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

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Baran Hashemi
Origins Cluster Munich, TUM

Are you looking for …

HEP Side:

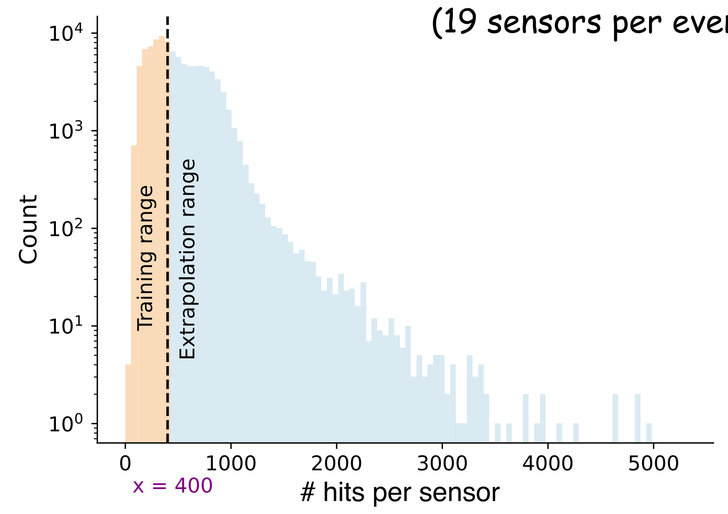
- ☑ *Simulating a **full detector** with **irregular** and **complex geometry**?*
- ☑ *Generating detector signatures for kinematic/luminosity regions where **data is very rare**?*

ML Side:

- ☑ *Reaching "**Context Extrapolation**" in Inverse design problems?*
- ☑ *Enhancing your VAE prior with **Self-Supervised Learning**?*

Setup

☑ PXD tracking detector → *Up to **110,000** hits per event* (19 sensors per event)



Train:

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

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

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

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

$\mathbf{N}^{(m)}$: Variable for each sample

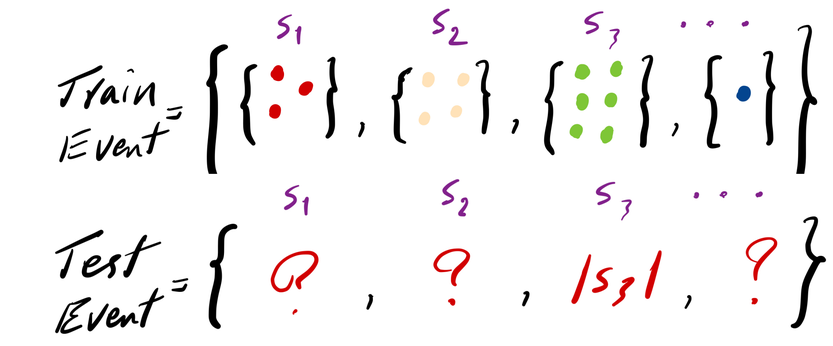
Test:

Zero-Shot

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

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

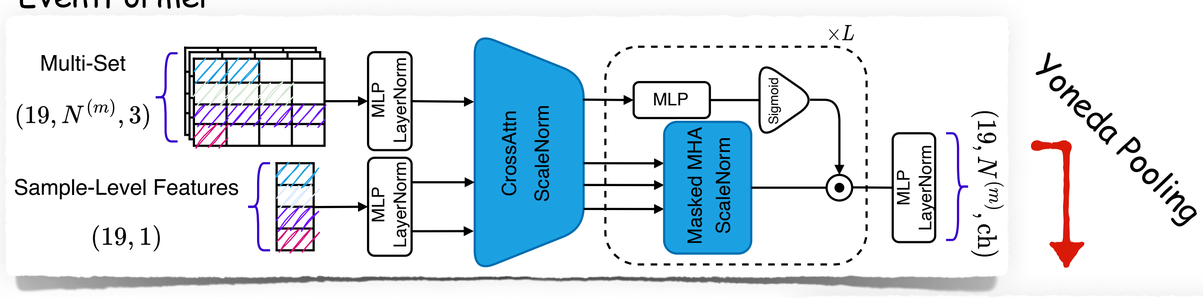
s.t $N_{max}^{test} \gg N_{max}^{train}$



