LUDWIG-
MAXIMILIANS
UNIVERSITY
MUNICH

YonedaVAE: Self-Supervised Out-of-Distribution Multi-Set Generation for Amortized Simulation and Inverse Problems

Baran Hashemi

Ludwig-Maximilians-Universität München

Currently at ***The ORIGINS Cluster Munich (TUM)***

Motivation

From HEP side:

To do a **full detector** simulation with **irregular geometry** and **ultra-high granularity**?

To Generate detector signatures for kinematic/luminosity regions where **data is very rare**?

Motivation

From HEP side:

To do a **full detector** simulation with **irregular geometry** and **ultra-high granularity**?

To Generate detector signatures for kinematic/luminosity regions where **data is very rare**?

From ML side:

To Reach "**Context Extrapolation**" in Inverse design problems?

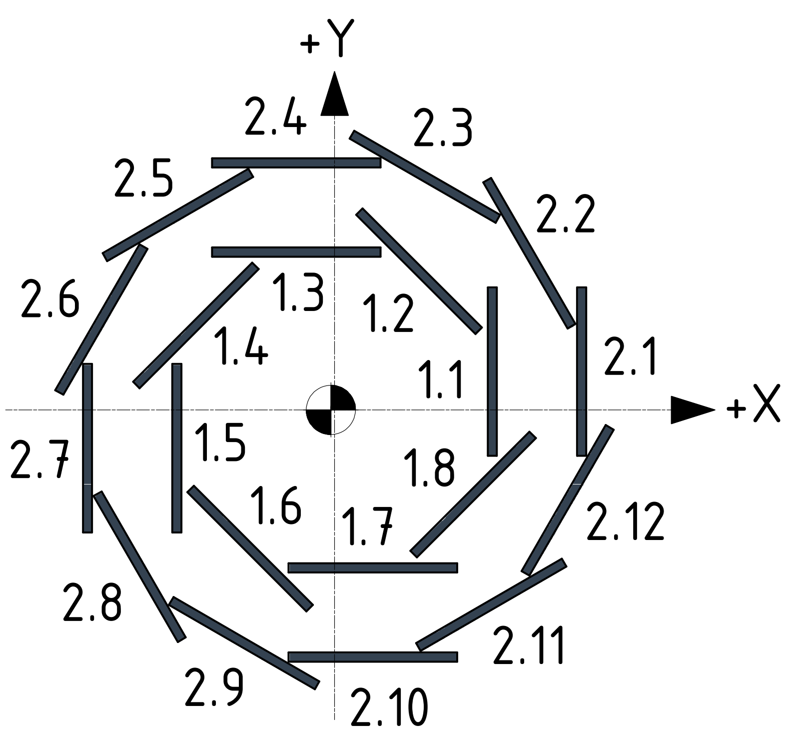
Enhance the VAE prior with **Self-Supervised Learning**?

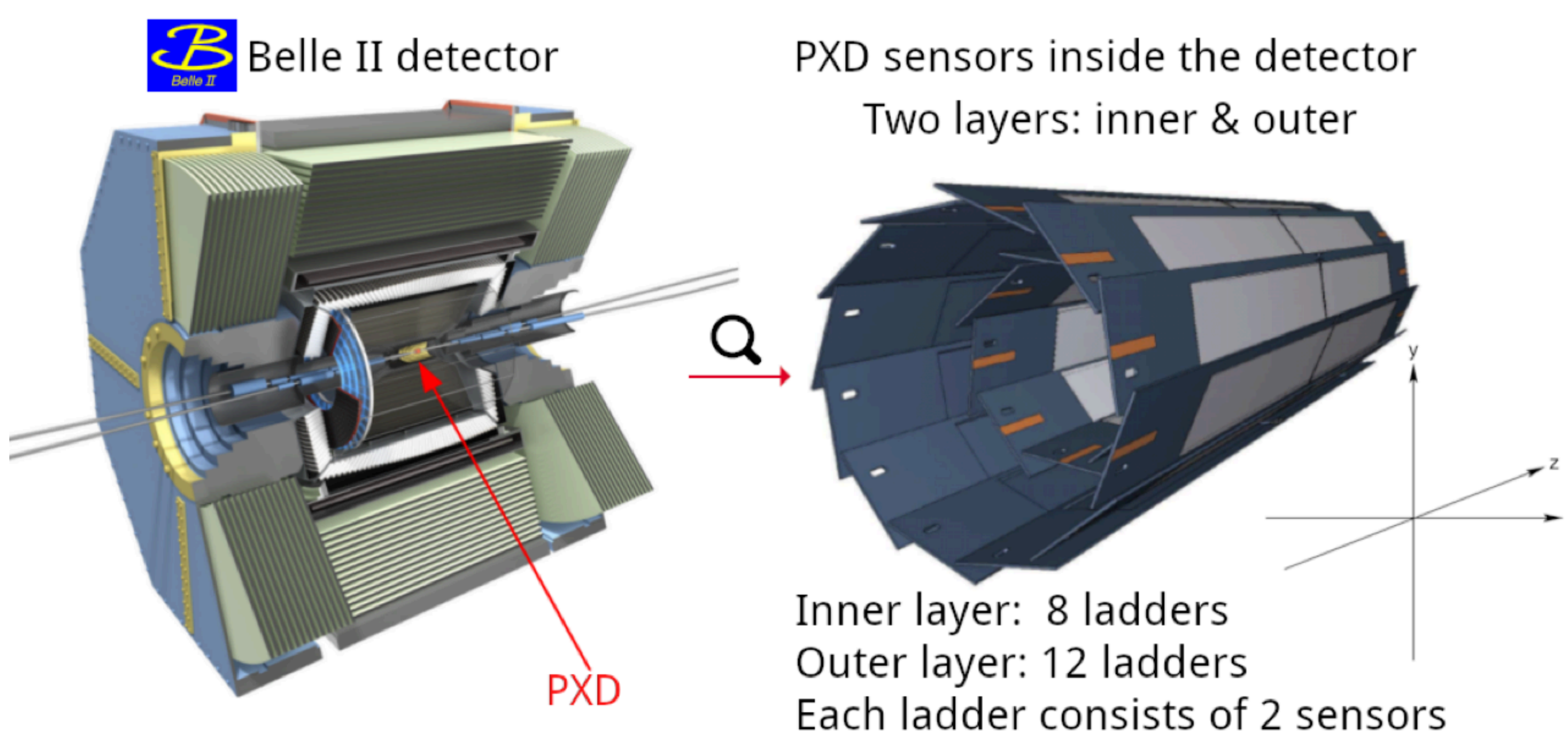
Experiment

- **The Pixel Vertex Detector (PXD):** sub-detector for charged tracks at Belle II.
- **+7.5M channels, [40,250,768],** per event —> "**Ultra-High Granularity**"
- Having a **Toroid** topology:

The inner layer: 16 sensors, **The outer layer:** 24 sensors

(only 19 sensors were installed)





PXD Generation: YonedaVAE

- Problem description:

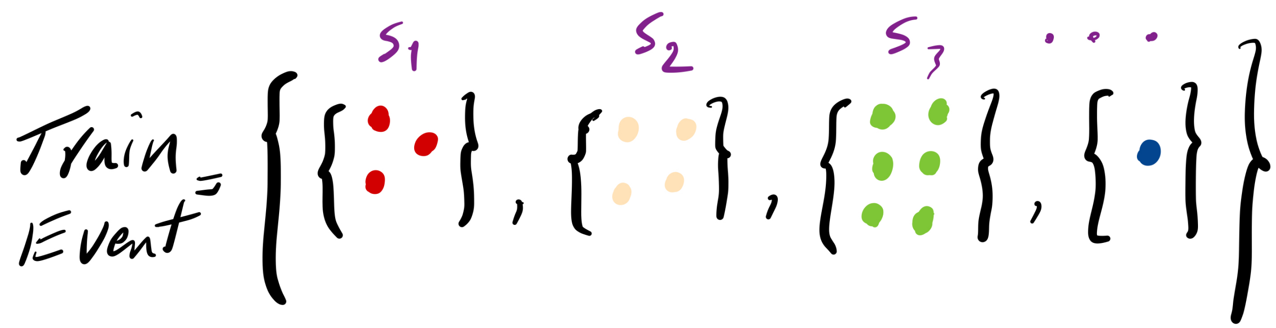
A. Training:

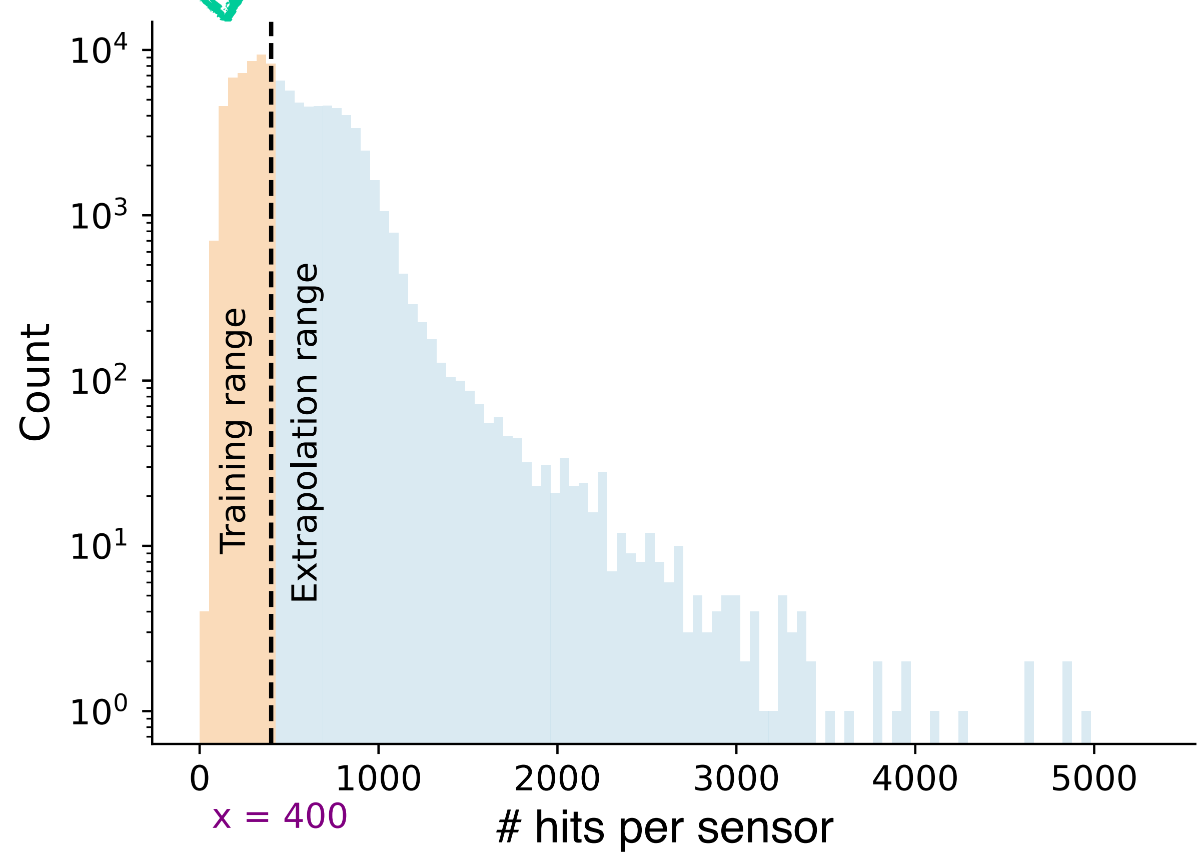
$$\mathscr{D}_{train} = \{(\mathbf{X}^{(m)}; \mathbf{e}^{(m)}; \mathbf{c}^{(m)})\}_{m=1}^{19}$$

$\mathbf{X}$: Set of hits, $|\mathbf{X}^{(m)}| = \mathrm{N}^{(m)}$

$\mathbf{e}$: Event-level attribute

$\mathbf{c}$: Sensor-level condition





PXD Generation: YonedaVAE

- Problem description:

A. Training:

$$\mathscr{D}_{train} = \{(\mathbf{X}^{(m)}; \mathbf{e}^{(m)}; \mathbf{c}^{(m)})\}_{m=1}^{19}$$

$\mathbf{X}$: Set of hits, $|\mathbf{X}^{(m)}| = \mathrm{N}^{(m)}$

$\mathbf{e}$: Event-level attribute

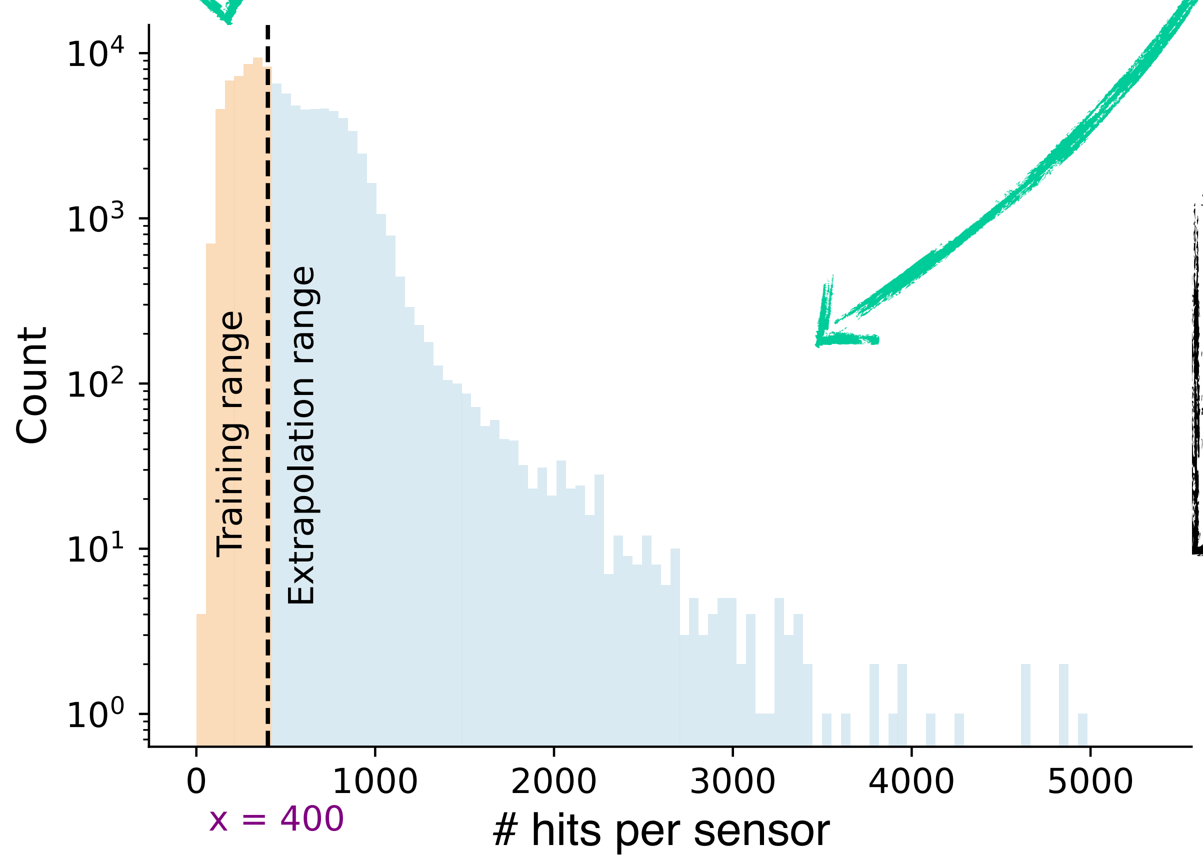
$\mathbf{c}$: Sensor-level condition

B. Test: Zero-Shot

$$\mathscr{D}_{test} = \{\mathbf{e}^{(m)}\}_{m=1}^{19}$$

$\mathbf{e}$: Event-level attribute

s.t $N_{max}^{test} \gg N_{max}^{train}$



Up to 110,000 hit points per event

PXD Generation: YonedaVAE

- Problem description:

