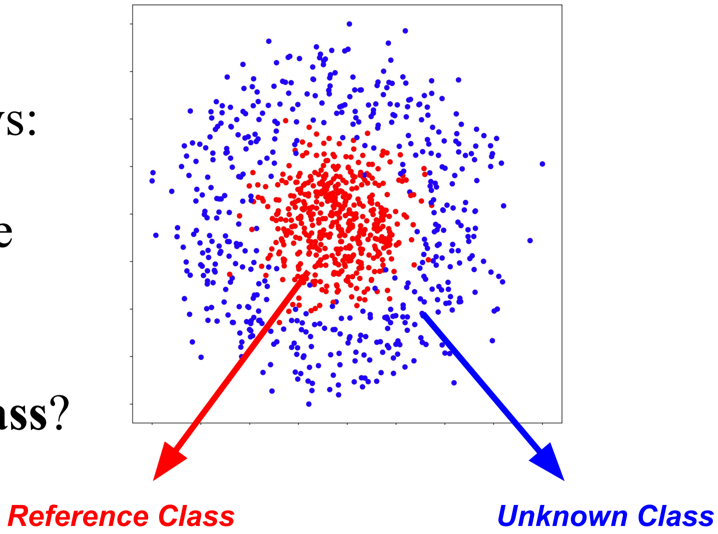


OCC and AD

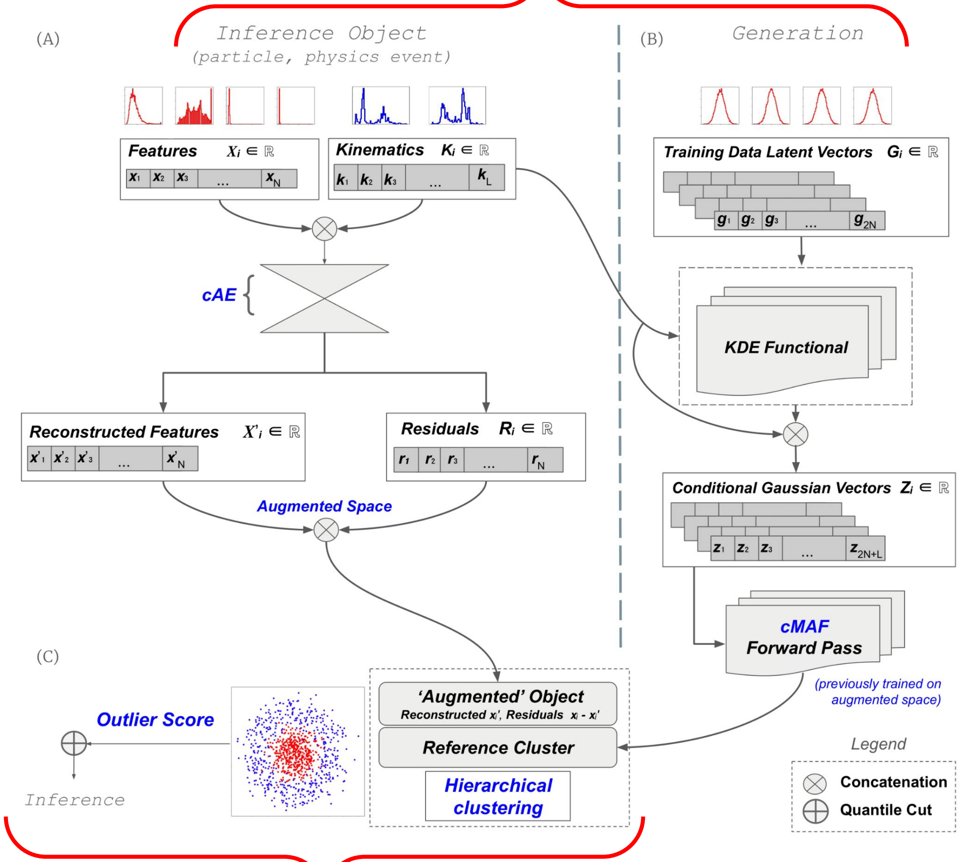
Suppose we have two classes distributed as follows:

1. Can we use deep learning to separate the two more efficiently than standard rectangular cuts?

2. Can we remain agnostic towards the **unknown class**?
   - Agnostic threshold selection



“Flux



\+ Mutability”

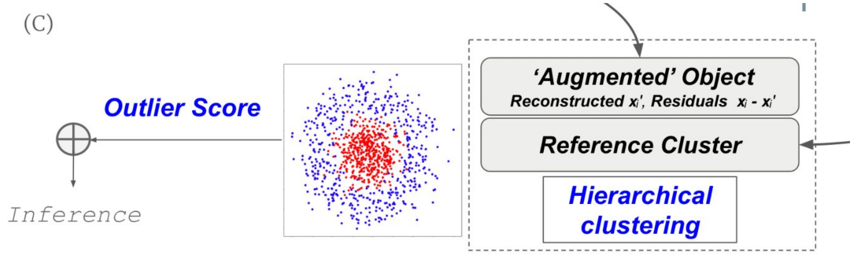
(A) Inference Object fed through cAE
- Features $\otimes$ Kinematics
- Reconstructions $\otimes$ Residuals (x’ - x)