YonedaVAE: Encoder

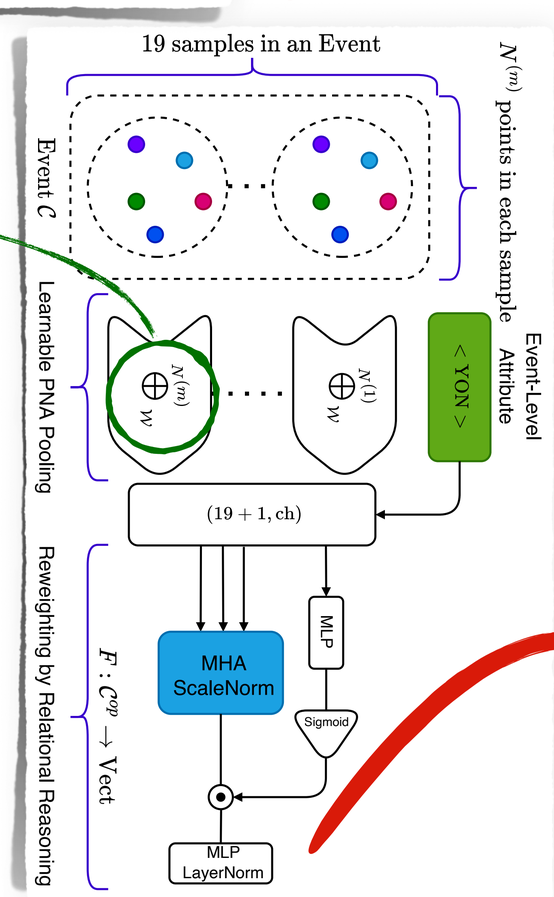
1. Encoder = *EventFormer + Yoneda Pooling*

EventFormer



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

Learnable PNA pooling operation



⇒ Going from "**syntax**" to "**semantics**"

Syntax: "which sensor is near which sensor,"

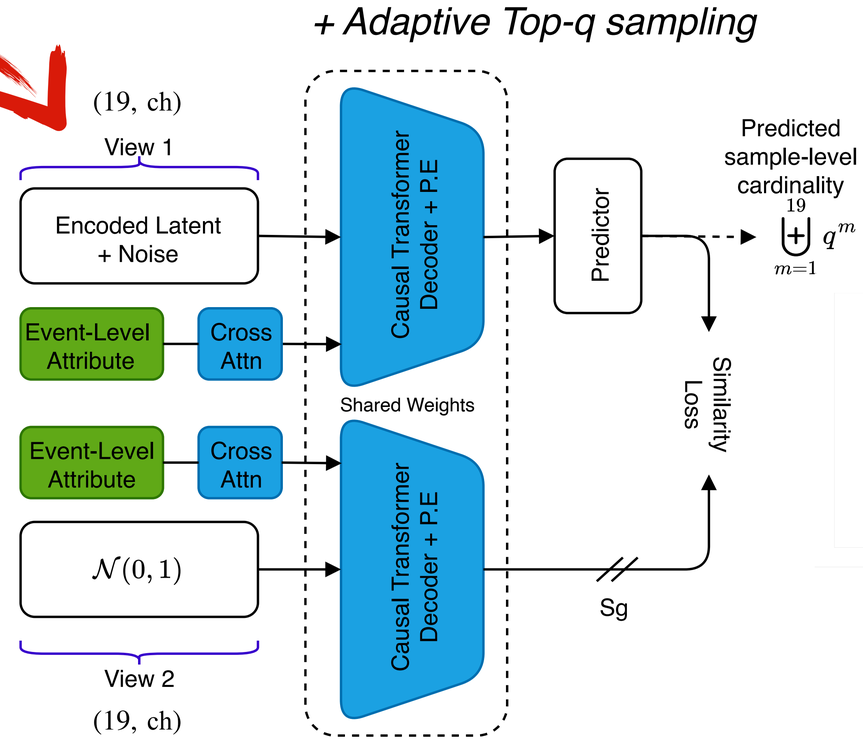
Semantics: "what sensors mean in an event."