A. Training:

$$\mathscr{D}_{train} = \{(\mathbf{X}^{(m)}; \mathbf{e}^{(m)}; \mathbf{c}^{(m)})\}_{m=1}^{19}$$

$\mathbf{X}$: Set of hits, $|\mathbf{X}^{(m)}| = \mathrm{N}^{(m)}$

$\mathbf{e}$: Event-level attribute

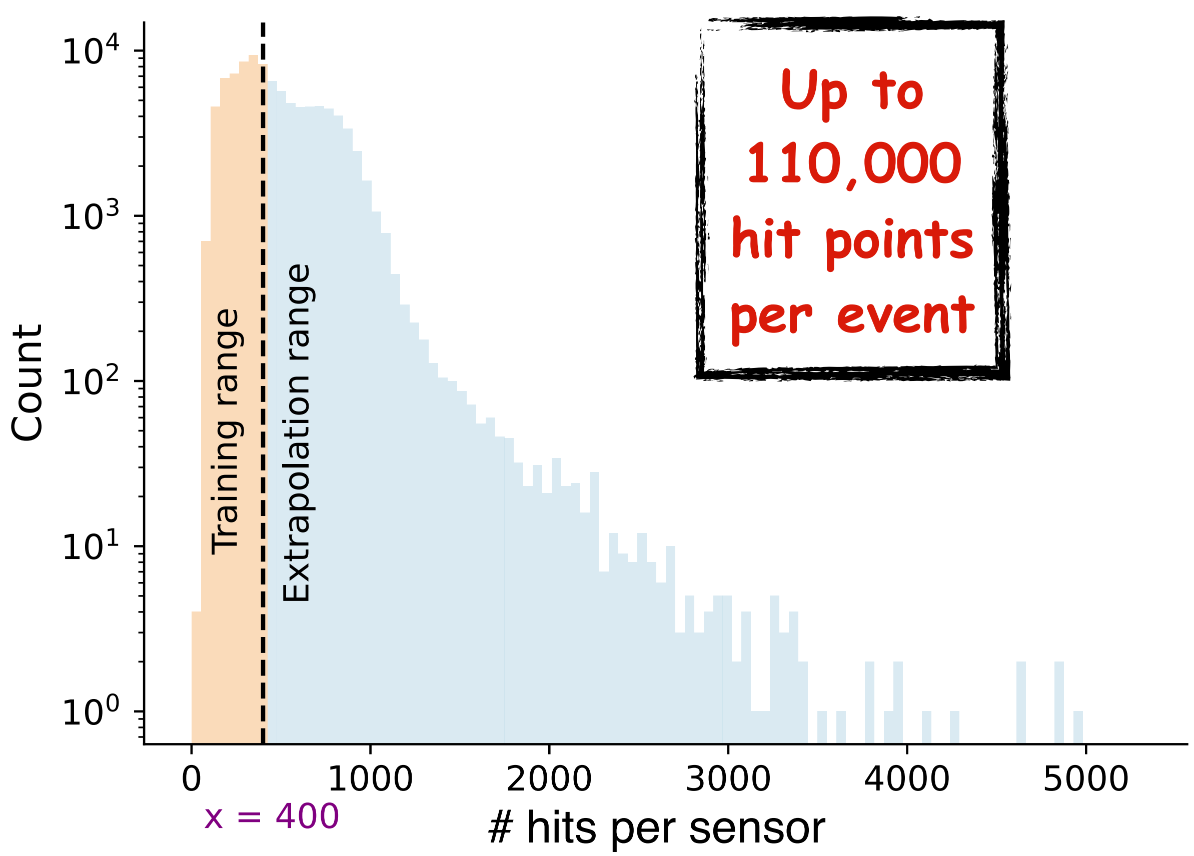
$\mathbf{c}$: Sensor-level condition

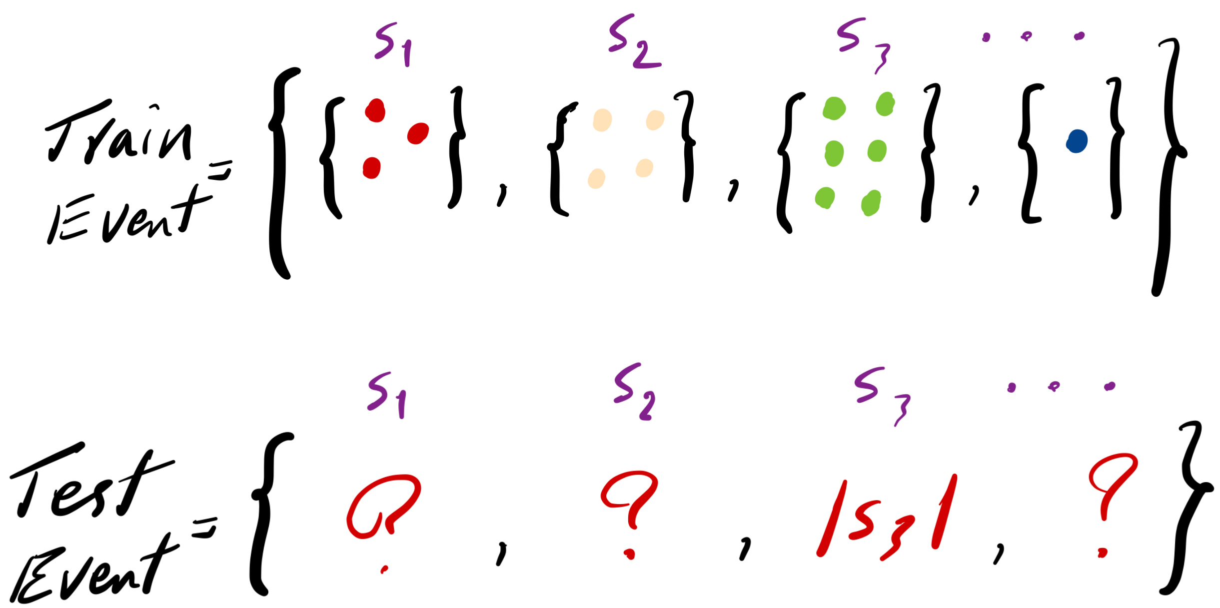
B. Test: Zero-Shot

$$\mathscr{D}_{test} = \{\mathbf{e}^{(m)}\}_{m=1}^{19}$$

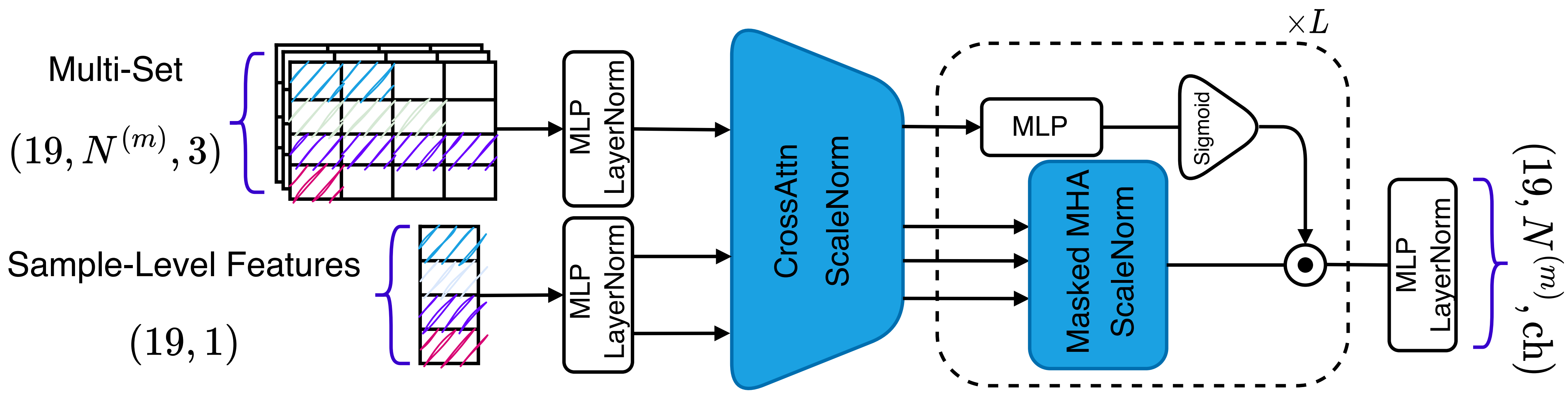
$\mathbf{e}$: Event-level attribute

s.t $N_{max}^{test} \gg N_{max}^{train}$



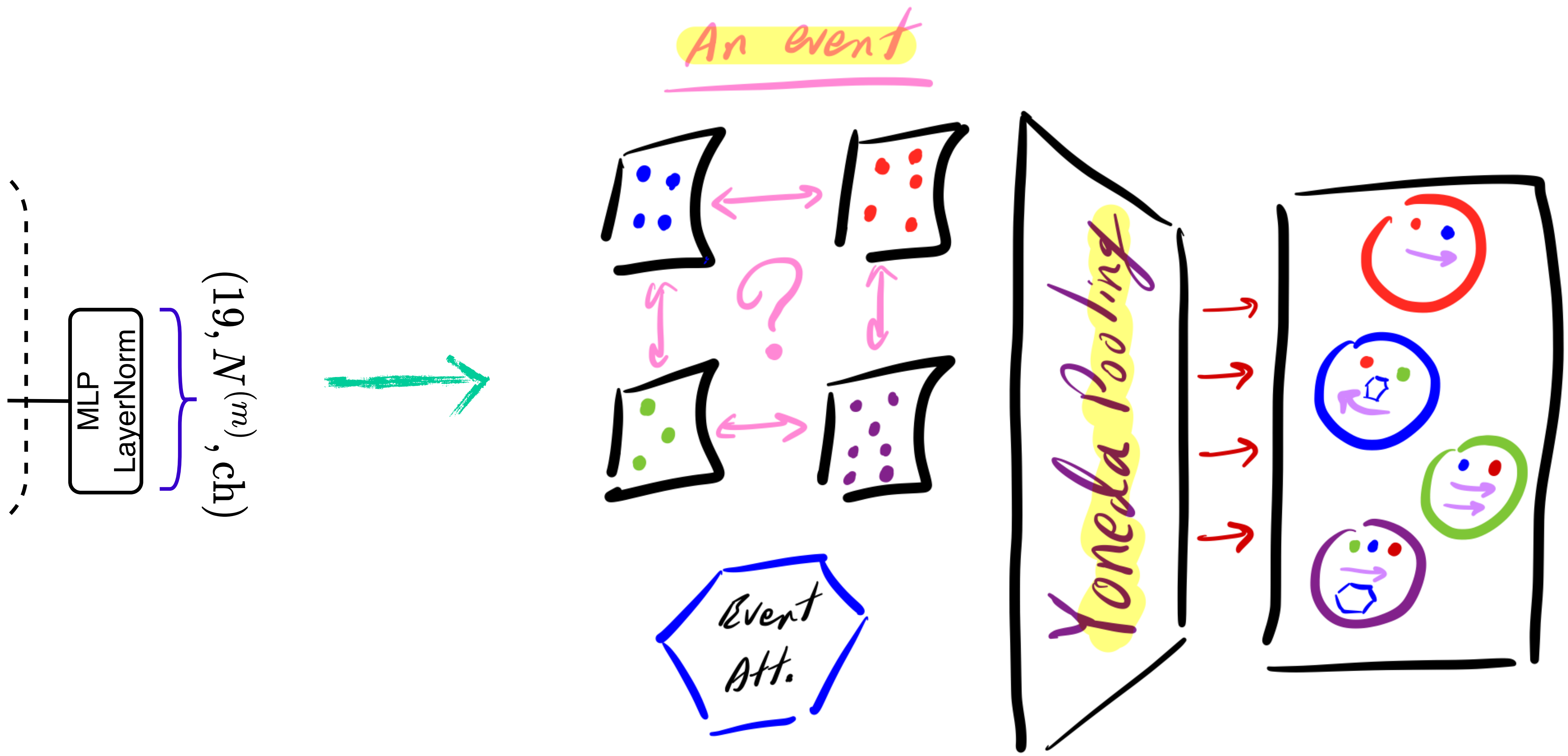


1. **Encoder:** EventFormer



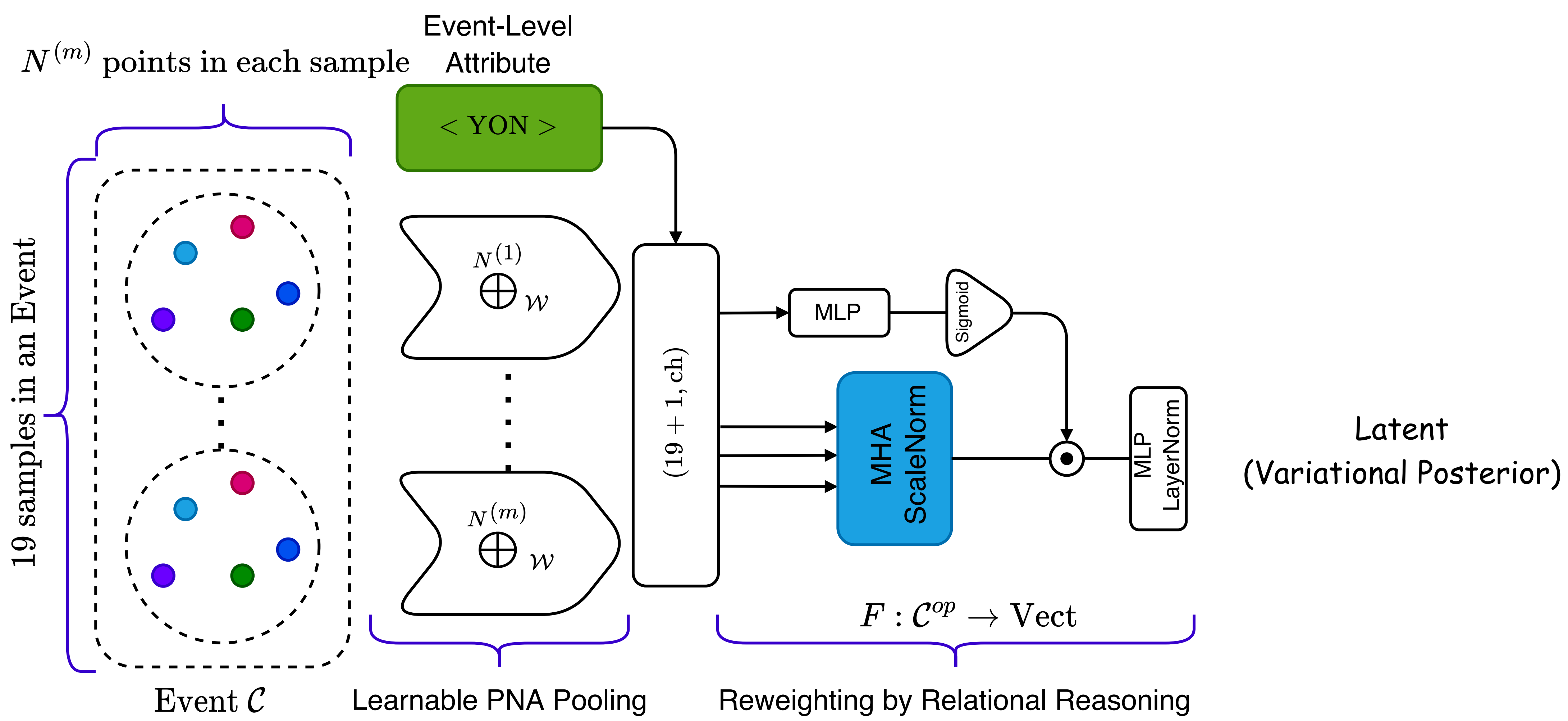
YonedaVAE: Encoder

1. **Encoder:** EventFormer —> Yoneda Pooling



YonedaVAE: Encoder

1. Encoder: EventFormer —> Yoneda Pooling



YonedaVAE: Set Gen.

