# LATENT SPACES IN GENERATIVE MODELS

**Nesar Ramachandra**
**nramachandra@anl.gov**
**CPS Division, Argonne National Laboratory**

U.S. DEPARTMENT OF ENERGY | Office of Science

Argonne NATIONAL LABORATORY

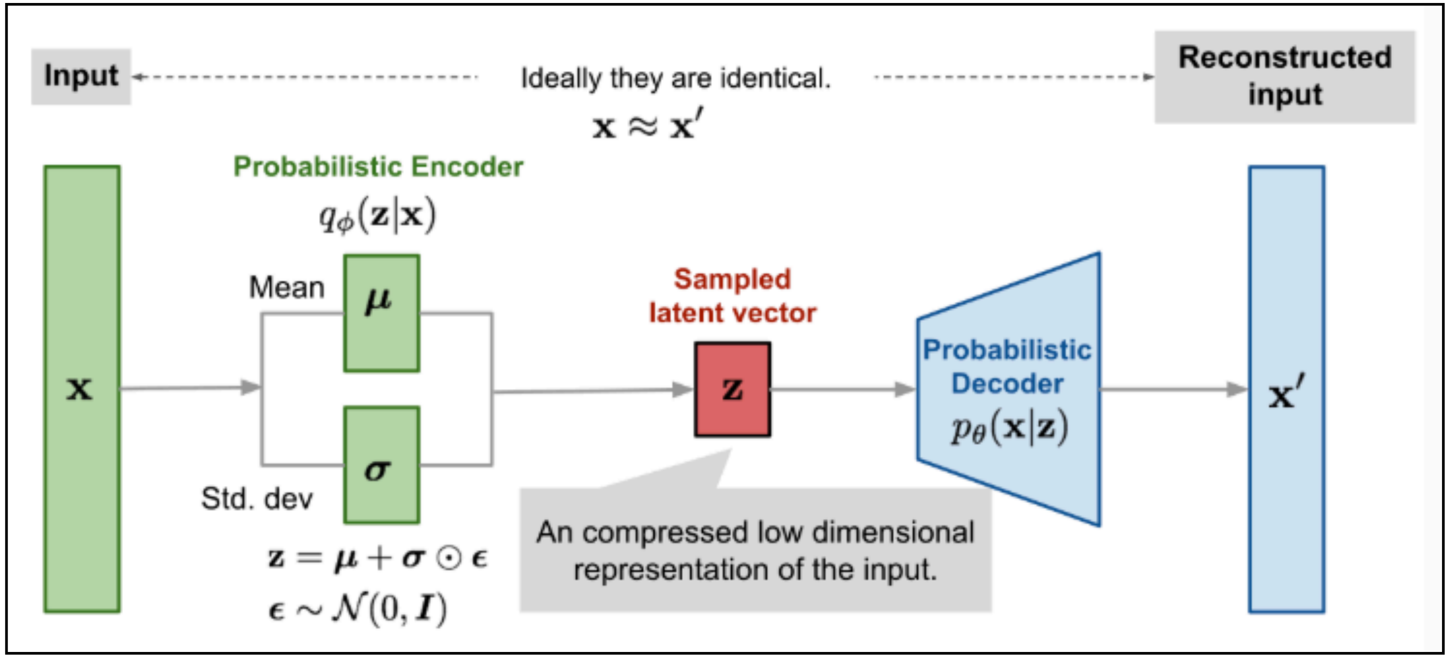
# INTRODUCTION

This talk highlights a specific way of adapting and leveraging AI algorithms for scientific studies.