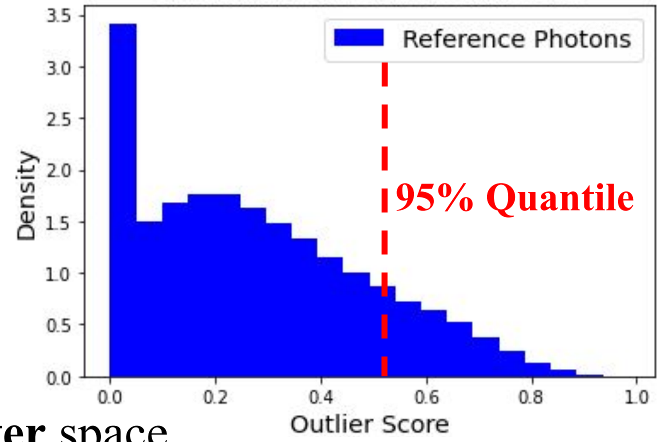
(B) Continuous Conditional Generation
- Pre-fit KDE Objects in kinematic bins
- Map inference kinematics to KDE object
- Sample new Gaussian vectors from restricted domain
- Gaussian Vectors $\otimes$ Inference Kinematics
- **Conditionally generate reference population** via cMAF

(C) Compare inference object to **reference population** via Hierarchical clustering and quantile cuts

Unsupervised OCC and AD

HDBScan and Quantile Cuts



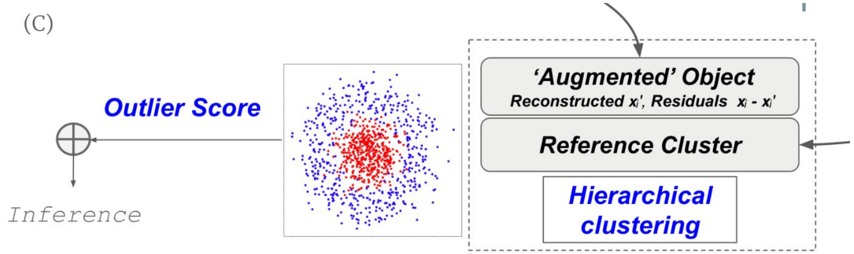


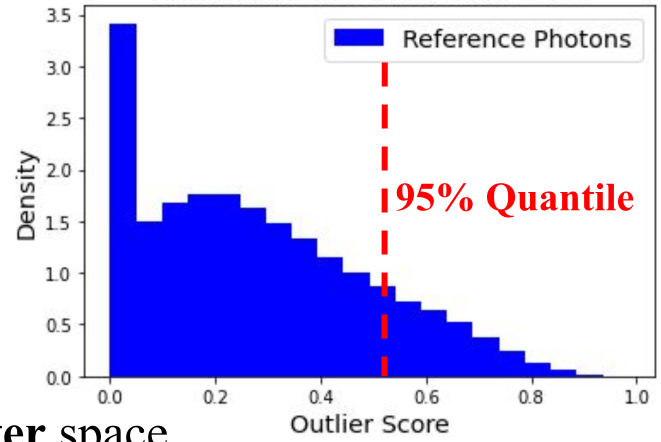
- Augment the inference particle into the **reference cluster** space
  - Two notions of *membership* - density based, distance based
- Combine the two PMF's and extract a probability of membership ($\boldsymbol{P_{in}}$)
- Define *Outlier Score* as complementary probability $\boldsymbol{P_{out} = 1 - P_{in}}$
- Extract **reference population** outlier score corresponding to a desired quantile