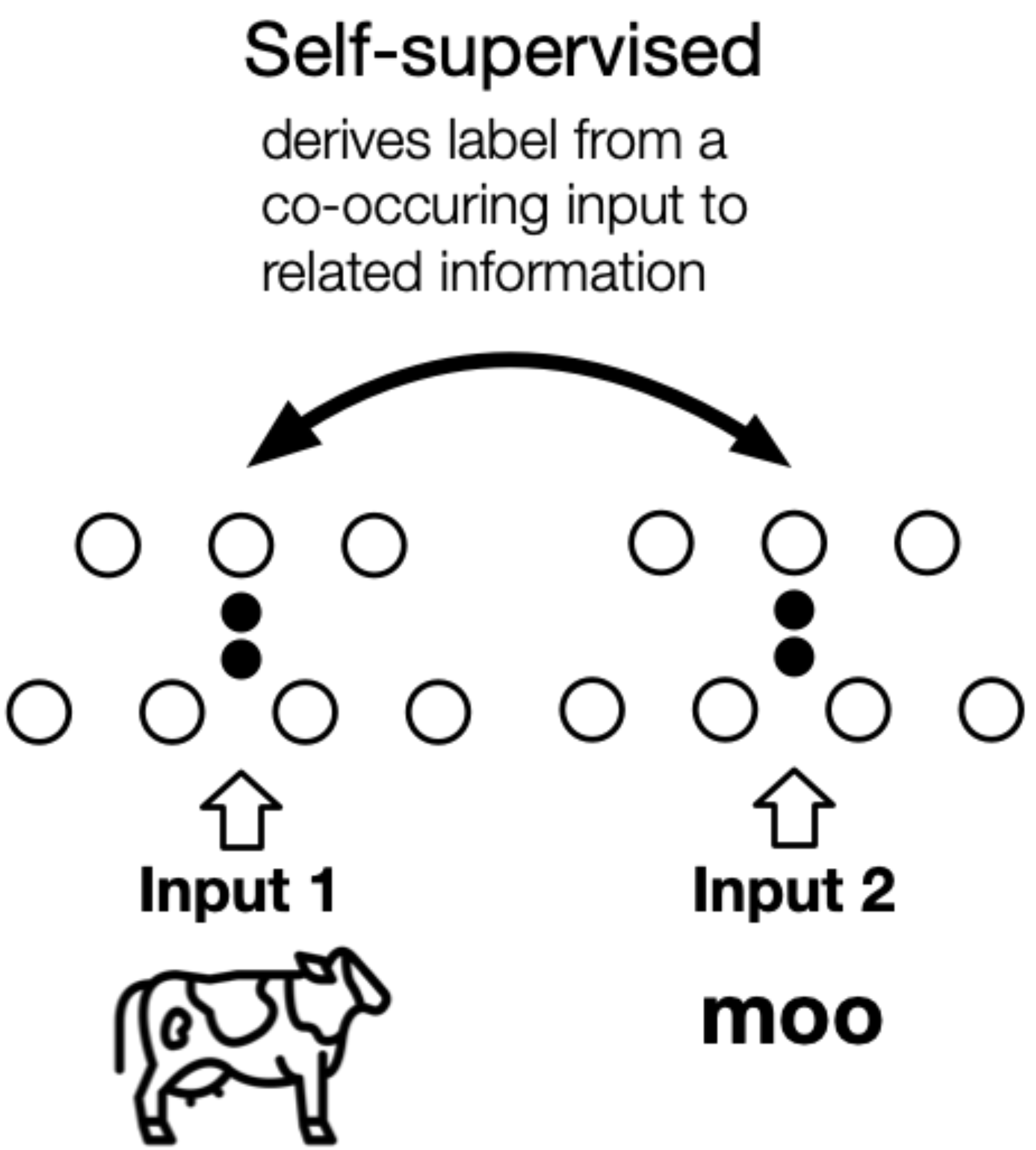
2. **Set Generator:**

Question:

How to learn a better prior to help the model generalize better to unseen data at test time?

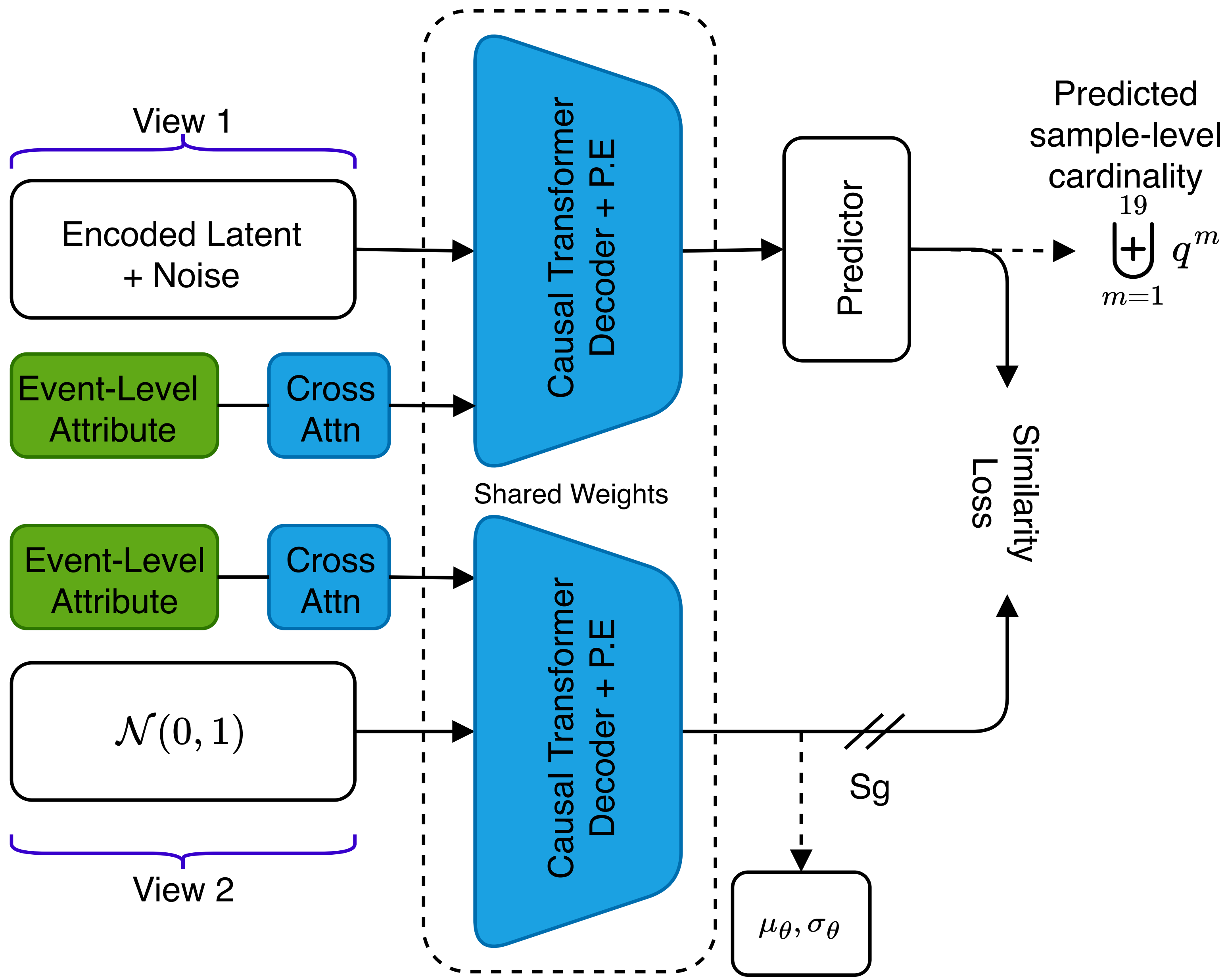
YonedaVAE: Set Gen.

2. **Set Generator:** Learning prior à la Self-Supervised Learning



YonedaVAE: Set Gen.

2. **Set Generator:** Learning prior à la Self-Supervised Learning



They go as the Prior to match the Variational Posterior

2. **Set Generator:**

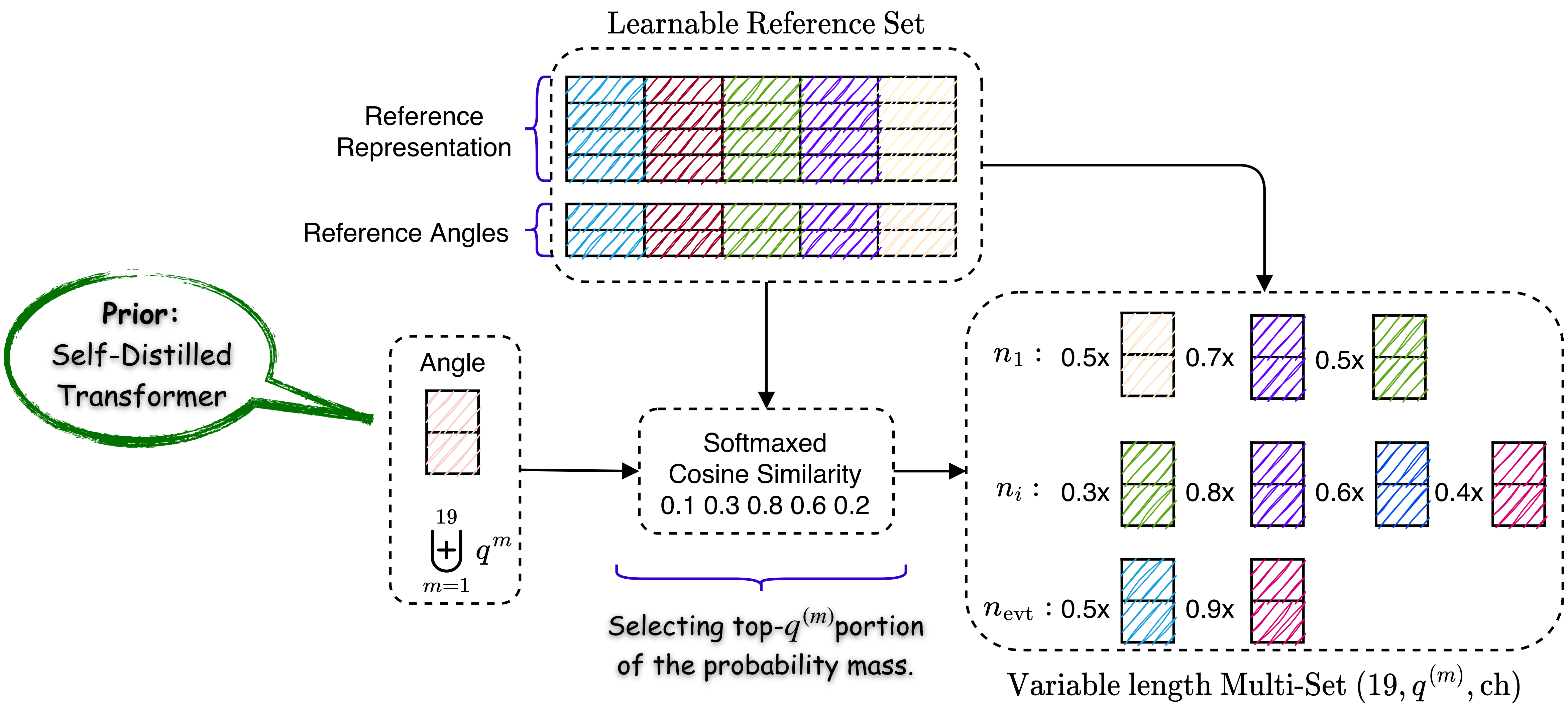
Now we learned a prior and predicted sample-level cardinality

Question:

How should we populate the points point cloud for an irregular setup?

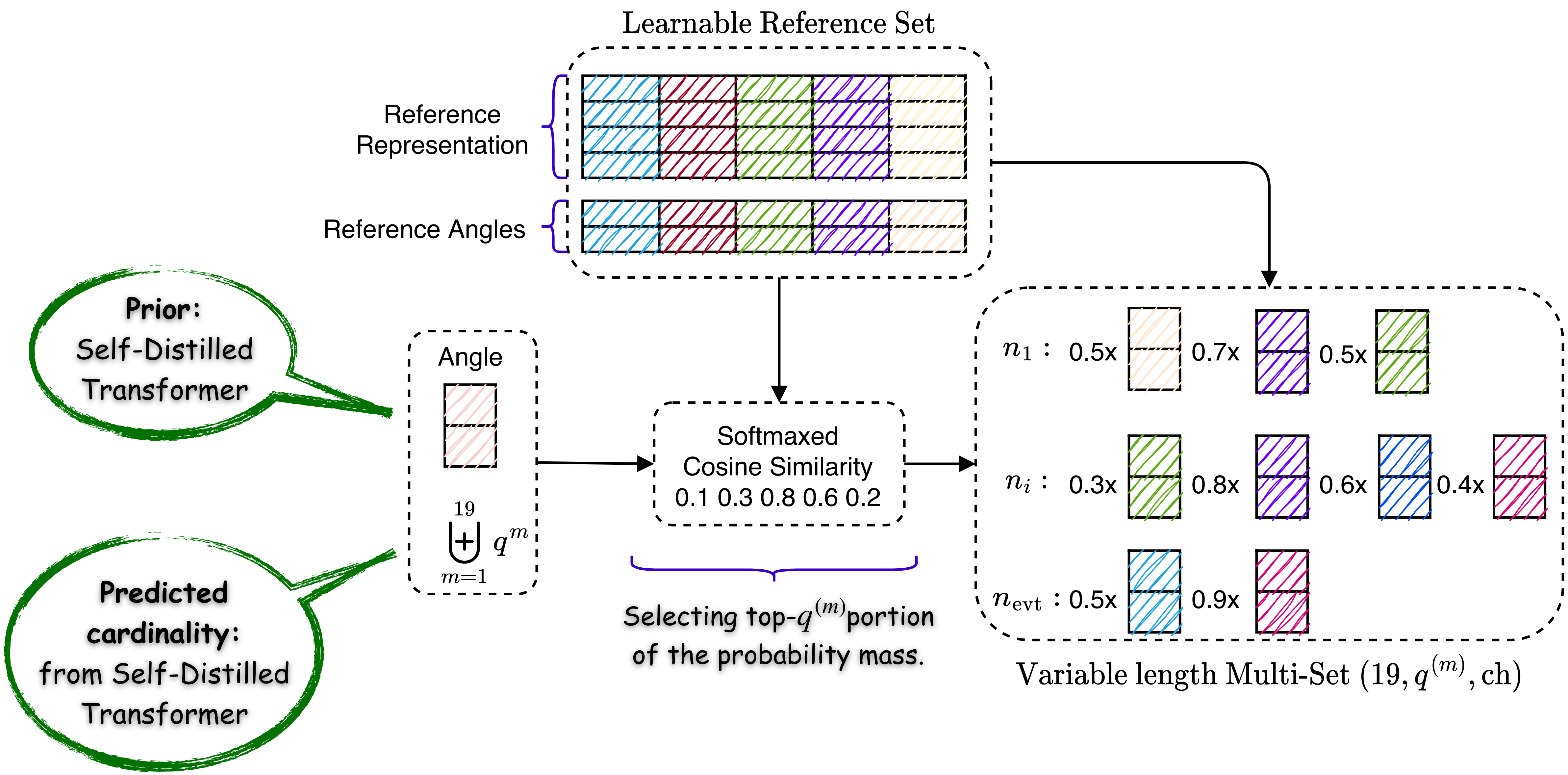
YonedaVAE: Set Gen.