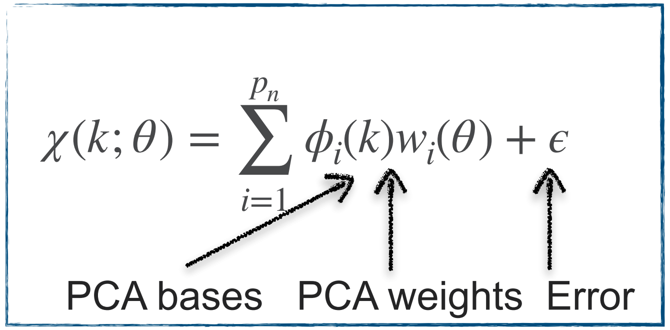
- What is a latent/feature/representation space? Why is it important?
- What scientific advancements are possible with representation learning?
- What are the issues with latent spaces resulting from completely un-supervised learning?
- How can one improve upon existing AI models to unravel unknown physics models?
  - Disentanglement and auxiliary information
  - Applications — 2 physics problems

# LATENT/FEATURE/REPRESENTATION SPACE

- Latent space is a compressed/encoded representation of an original higher dimensional dataset.
- Often used these days generative AI algorithms, but a wide variety of statistical/physical models work in lower dimensional representation space.

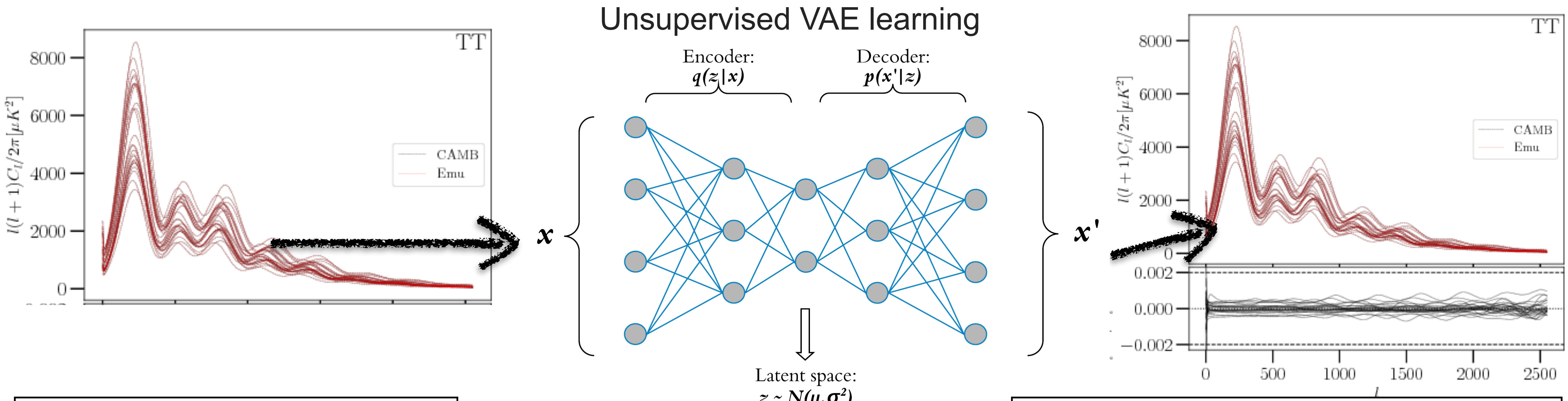


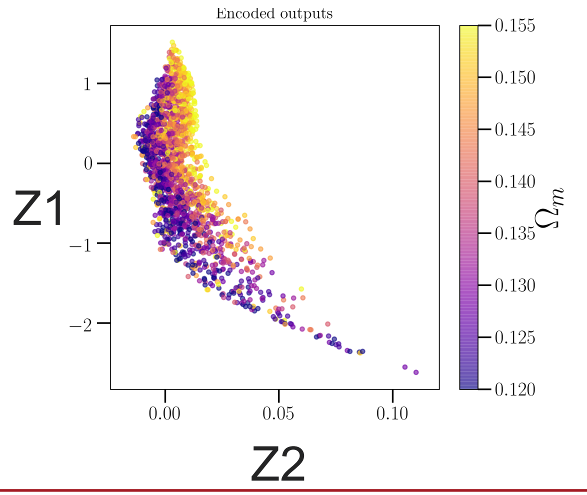

Variational Auto-encoder with a bottleneck layer




Truncated Principal Component Analysis

# IMPORTANCE OF LATENT SPACES



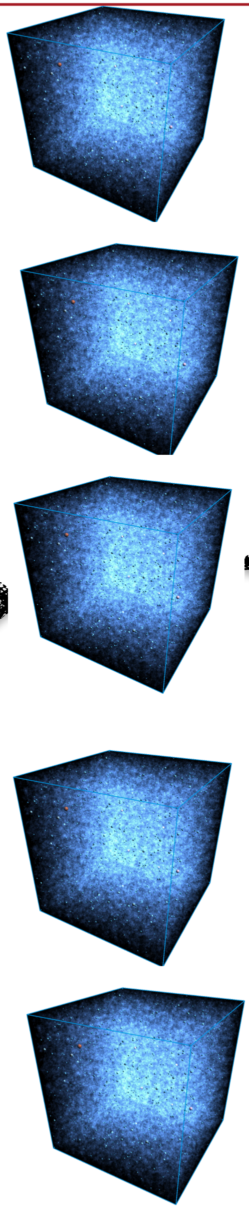

- **Feature extraction:** Correlations that are not apparent in original data space may be clearer in the reduced space.

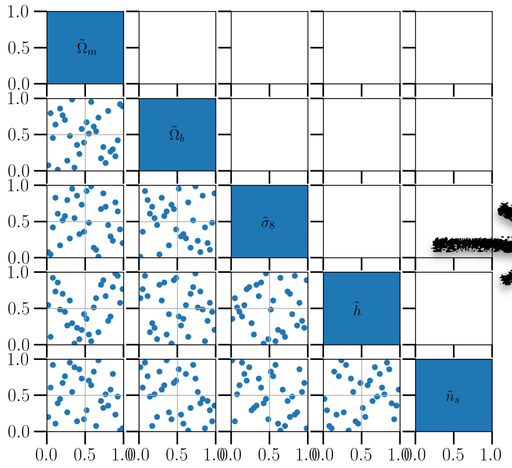
- **Compression:** Low dimensional representation can aid in smoother interpolation, easier generation, model performance.