University of Regina

Unsupervised OCC and AD

HDBScan and Quantile Cuts



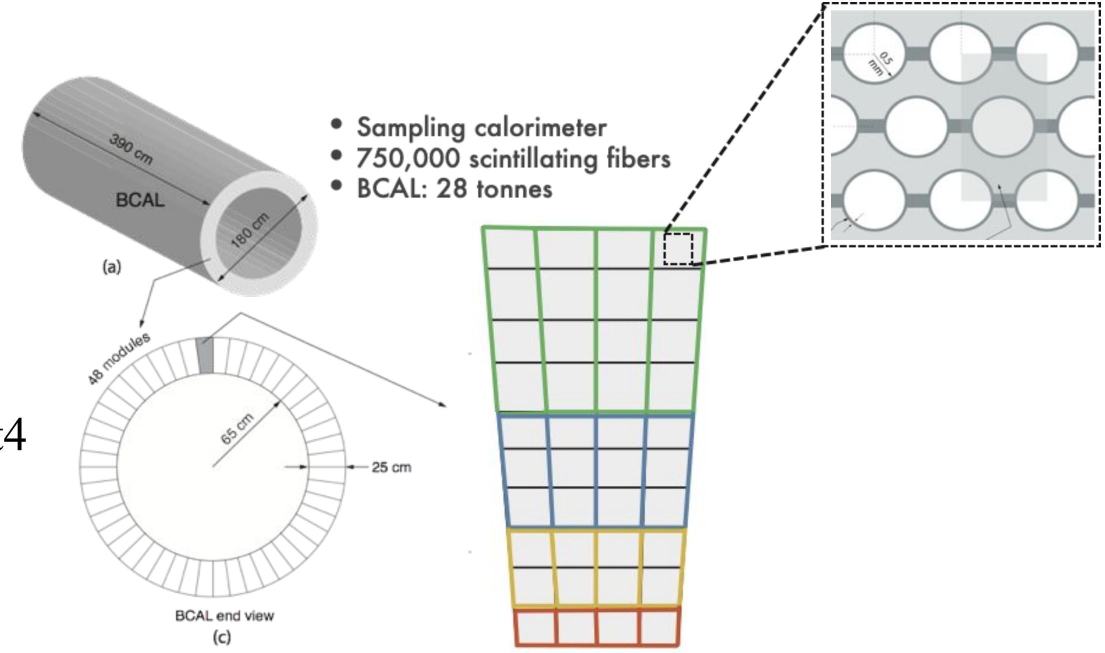


- Augment the inference particle into the **reference cluster** space
  - Two notions of *membership* - density based, distance based
- Combine the two PMF's and extract a probability of membership ($\boldsymbol{P_{in}}$)
- Define *Outlier Score* as complementary probability $\boldsymbol{P_{out} = 1 - P_{in}}$
- Extract **reference population** outlier score corresponding to a desired quantile

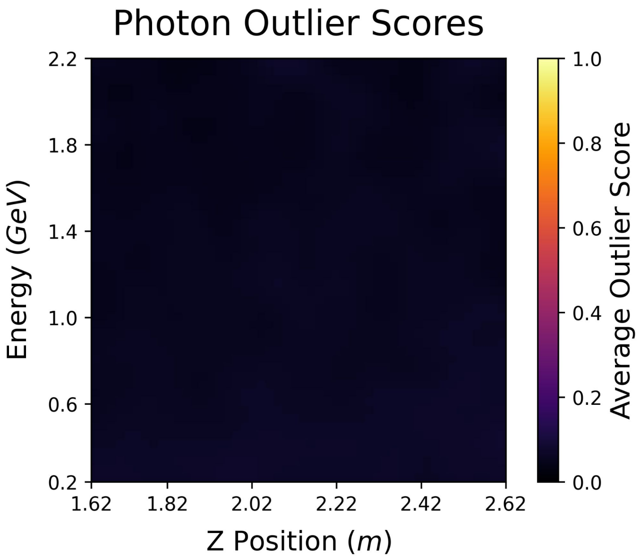
We have defined a dynamic threshold as function of the kinematics, completely agnostic towards the unknown class.

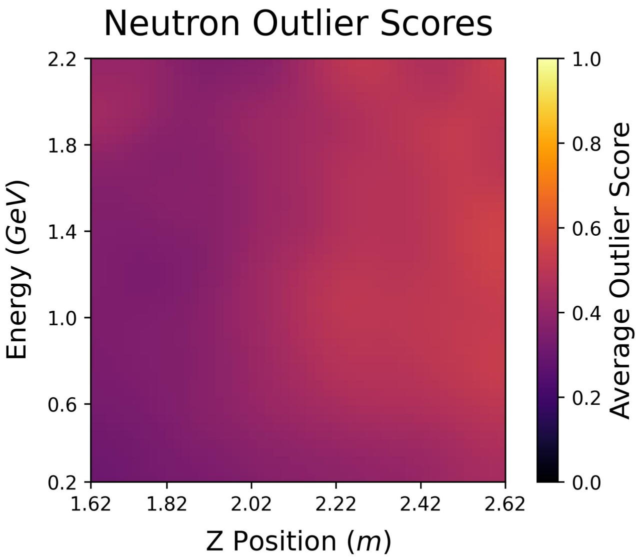
γ/n Separation at GlueX - OCC

- High confidence on **one class**
- Isolate highly active phase space within BCAL
- Reconstructed energy (E) and z position (z) as kinematic conditions
- Simulated photon (**reference**) and neutron (**unknown**) showers - Geant4
- **Strict** preselection cuts
- Deploy fiducial cuts to extract only neutron showers which highly resemble photons
- 14 input features comprising of detector response variables



γ/n Separation at GlueX - Results





Quantile	TPR	TNR
1σ (68%)	**68.28 ± 0.18 %**	**87.44 ± 0.13%**
2σ (95%)	**95.09 ± 0.08 %**	**52.40 ± 0.19%**
3σ (99%)	**98.97 ± 0.04 %**	**34.95 ± 0.18%**