"*Tell me how you relate to everything else, and I will tell you who you are*"

Yoneda lemma

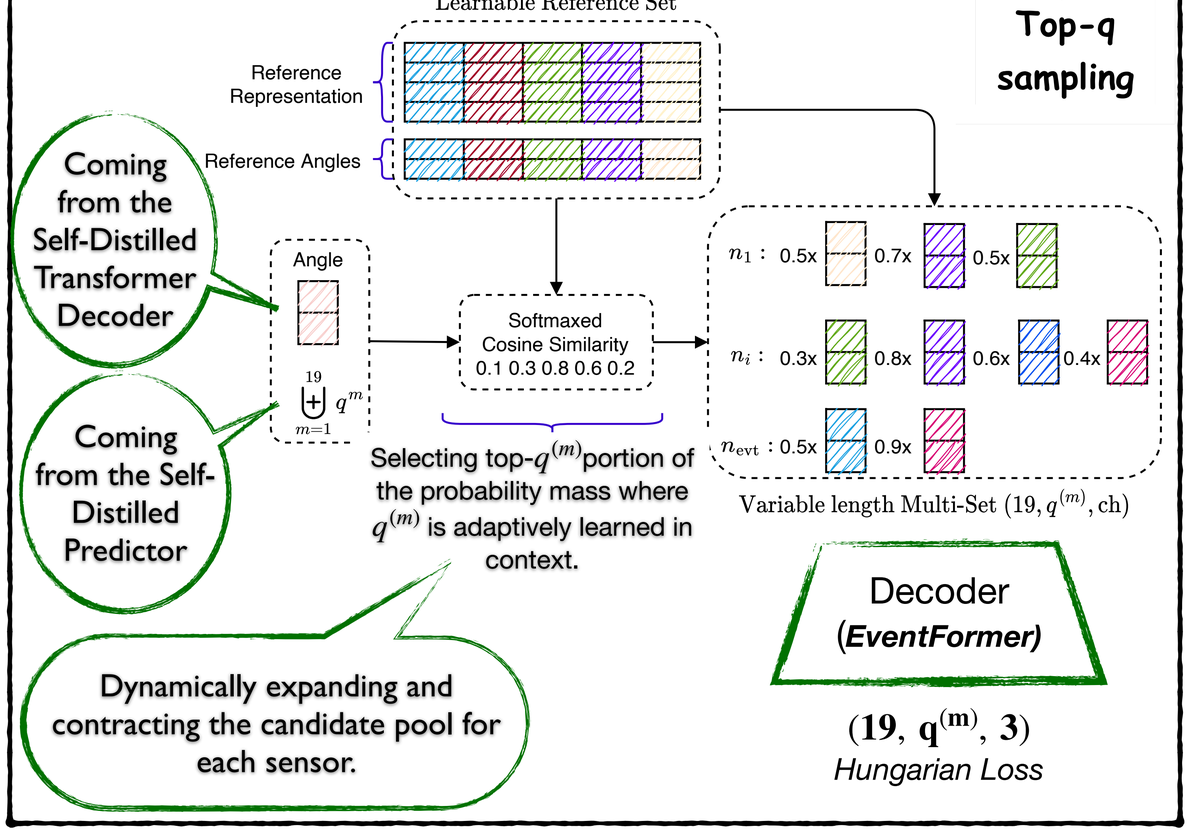
YonedaVAE: Set Generator & Decoder

2. Set Generator = *Self-Supervised Causal Transformer + Adaptive Top-q sampling*



Learning VAE prior à la Self-Supervised Learning

Adaptive Top-q sampling

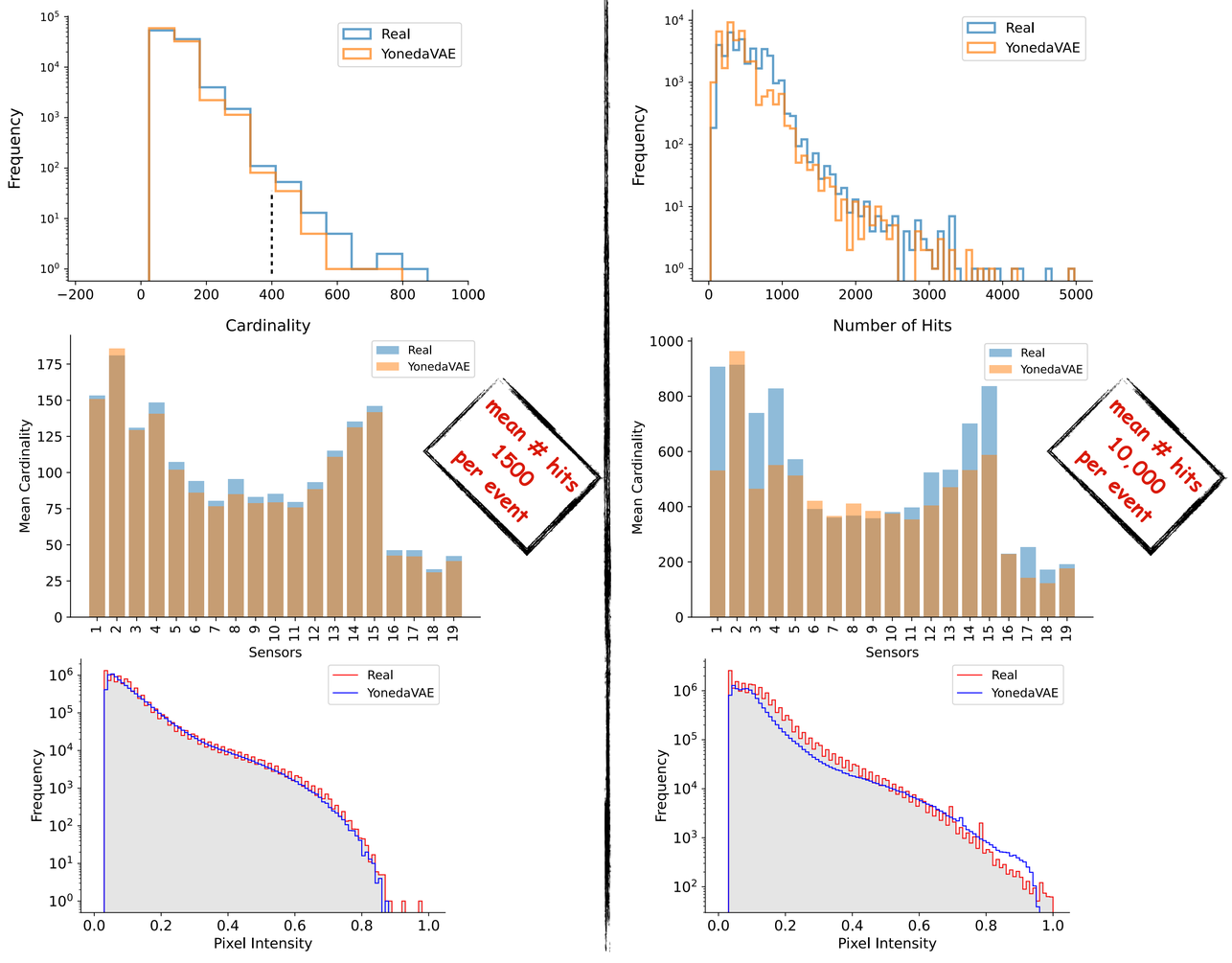


Evaluation (More plots on Friday ;D)

	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	$\mathbf{130} \pm \mathbf{2}$	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.

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



Belle II · TUM LMU ORIGINS Excellence Cluster · Bundesministerium für Bildung und Forschung · ErUM-Data IDT · Belle II