# LATENT SPACE WALKS — APPLICATION 1

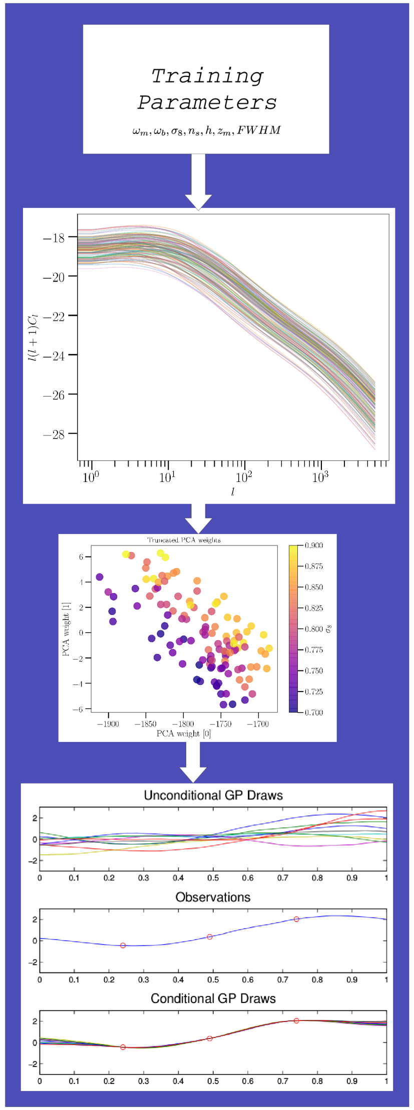
Argonne NATIONAL LABORATORY

Run training simulations, generate summary statistics

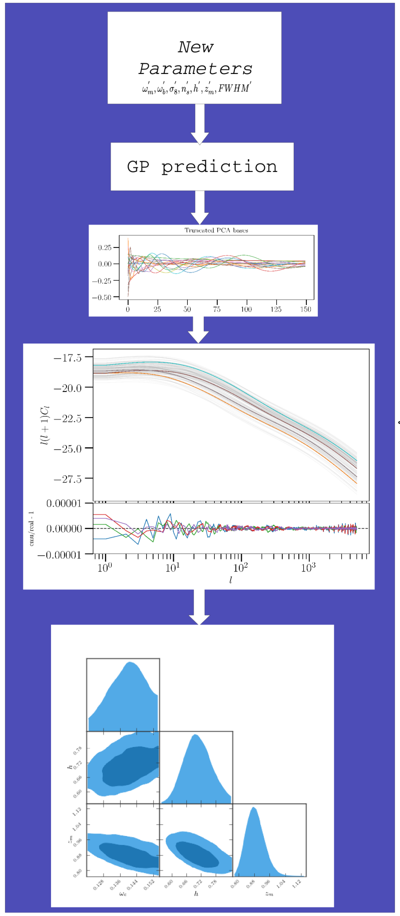




Experimental design: space filling latin hypercube




PCA/VAE reduction, GP fitting




Latent-walk based surrogate model for faster Likelihood calculation

$$\mathscr{L}(D\,|\,\theta) \propto \exp\left[-\frac{1}{2}\sum_{i,j}\left(D - f(\theta)\right)_i C_{ij}^{-1}\left(D - f(\theta)\right)_j\right]$$

Posterior estimation using emulated Likelihood via MCMC

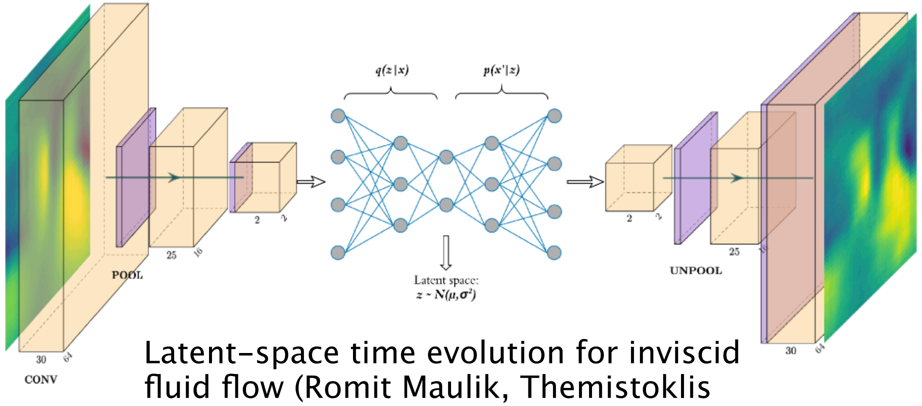
$$P(\theta\,|\,D) \propto \mathscr{L}(D\,|\,\theta)P(\theta)$$

**NR**, Georgios Valogiannis et al (arxiv:2010.00596)

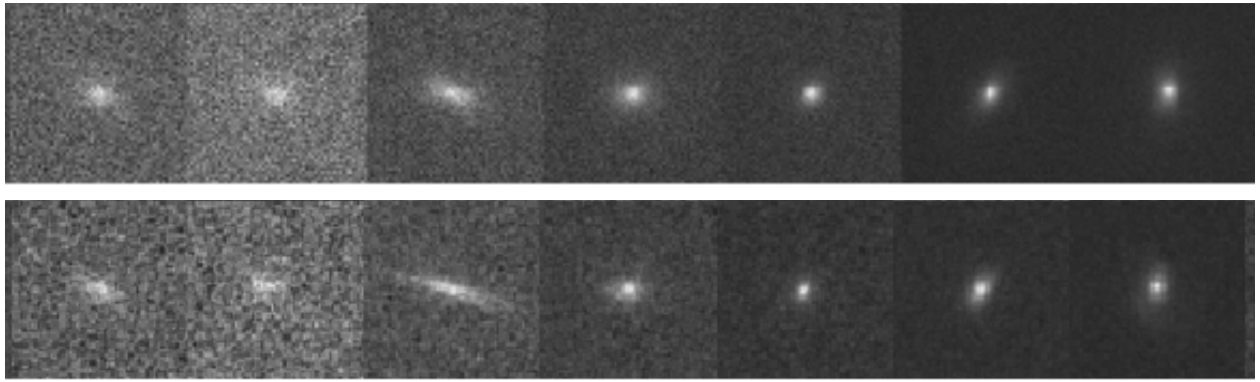
U.S. DEPARTMENT OF ENERGY | Office of Science