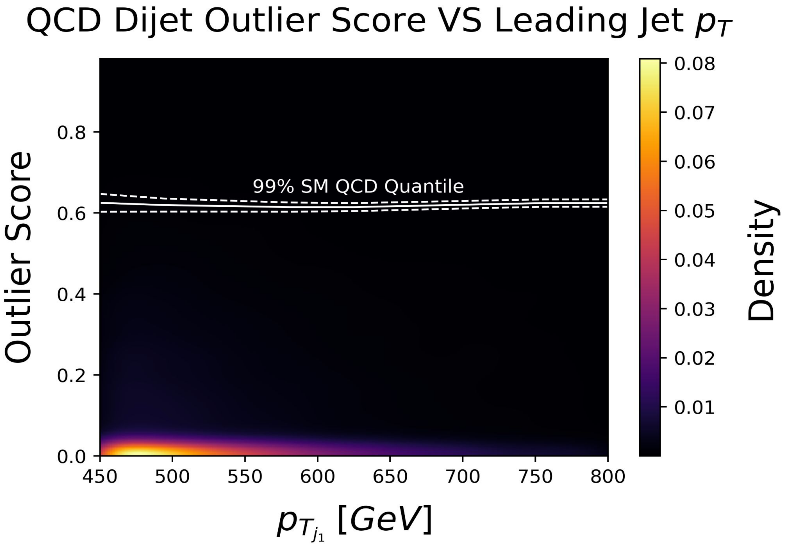
University of Regina

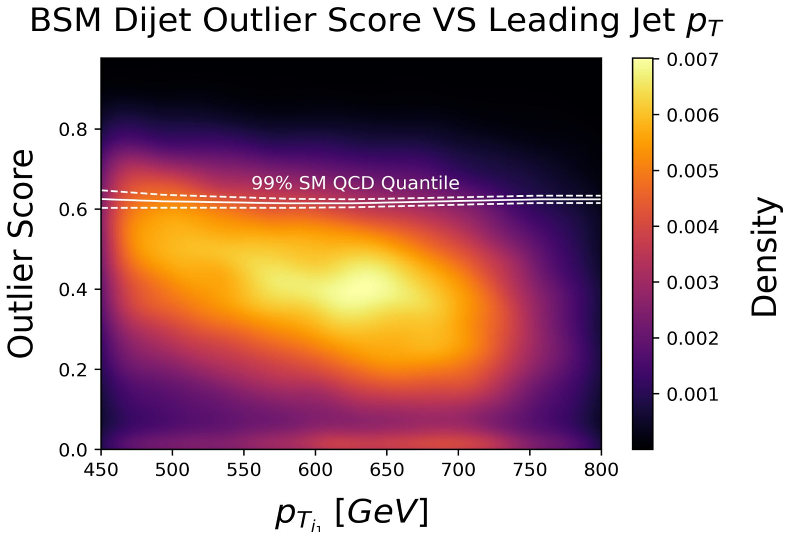
BSM/SM Dijet Separation at LHC - AD

- Consider QCD dijet events as **reference**
- Isolate $Z' \rightarrow t\bar{t}$ dijets as **unknown**
- Publicly available datasets generated via *MADGRAPH* and *Pythia8* using the *DELPHES* framework for fast detector simulation
- Require leading jet transverse momenta 450 GeV < $\boldsymbol{p_T}$ < 800 GeV and sub-leading jet $\boldsymbol{p_T}$ > 200 GeV
- Consider leading jet $\boldsymbol{p_T}$ as single kinematic condition
- 15 input features
    - Remaining 4 vector properties of the leading jet and n-subjettiness variables
    - Sub-leading jet 4 vector and n-subjettiness variables

University of Regina

BSM/SM Dijet Separation at LHC - Results





Quantile	TPR	TNR
1σ (68%)	**68.18 ± 0.22 %**	**93.20 ± 0.06%**
2σ (95%)	**95.15 ± 0.10 %**	**42.40 ± 0.22%**
3σ (99%)	**99.03 ± 0.05 %**	**11.82 ± 0.14%**
Fiducial cuts (99%)	**98.92 ± 0.05 %**	**2.35 ± 0.06%**

	Ours	Fraser et al.	Cheng et al.
AUC	**0.885 ± 0.003**	0.87	0.89

University of Regina

Summary

- Our architecture removes the need for semi-supervised approaches
  - Agnostic threshold selection
  - Totally unsupervised
- Highly dependable return rate (TPR = Quantile)
- Flexible - deployable on various problems
- Lends itself naturally to the role of data monitoring within detectors

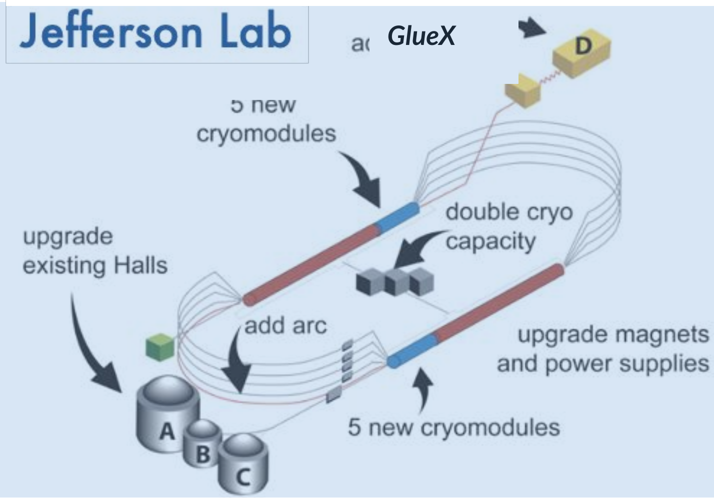
Thank you!