2. **Set Generator:** Adaptive Top-q Sampling



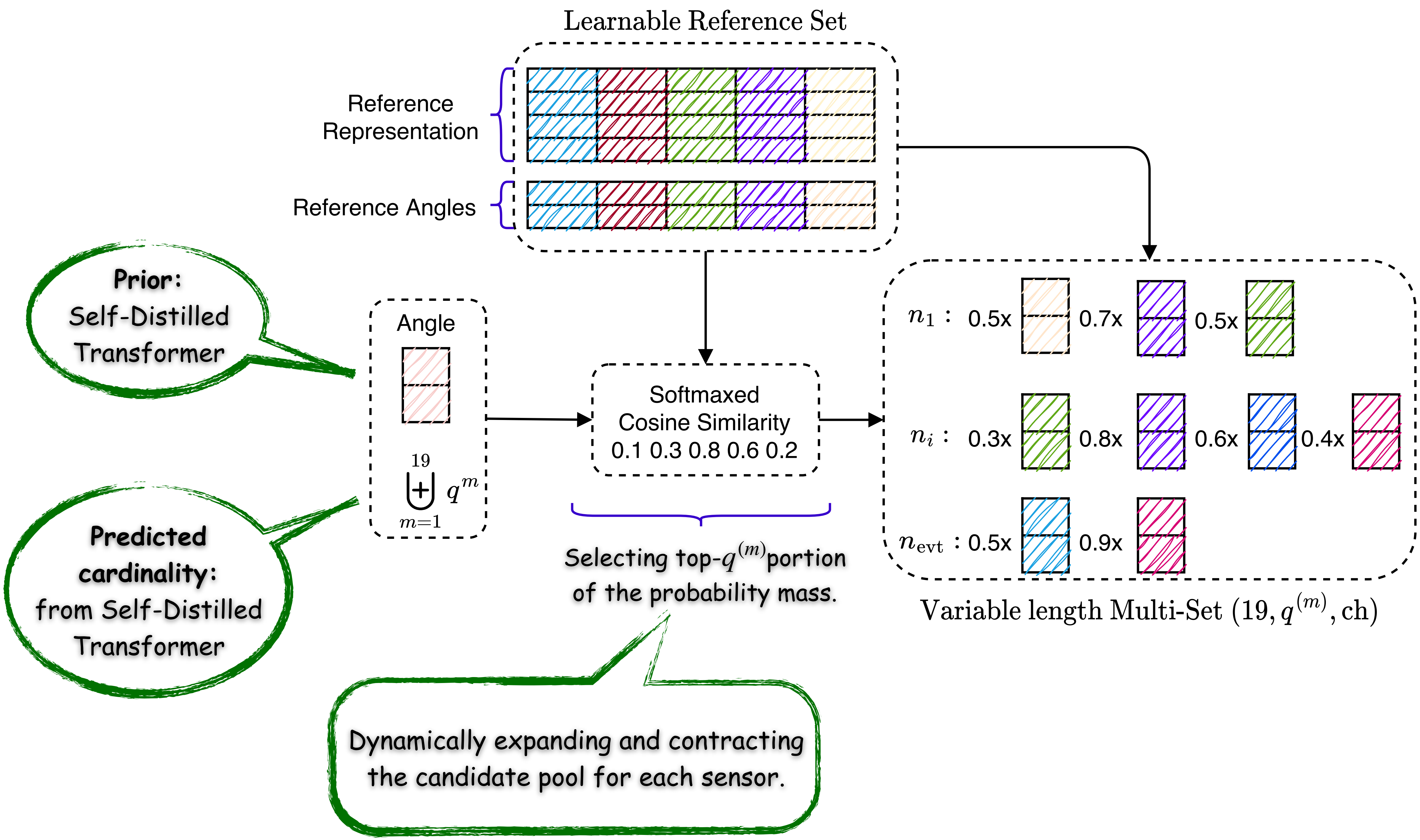
YonedaVAE: Set Gen.

2. **Set Generator:** Adaptive Top-q Sampling



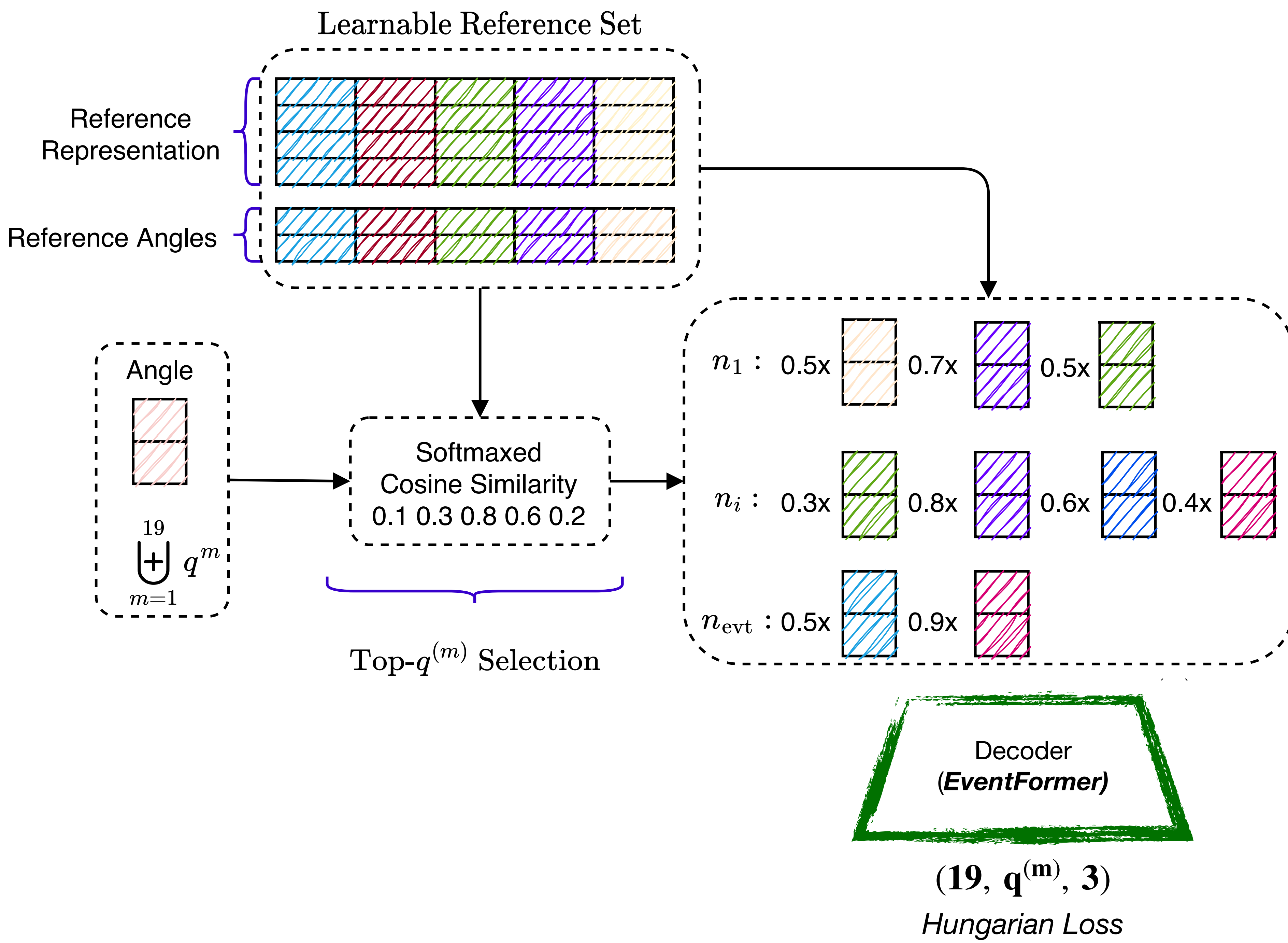
YonedaVAE: Set Gen.

2. **Set Generator:** Adaptive Top-q Sampling



YonedaVAE: Decoder

3. Decoder: Modified EventFormer



YonedaVAE: Results

YonedaVAE: Results

NN-based Metrics

FID/KID backbone trained on the complete dataset

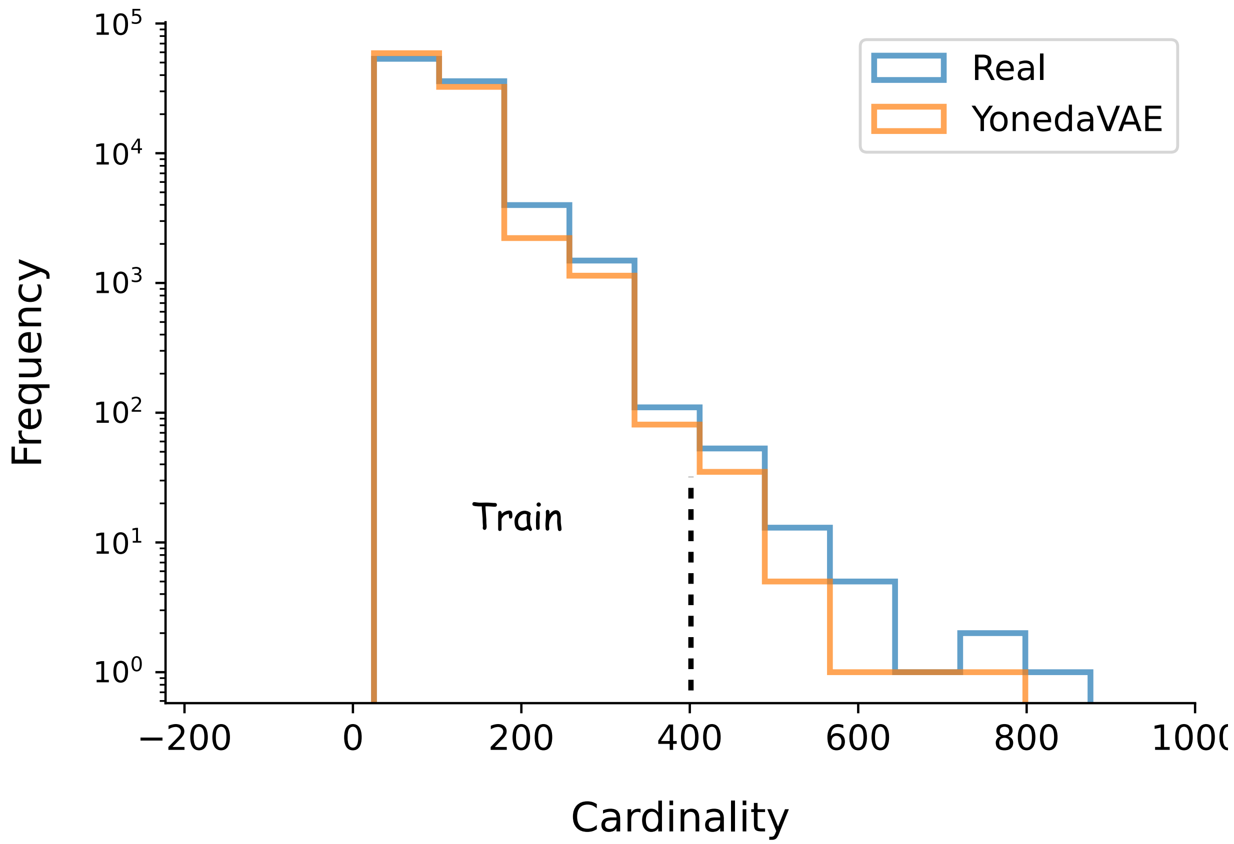
The Lower the FID/KID, the better sample quality

	TSPN (i.i.d)[1]	TSPN (Top-k)[2]	IEA-GAN[3]	Set-VAE[4]	YonedaVAE	Test Data
FID	49.46 ± 0.29	41.40 ± 0.48	37.84 ± 0.98	33.49 ± 0.11	$\mathbf{20.19} \pm 0.31$	0
KID$^{(\times 10^{-4})}$	339 ± 7	312 ± 1	283 ± 8	181 ± 2	$\mathbf{130 \pm 2}$	0

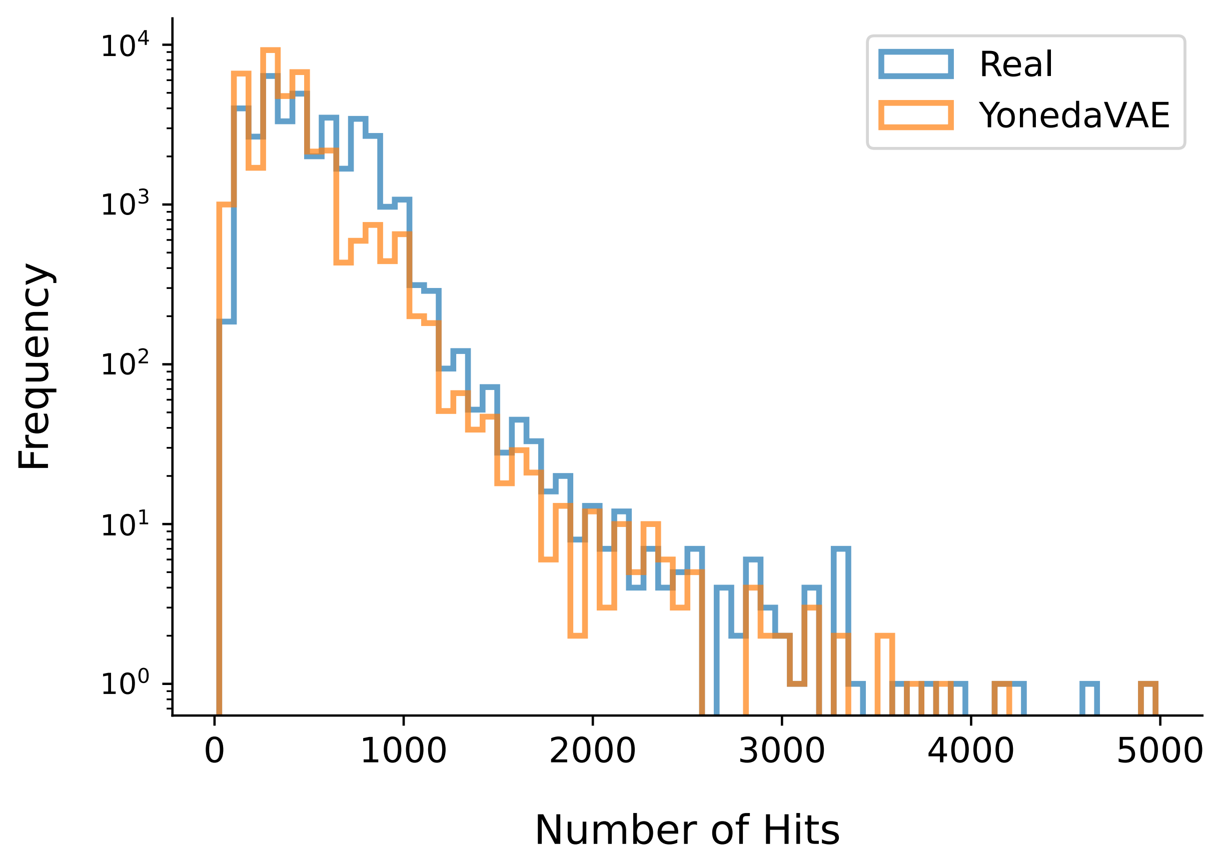
1: 2006.16841, Kosiorek et al. **2:** 2110.02096 Vignac et al. **4:** 2303.08046 Hashemi et al. **5.** 2103.15619 Kim et al.