# LATENT SPACE WALKS — APPLICATION 2

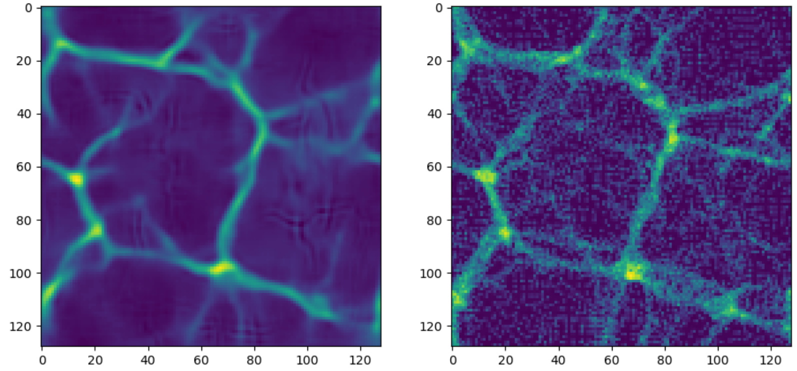
Deep-learning based compressions can be applied to variety of datasets — images, time-series, n-D simulations, graphs, texts




Latent-space time evolution for inviscid fluid flow (Romit Maulik, Themistoklis Botsas, **NR** et al: arxiv:2007.12167)



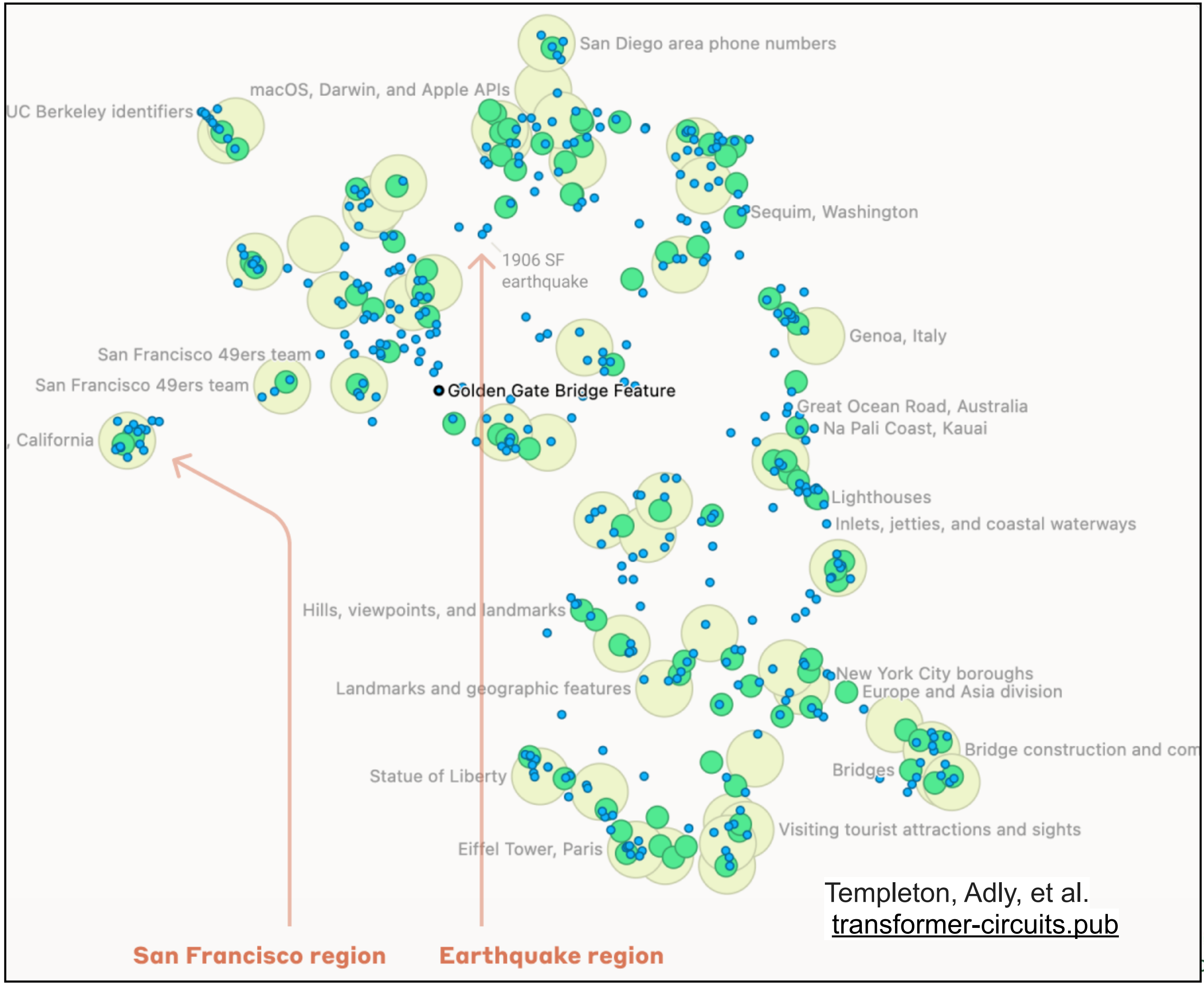
Galaxy image emulation (Claire Guilloteau, **NR** et al)