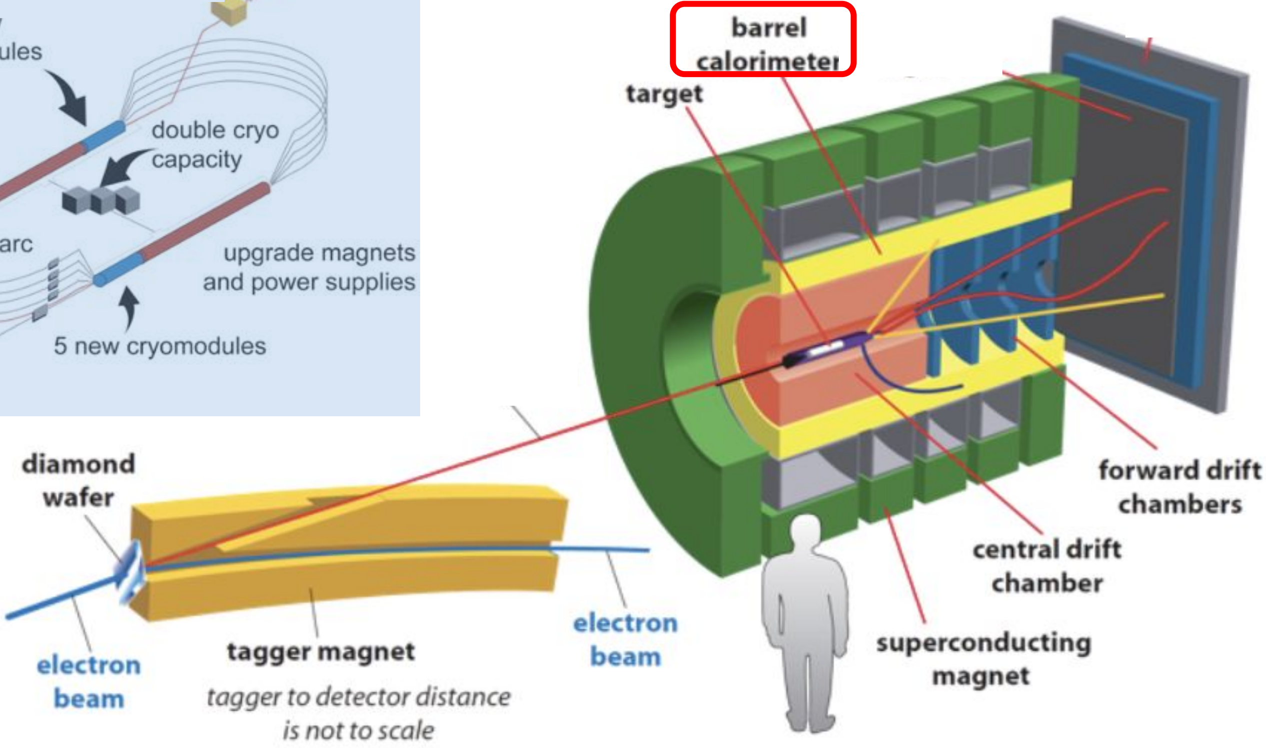
NSERC CRSNG compute canada | calcul canada U.S. DEPARTMENT OF ENERGY Office of Science