YonedaVAE: Results

Train/Val Set (ID data), Lum 1.42×10^{34} cm^{-2}s^{-1}



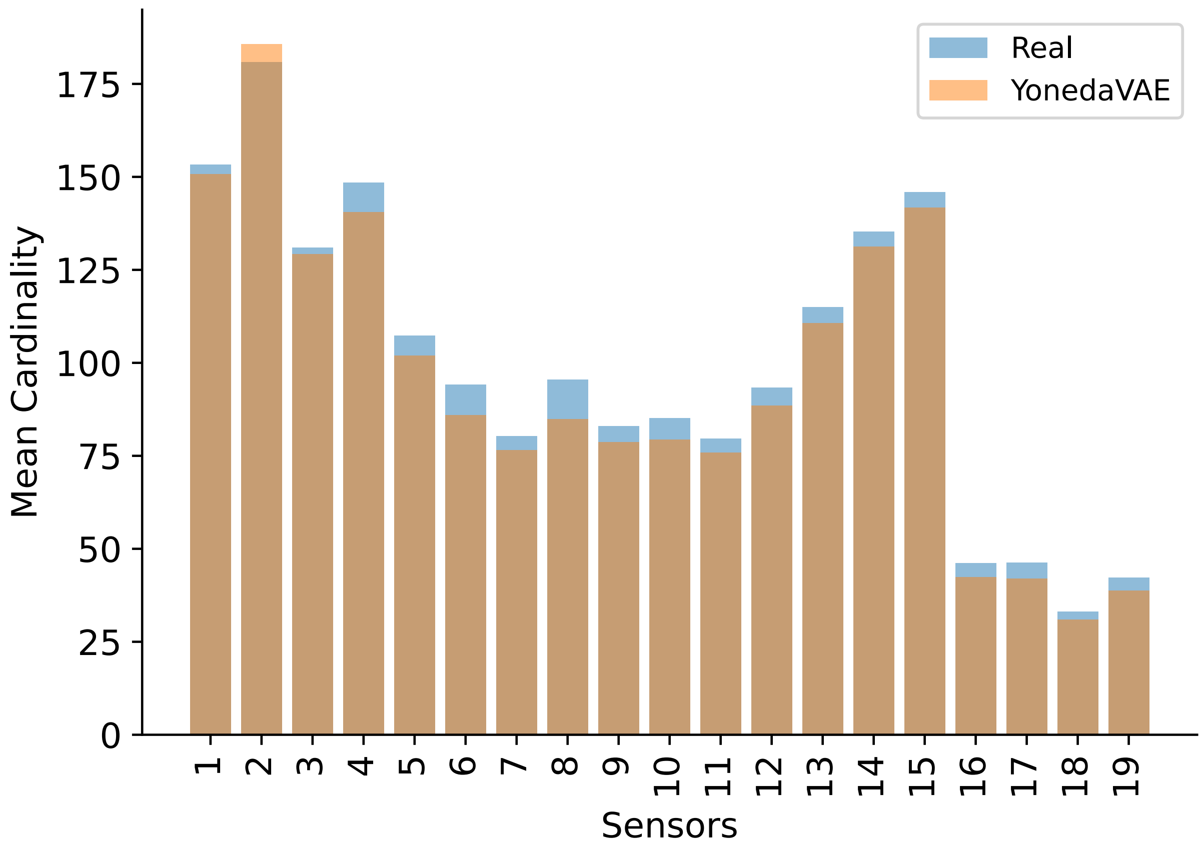
Test (OOD data), Lum. $\mathbf{2.68 \times 10^{34}}$ cm^{-2}s^{-1}



YonedaVAE: Results

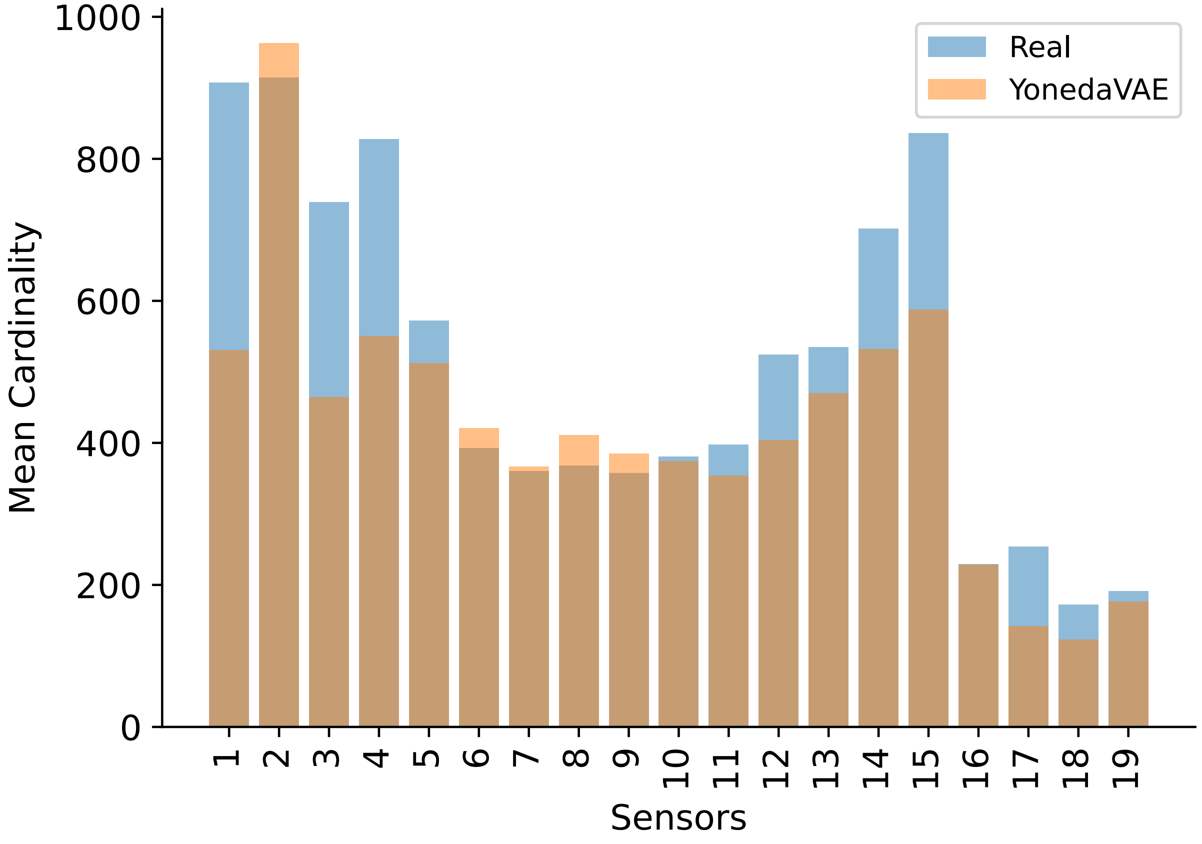
Train/Val Set (ID data), Lum 1.42×10^{34} cm^{-2}s^{-1}

mean # hits
1500
per event



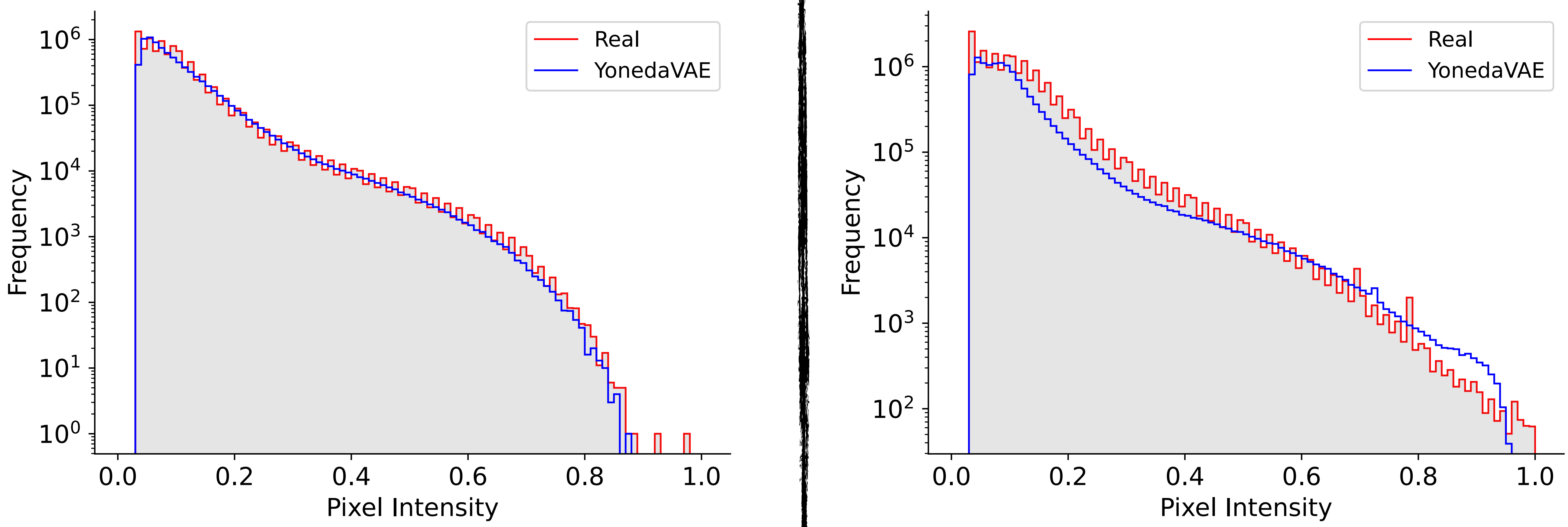
Test (OOD data), Lum. $\mathbf{2.68 \times 10^{34}}$ cm^{-2}s^{-1}

mean # hits
10,000
per event



YonedaVAE: Results

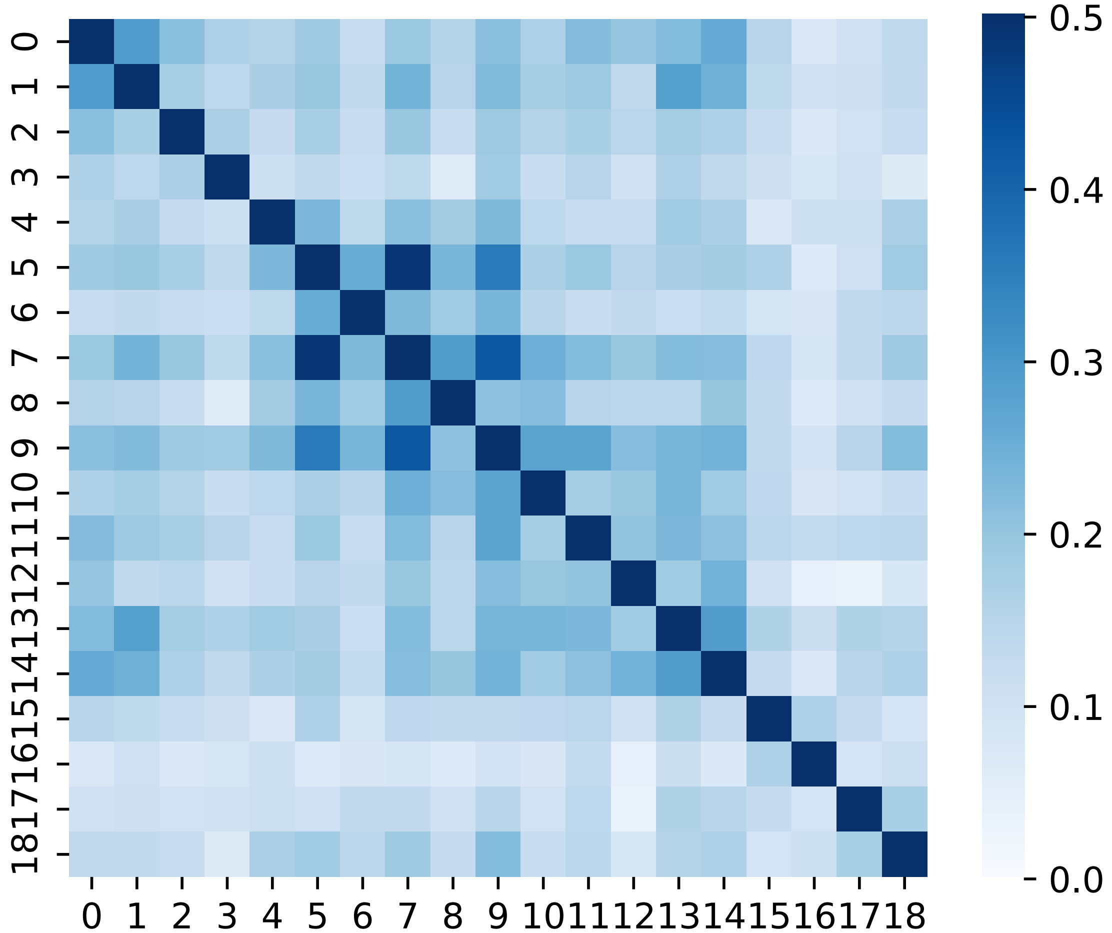
Train/Val Set (ID data), Lum 1.42×10^{34} cm^{-2}s^{-1}	Test (OOD data), Lum. $\mathbf{2.68 \times 10^{34}}$ cm^{-2}s^{-1}