3D cosmic density field reconstruction (Xiaofeng Dong, **NR** et al, 2021 arxiv:2111.12118)

# LATENT SPACE WALKS — APPLICATION 3

Foundation models and LLMs deal with enormous quantity of data, and looking into latent spaces can offer better insights.

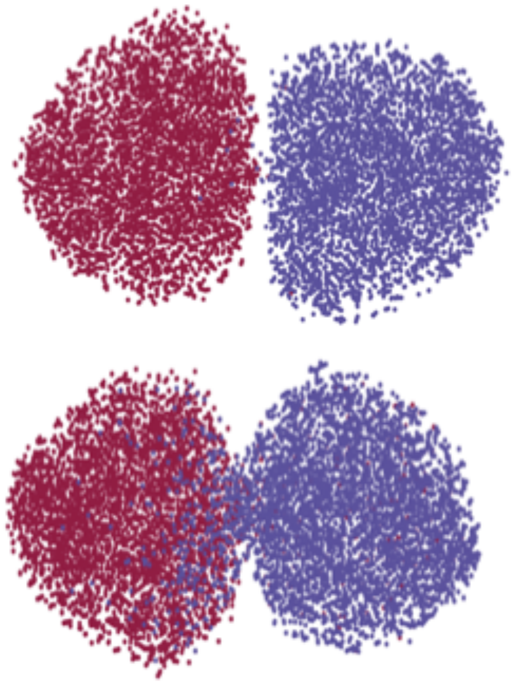



Templeton, Adly, et al.
transformer-circuits.pub

U.S. DEPARTMENT OF ENERGY | Office of Science

# LATENT SPACE WALKS — APPLICATION 4

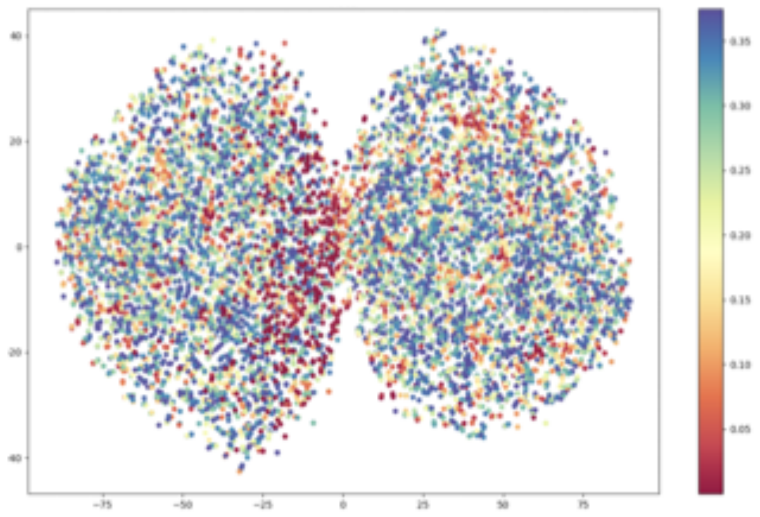
Latent spaces combined with robust UQ can help isolate misidentifications.