arXiv

Backup Slides

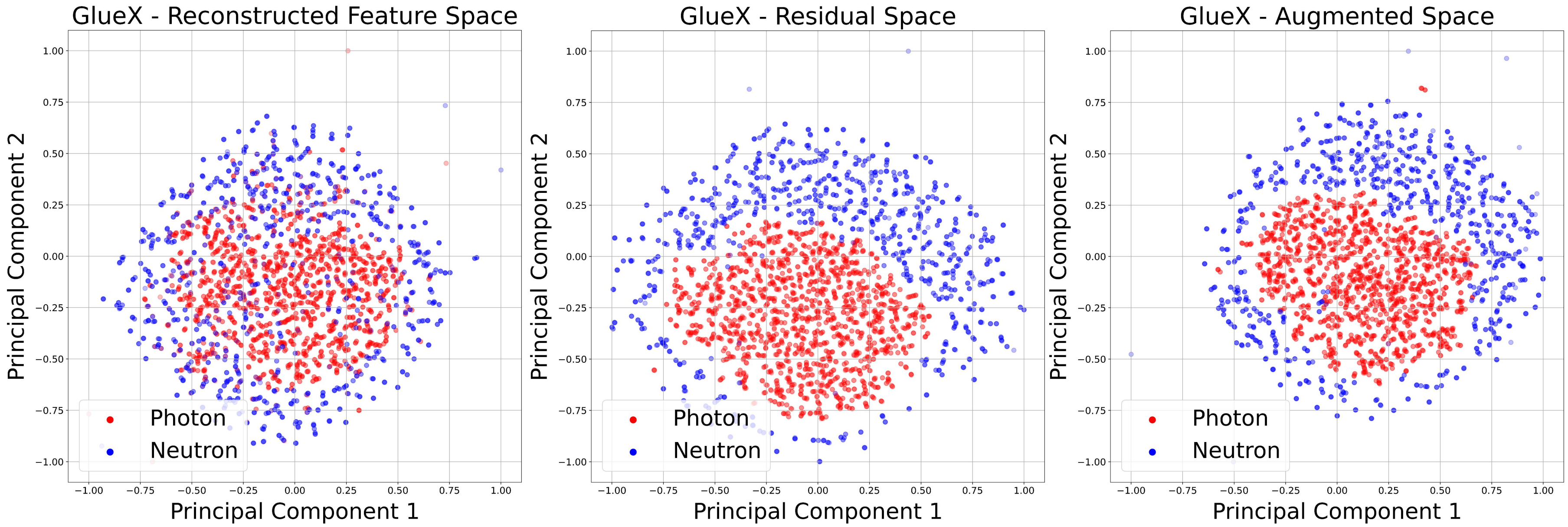
The Beamline





University of Regina

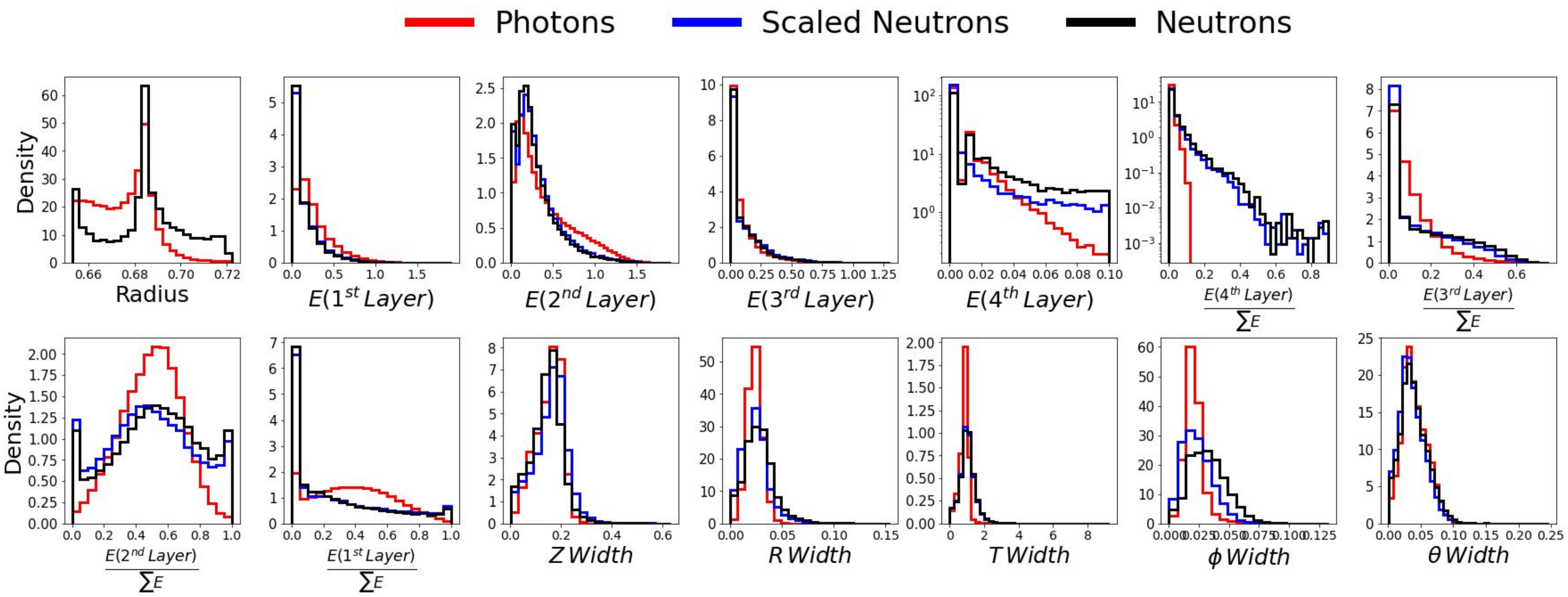