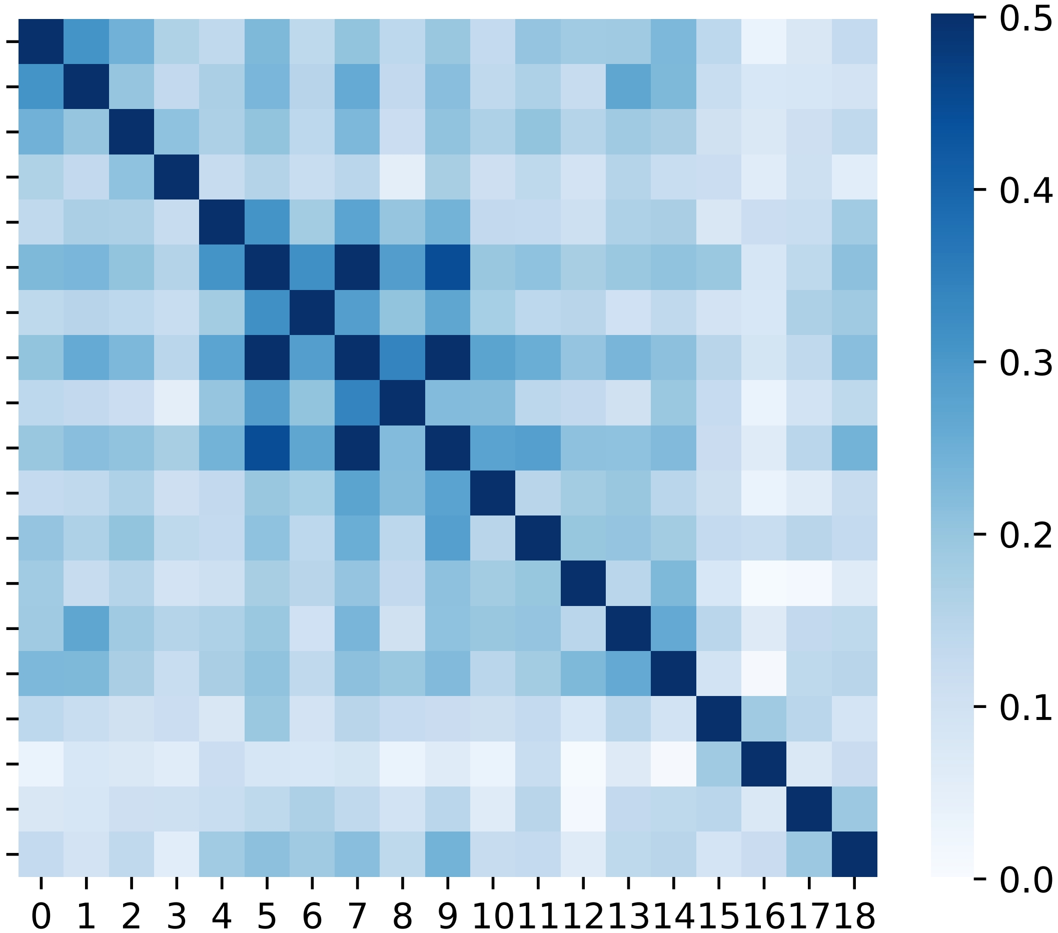


YonedaVAE: Results

Train/Val Set (ID data), Lum 1.42×10^{34} cm^{-2}s^{-1}

Correlation between the number of hits between 19 sensors of PXD

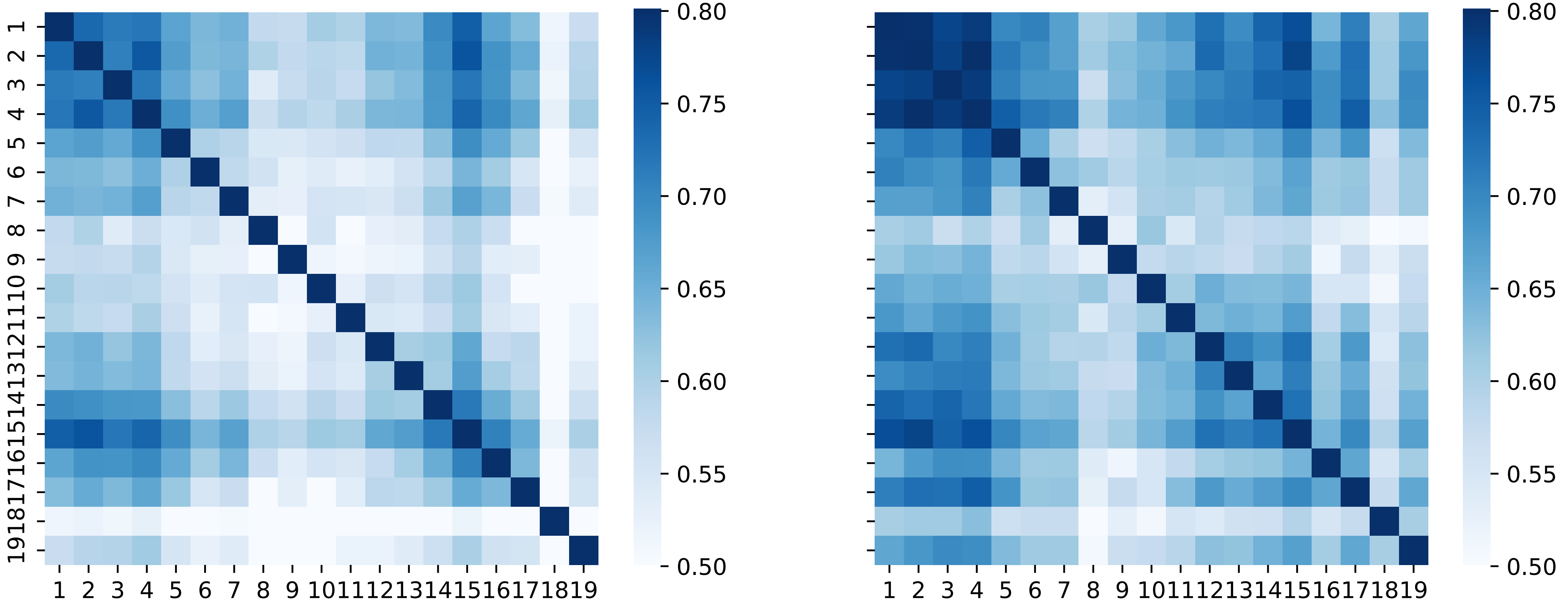




YonedaVAE: Results

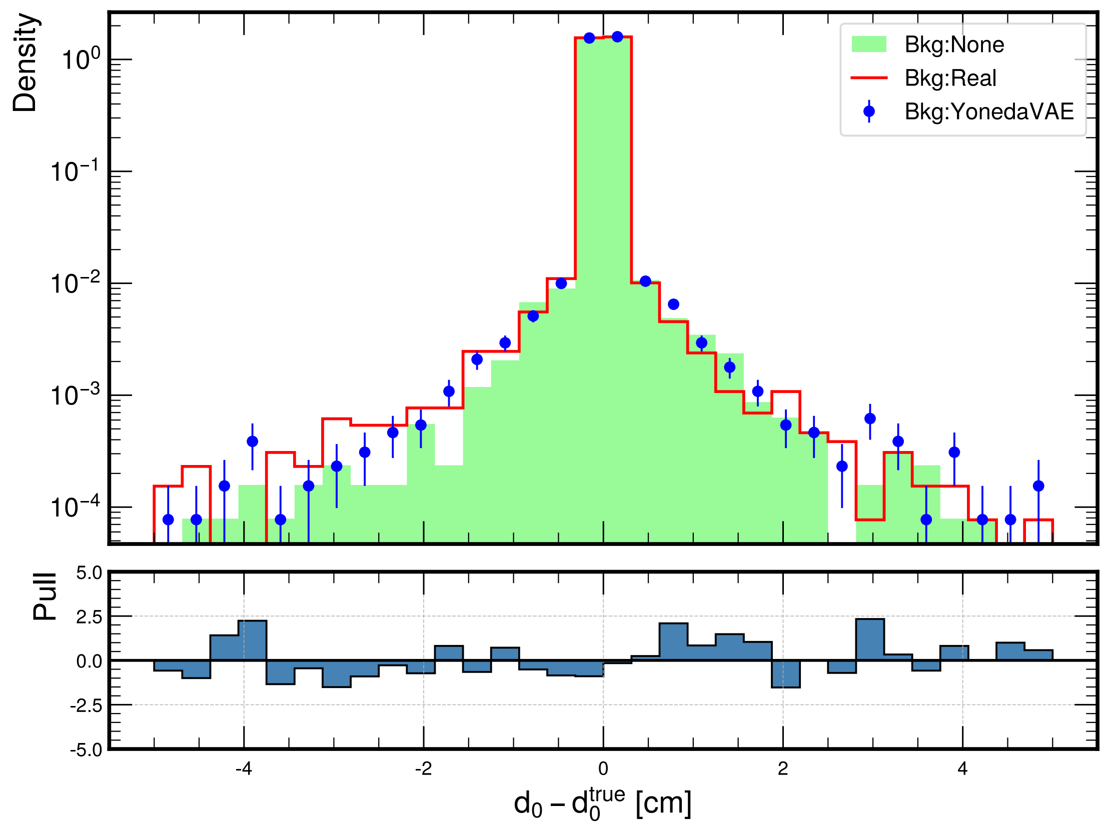
Test (OOD data), Lum. 2.68×10^{34} cm^{-2}s^{-1}

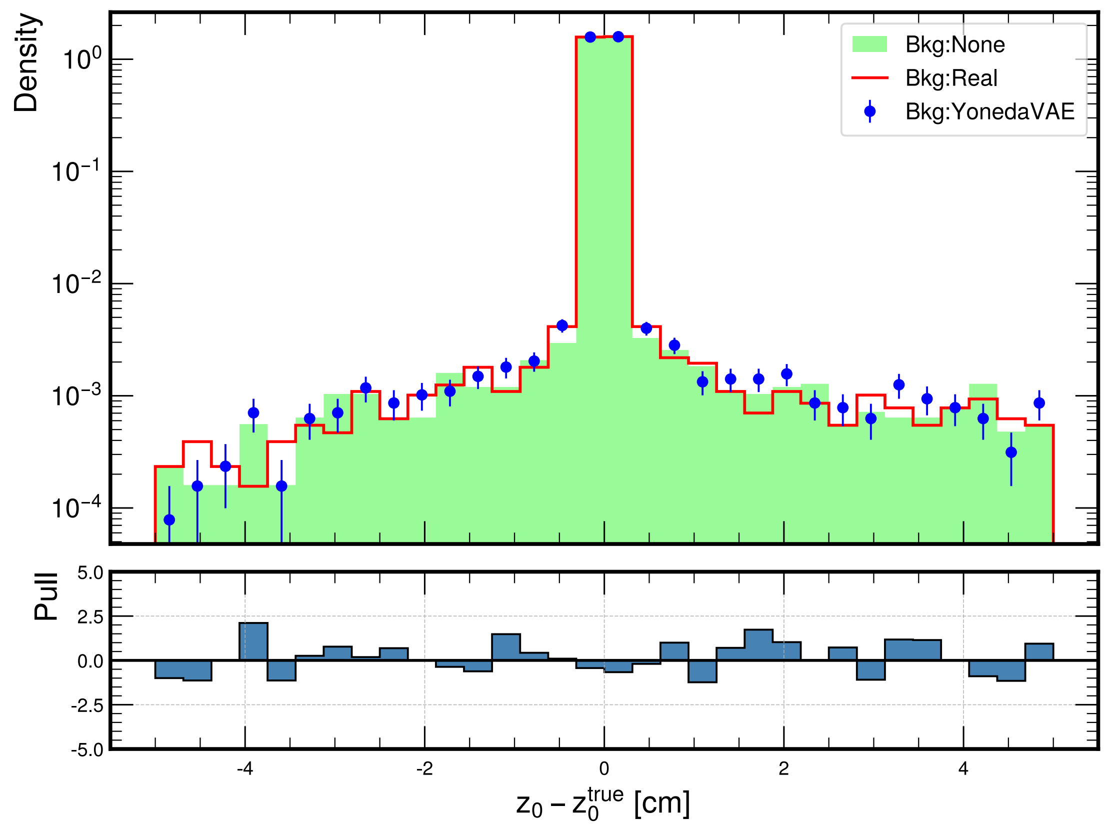
Correlation between the number of hits between 19 sensors of PXD



YonedaVAE: Results

Test (OOD data), Lum. 2.68×10^{34} cm^{-2}s^{-1}





$\sigma_{\text{YonedaVAE}}(\Delta d_0) = 0.1674 \pm 0.0011$

$\sigma_{\text{Real}}(\Delta d_0) = 0.1643 \pm 0.0012$

Real vs YonedaVAE:

KS Statistic Δd_0 : 0.0058, p-value: 0.4910

$\sigma_{\text{YonedaVAE}}(\Delta z_0) = 5.9990 \pm 0.0409$

$\sigma_{\text{Real}}(\Delta z_0) = 5.8735 \pm 0.0399$

Real vs YonedaVAE:

KS Statistic Δz_0 : 0.0040, p-value: 0.8993

Summary and Outlook

From HEP side:

- ☑ Simulate a **full detector**, with **irregular geometry** and **ultra-high granularity**
- ☑ Generate detector signatures for luminosity regions well beyond the training data

From ML side:

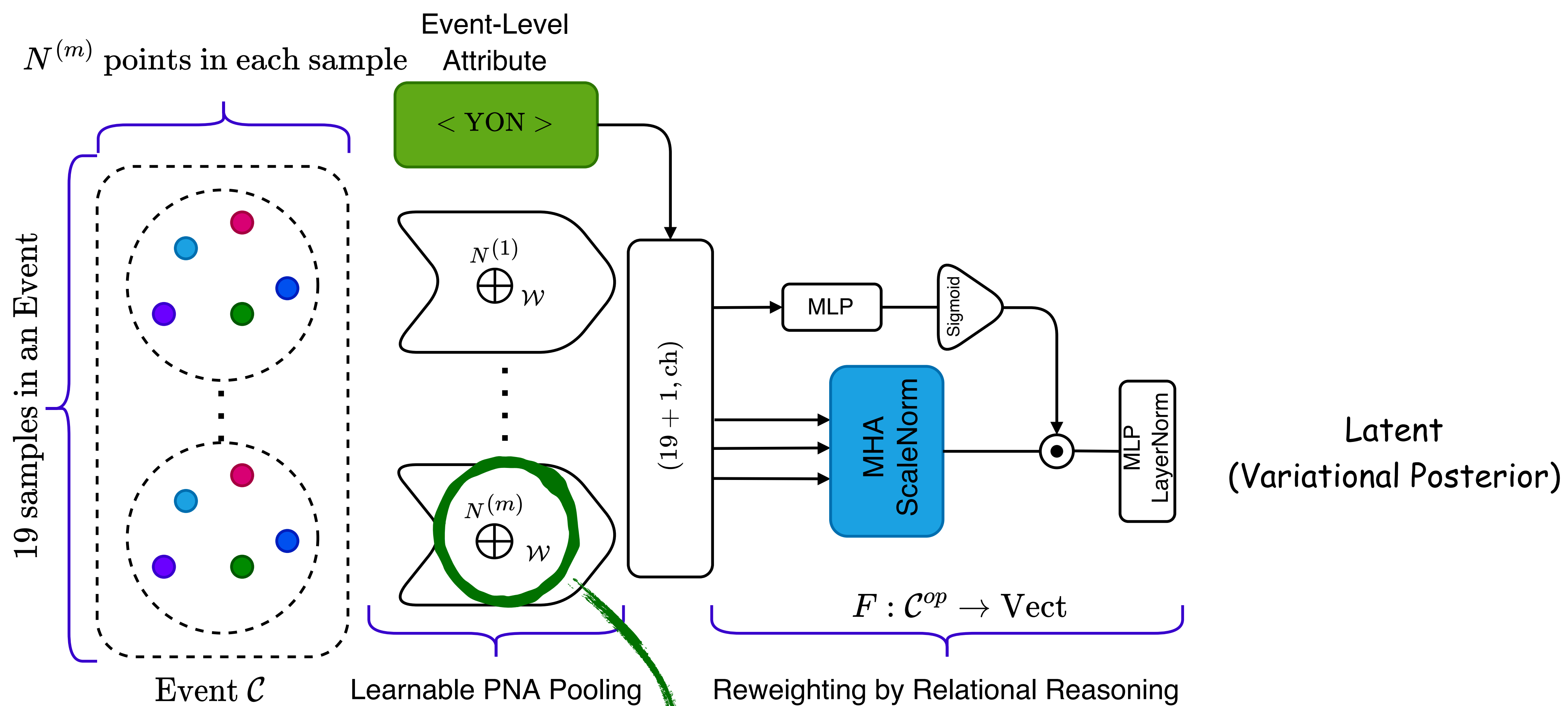
- ☑ To Reach "**Context Extrapolation**" in Inverse design problems?
- ☑ Introduce a learnable VAE prior with **Self-Supervised Learning and Transformer**?

- ☐ What needs to be done: **More in depth uncertainty quantification.**
- ☐ Stay Tuned for the full results!