γ/n Separation at GlueX - Residuals



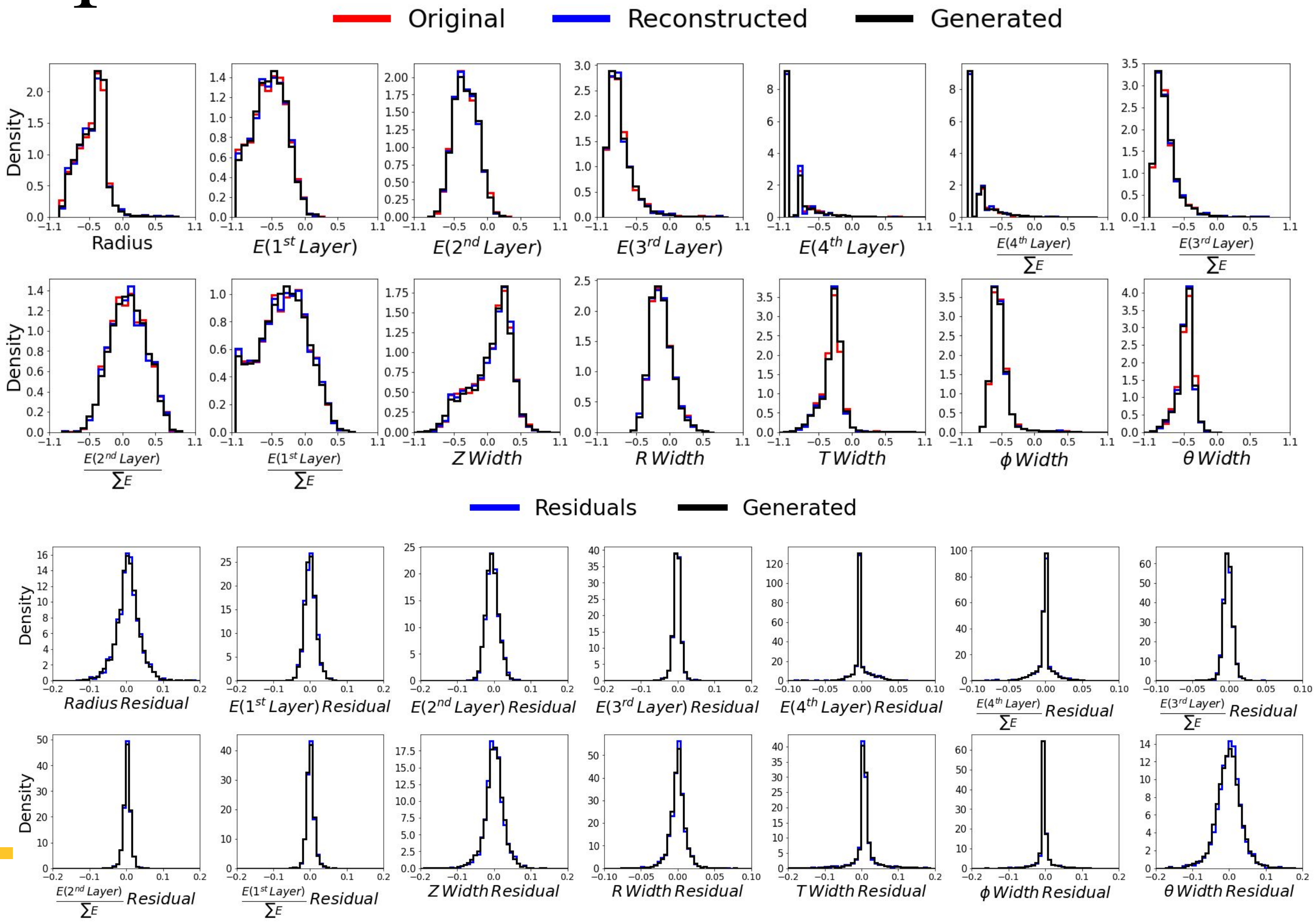
Features localize the space, residuals push nested clusters radially outward.

University of Regina

γ/n Separation at GlueX - Features

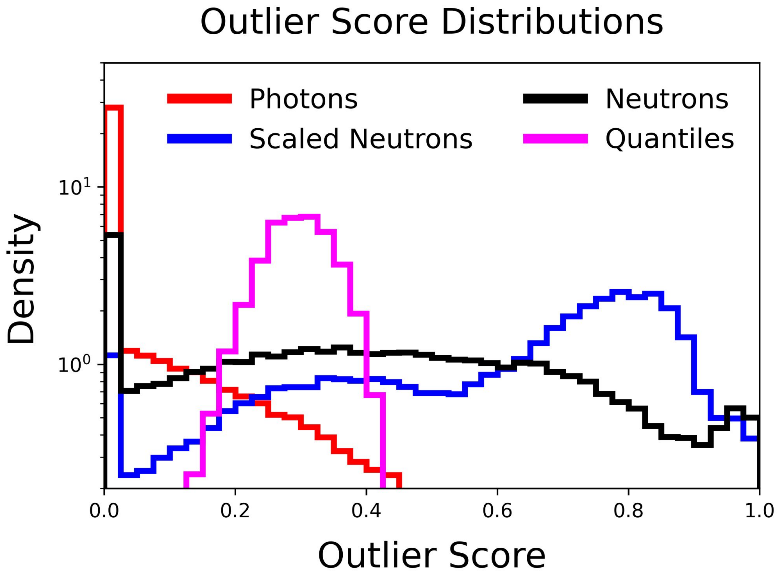


γ/n Separation at GlueX - Generations



Benefits of Conditional Learning

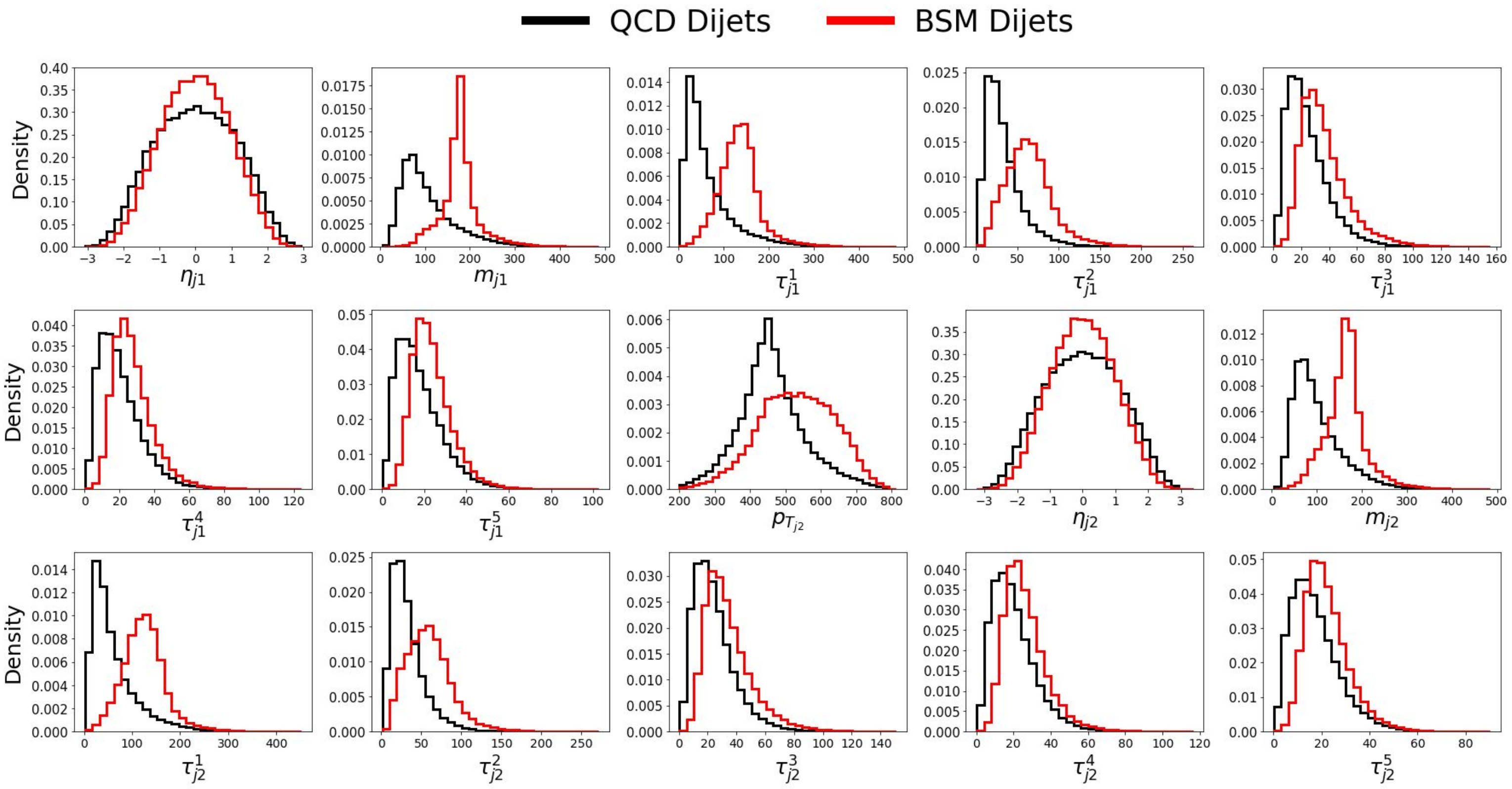
- Perturb neutrons such that they are almost indistinguishable from photons
  - Considered "Actual" detector response
- F+M trained on only photons
- XGBoost trained on unperturbed neutron sample along with photons
- XGBoost given access to E and z as features
- Neutron kinematic correlations picked up via F+M residuals - average outlier score increased



	Simulation		"Actual" Detector Response	
Algorithm	**TPR**	**TNR**	**TPR**	**TNR**
XGBoost	$92.15 \pm 0.10\%$	$91.93 \pm 0.10\%$	$92.15 \pm 0.10\%$	$\mathbf{78.82} \pm 0.15\%$
F + M (Augmented)	$92.28 \pm 0.10\%$	$60.29 \pm 0.18\%$	$92.33 \pm 0.10\%$	$\mathbf{82.71} \pm 0.14\%$
F + M (Features)	$92.34 \pm 0.10\%$	$56.14 \pm 0.19\%$	$92.34 \pm 0.10\%$	$50.30 \pm 0.19\%$

University of Regina

BSM/SM Dijet Separation at LHC - Features



BSM/SM Dijet Separation at LHC - Generations