Backup Slides

YonedaVAE: Encoder

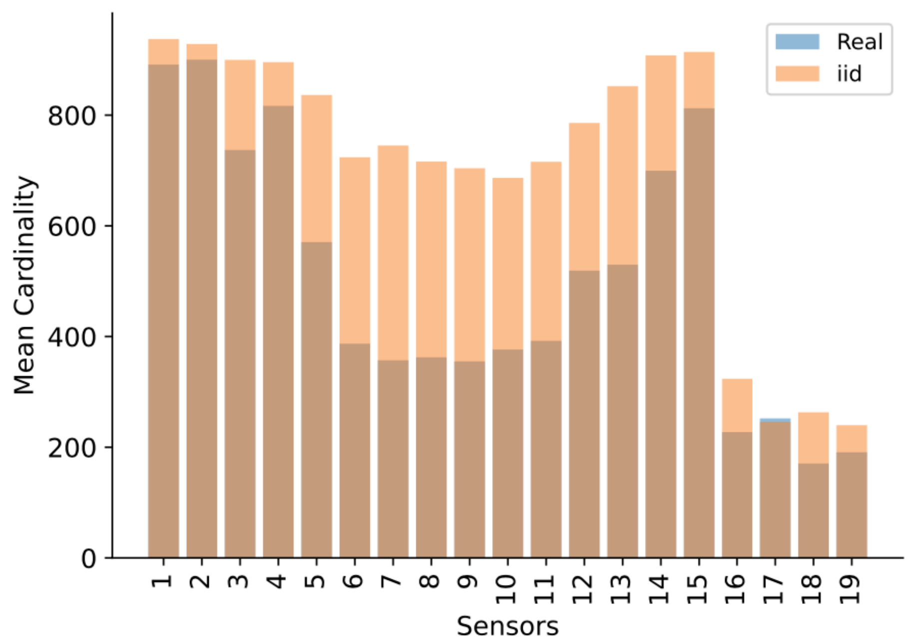
1. **Encoder:** EventFormer —> Yoneda Pooling

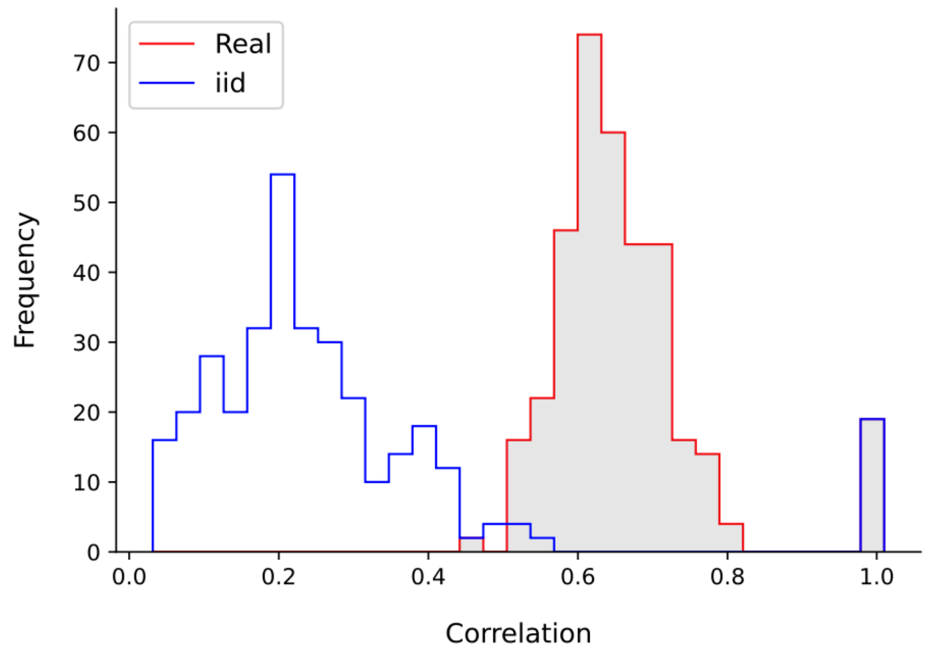


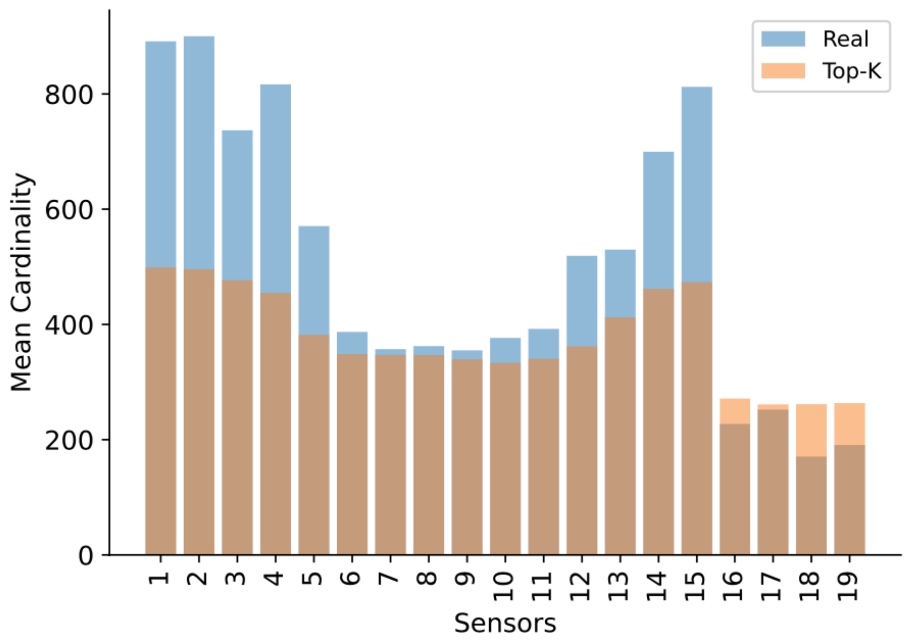
$$\bigoplus_{\mathscr{W}}^{N^{(m)}} = \underbrace{\begin{bmatrix} I \\ S(D,\alpha=1) \\ S(D,\alpha=-1)\end{bmatrix}}_{\text{scalers}} \otimes \underbrace{\begin{bmatrix} w_1\\ w_2\\ w_3\\ w_4\end{bmatrix}}_{\text{Learnable Weights}} \odot \underbrace{\begin{bmatrix}\mu\\ \sigma\\ \max\\ \min\end{bmatrix}}_{\text{aggregators}}$$

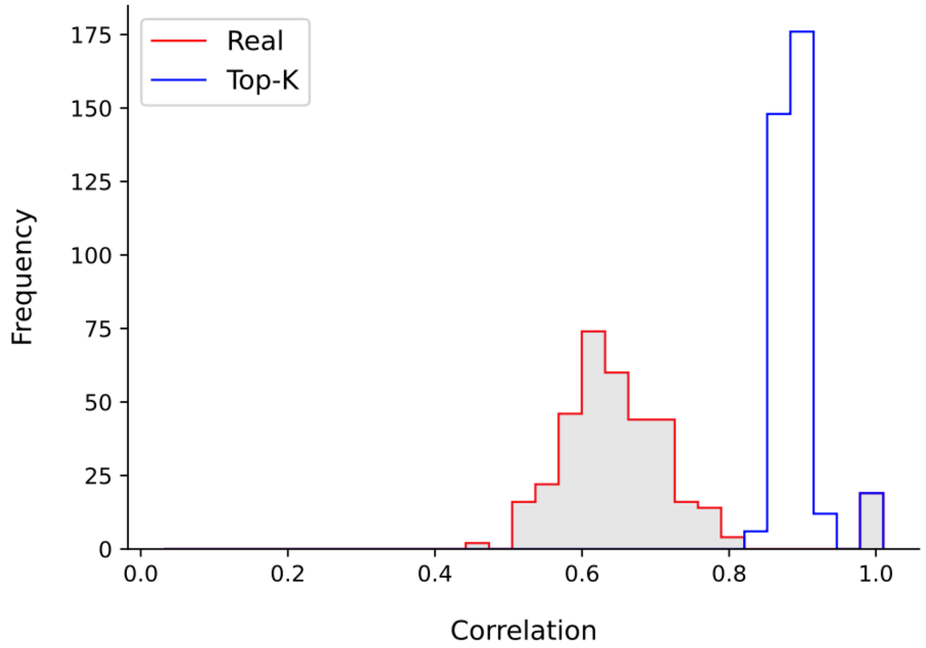
YonedaVAE: Ablation

1.Marginal Distributions:



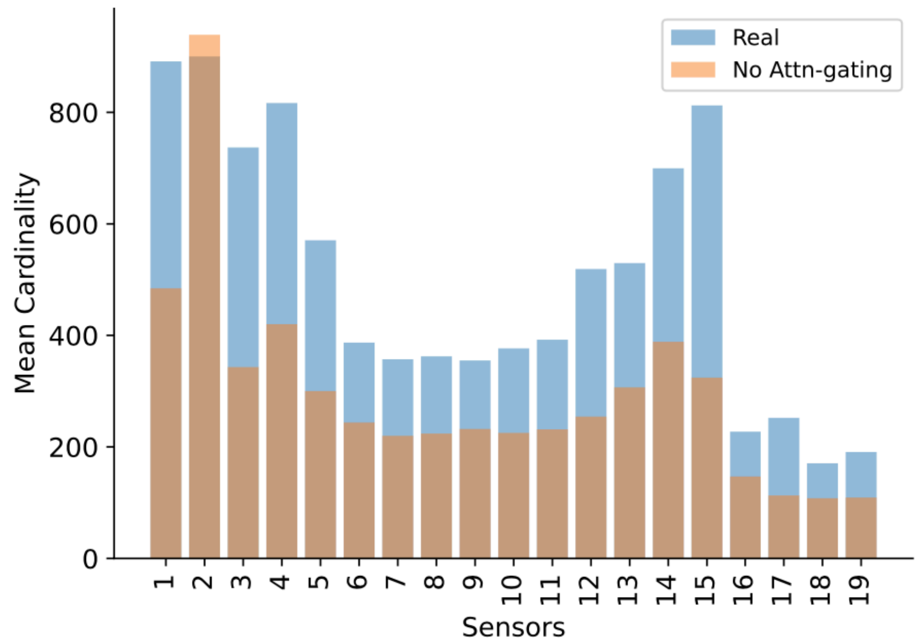


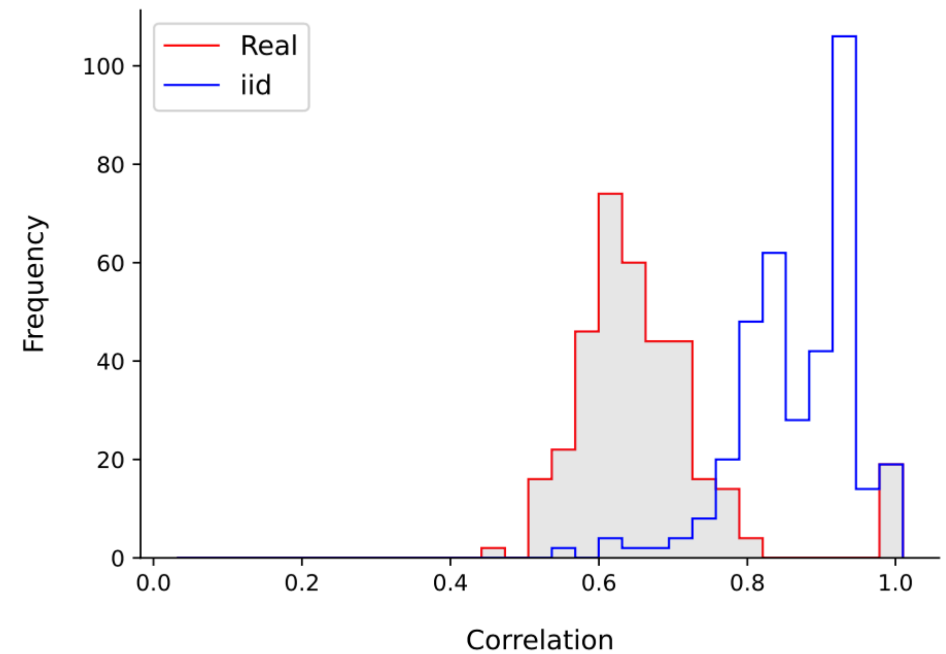


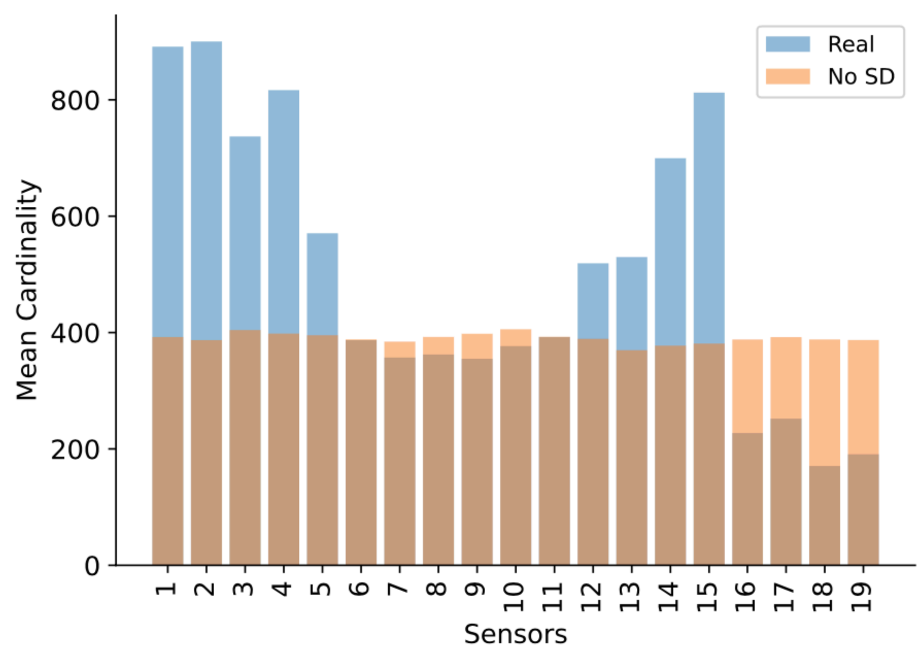


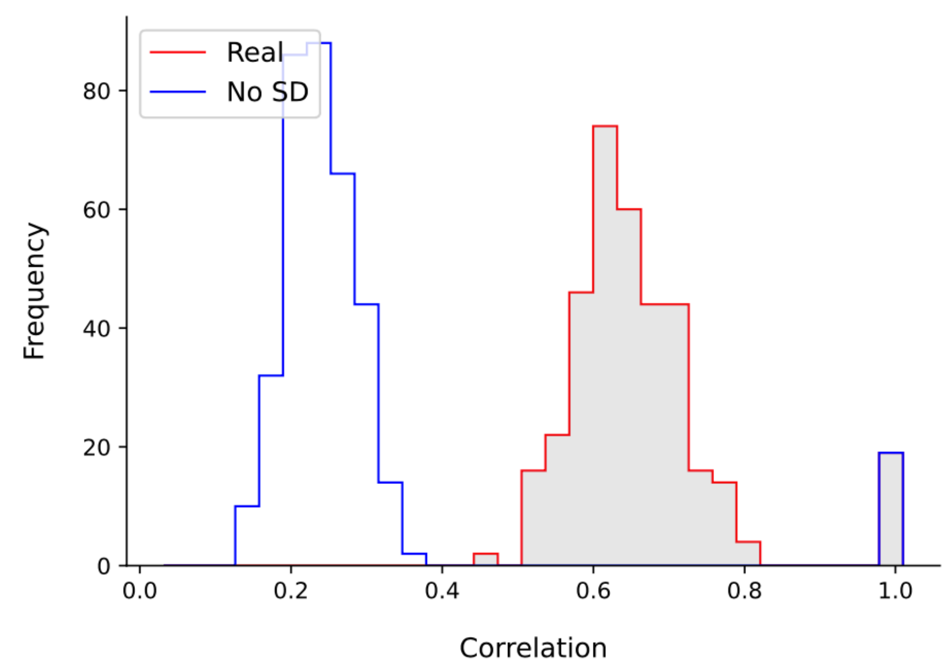
YonedaVAE: Ablation

1.Marginal Distributions:



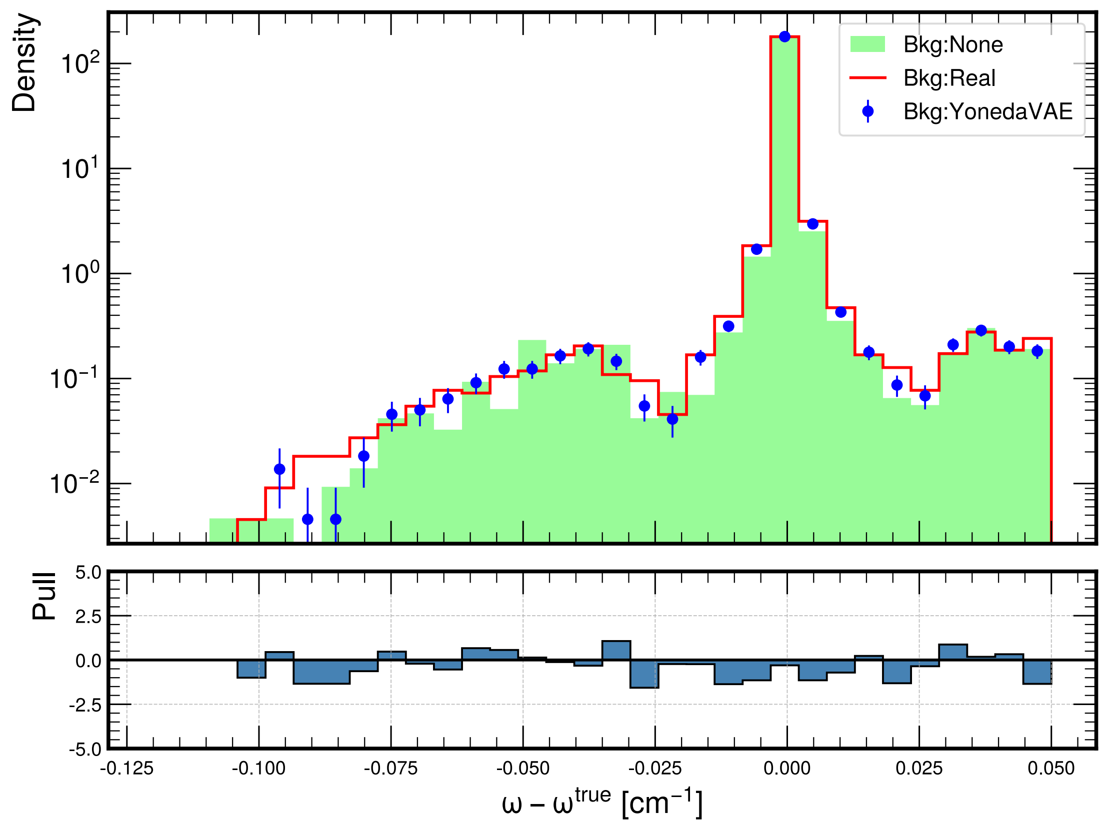






YonedaVAE: Results

Test (OOD data), Lum. 2.68×10^{34} cm^{-2}s^{-1}

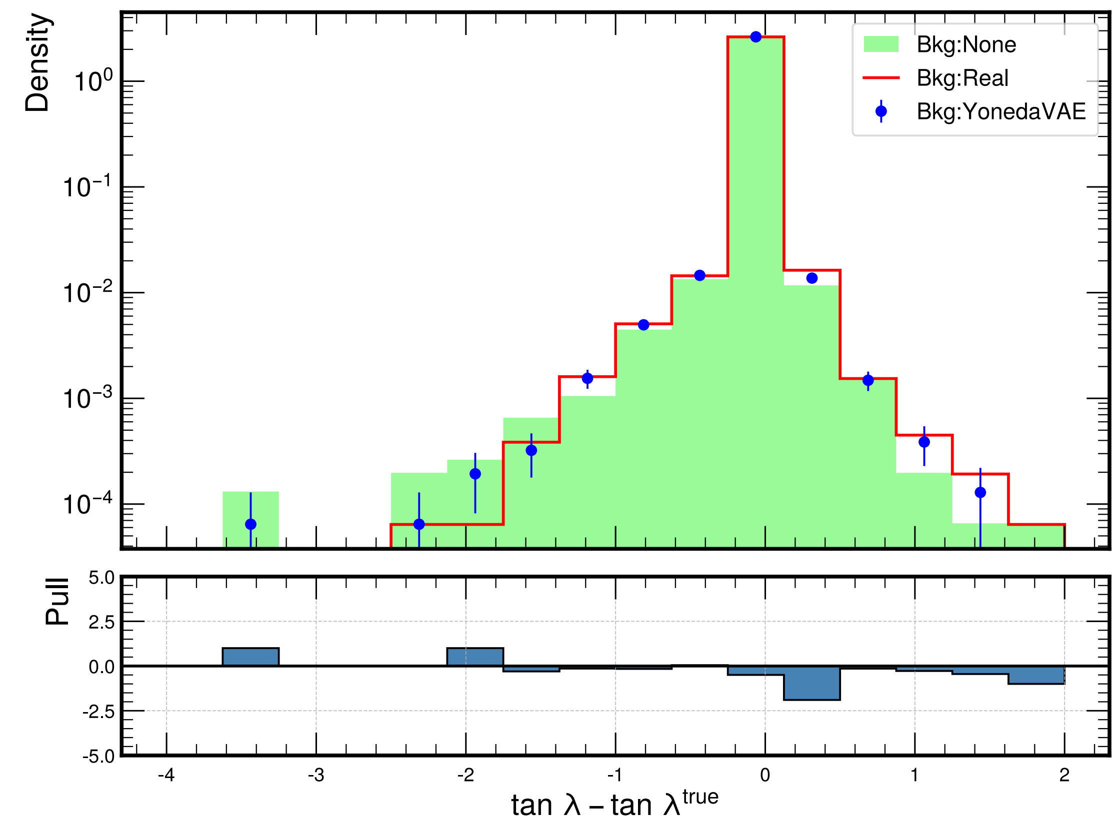


$\sigma_{\text{YonedaVAE}}(\Delta\omega) = 0.0066 \pm 0.0001$

$\sigma_{\text{Real}}(\Delta\omega) = 0.0065 \pm 0.0001$

Real vs YonedaVAE:

KS Statistic $\Delta\omega$: 0.0052, p-value: 0.6296



$\sigma_{\text{YonedaVAE}}(\Delta \tan\lambda) = 0.0753 \pm 0.0005$

$\sigma_{\text{Real}}(\Delta \tan\lambda) = 0.0726 \pm 0.0004$

Real vs YonedaVAE:

KS Statistic $\Delta \tan\lambda$: 0.0044, p-value: 0.8238

Validation of generated PXD images

- **Validation Metrics over the test set:**
  - **FID and KID:**
    - The use of activations of the last layer from the **Inception-V3** model trained on the PXD images to summarise each image, gives the score. **The lower the FID/KID the better the image diversity and Fidelity.**

	WGAN-gp	BigGAN-deep	ContraGAN	PE-GAN	IEA-GAN
FID	12.09	4.40 ± 0.88	3.14 ± 0.74	2.61 ± 0.91	$\mathbf{1.50 \pm 0.16}$
KID	0.0096	0.0031 ± 0.0001	0.0015 ± 0.0002	0.0021 ± 0.0004	$\mathbf{0.0010 \pm 0.0002}$

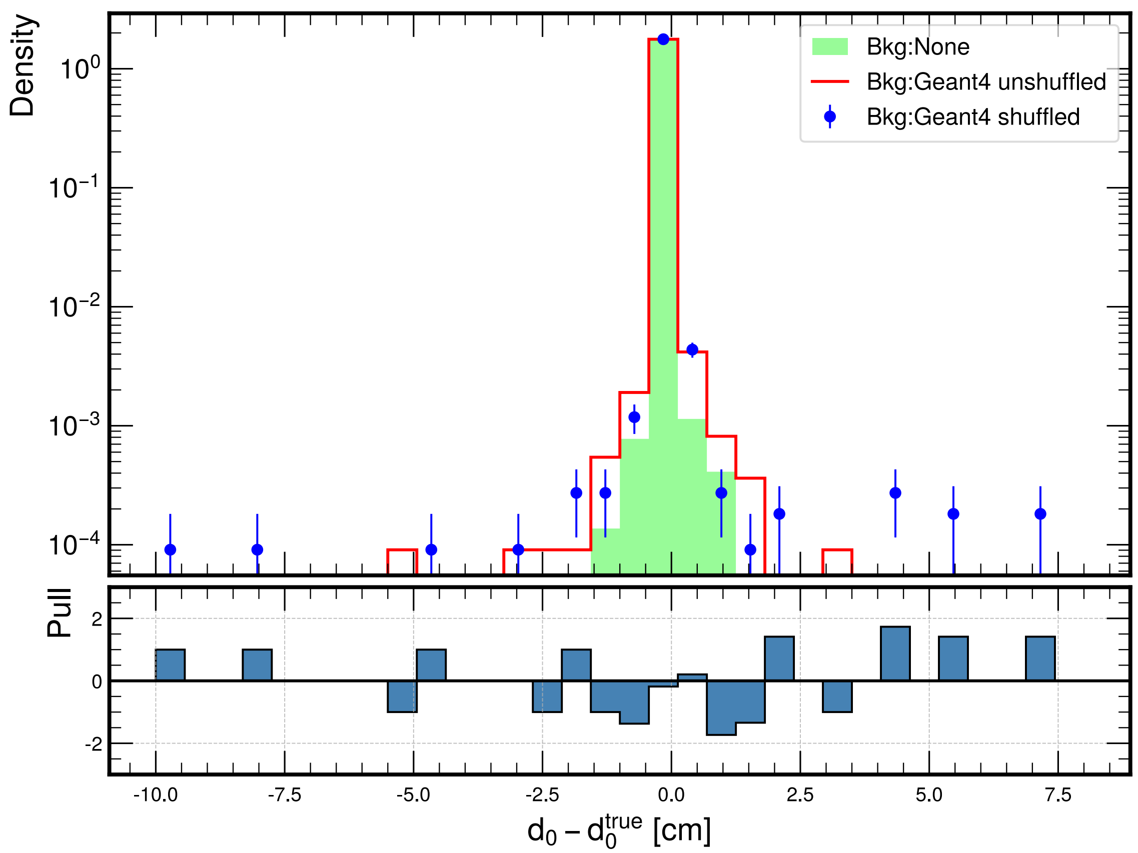
    - Possible interpretation of FID at the pixel level:

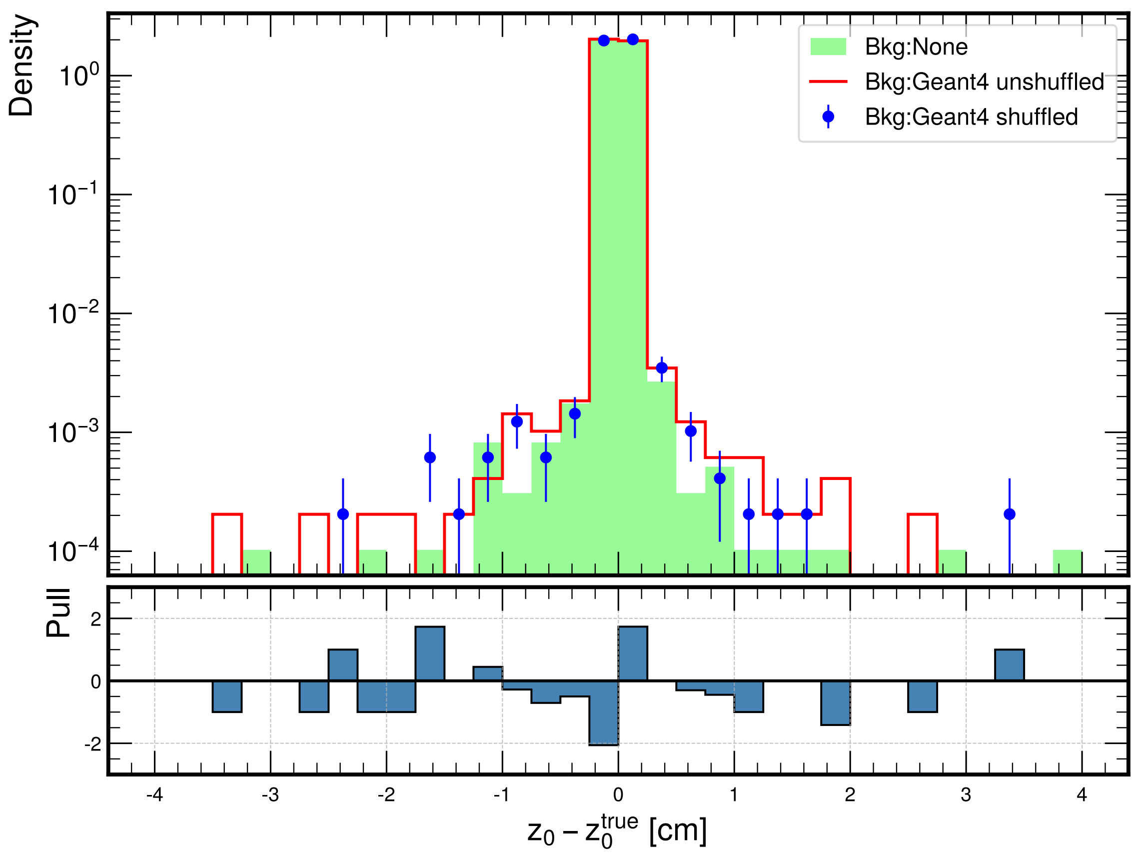
Image Jitterings	**FID**
None	0
Random Masking (dead zones)	14.58
Random Noise	87.23
Random Rotation (30 degrees)	23.69
Random Rotation (10 degrees)	2.81
Random Translation (0.1, 0.1)	1.99
Random Shear (10, 10)	23.53
Random Zoom	9.06
High Intensity smearing	3.16
Low Intensity smearing	47.24

Correlation vs No Correlation

How important are these correlations?

- For high momentum regime $p_T > 0.4\ GeV$
- **Shuffling the events —> losing the correlation**





Parameter	Unbiased Resolution ± error		KS statistic	p-value
	Shuffled Geant4	Unshuffled Geant4		
d_0	0.1343 ± 0.0007	0.0732 ± 0.0004	0.0067	0.7655
ϕ_0	0.2158 ± 0.0011	0.1859 ± 0.0009	0.0066	0.7899
z_0	5.0076 ± 0.0253	4.9341 ± 0.0249	0.0152	0.0211
ω	0.0010 ± 0.0000	0.0008 ± 0.0000	0.0138	0.0485
$\tan\lambda$	0.0388 ± 0.0002	0.0382 ± 0.0002	0.0167	0.0086

Overlay Problem

- **Realistic detector simulation has to take into account effects from background processes**
  - Simulation requires many PXD hitmaps with statistically independent background.
  - Overlay hits from simulated background or random trigger data to hits from signal MC.
  - PXD hits have the highest storage consumption.
  - Requires distributing over all sites where MC is produced.
- **Solution:** Generating the background data on the way of analysis with GANs instead of storing them.