Top: perfect classification - clear boundary in tSNE projection
Bottom: misidentified lenses at the junctions.

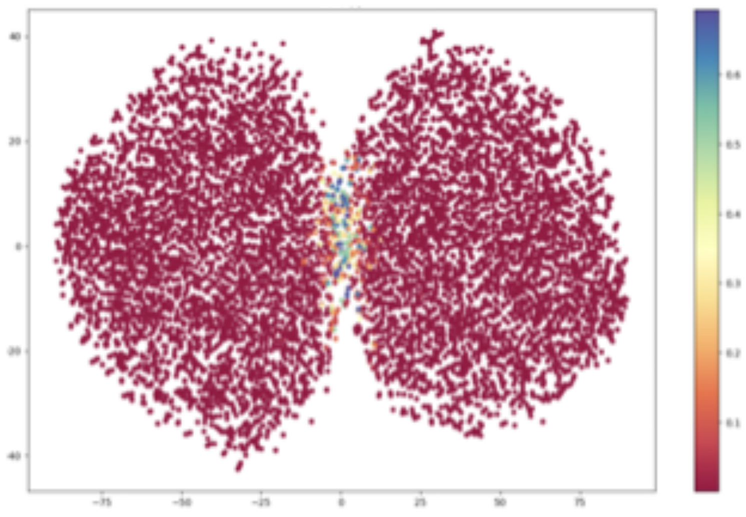
Measure of Uncertainty:

$$p(y, z|x) = \int dx\, p(x) e_\theta(z|x) q_\psi(y|z)$$



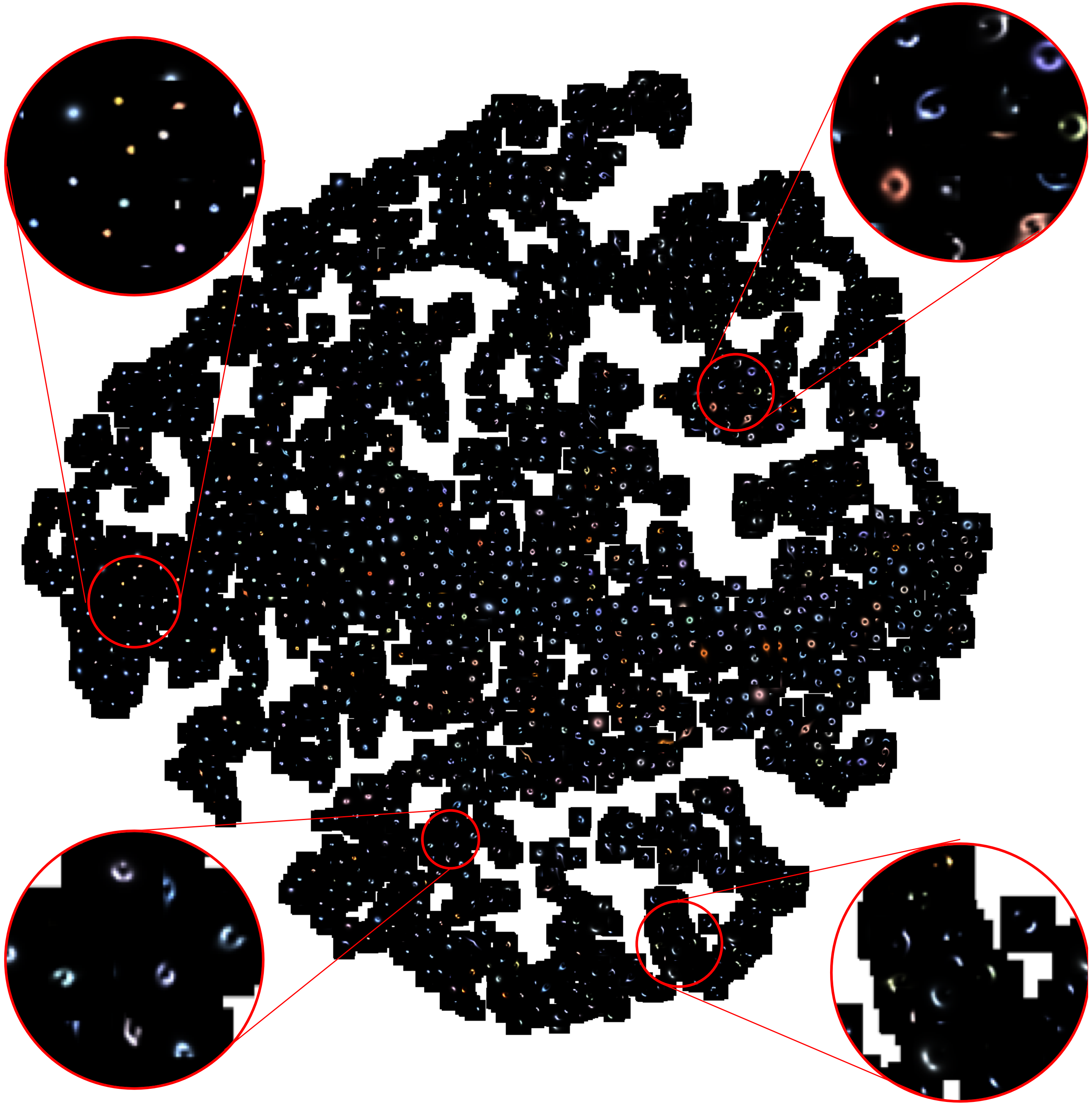
Entropy:

$$H(y|x) = -\sum_i p(y_i|x) \log p(y_i|x)$$



Measure of uncertainty and entropy both reveal lower confidence in mis-classified lenses

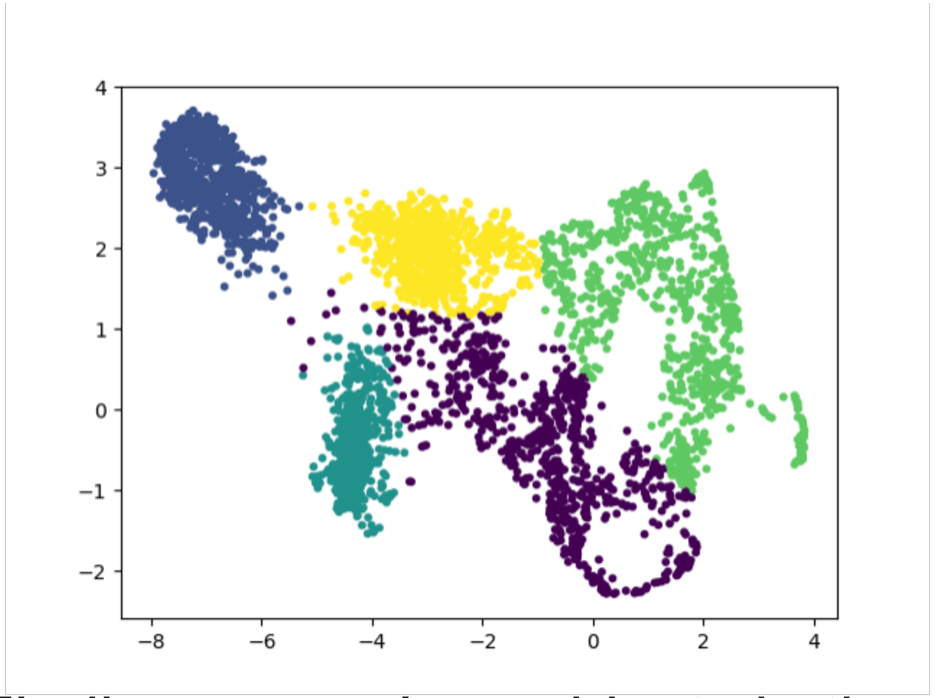
Sandeep Madireddy, **NR** et. al.: arxiv:1911.03867



Variational Information Bottleneck and representation learning shows arrangement of strong lenses based on geometrical features

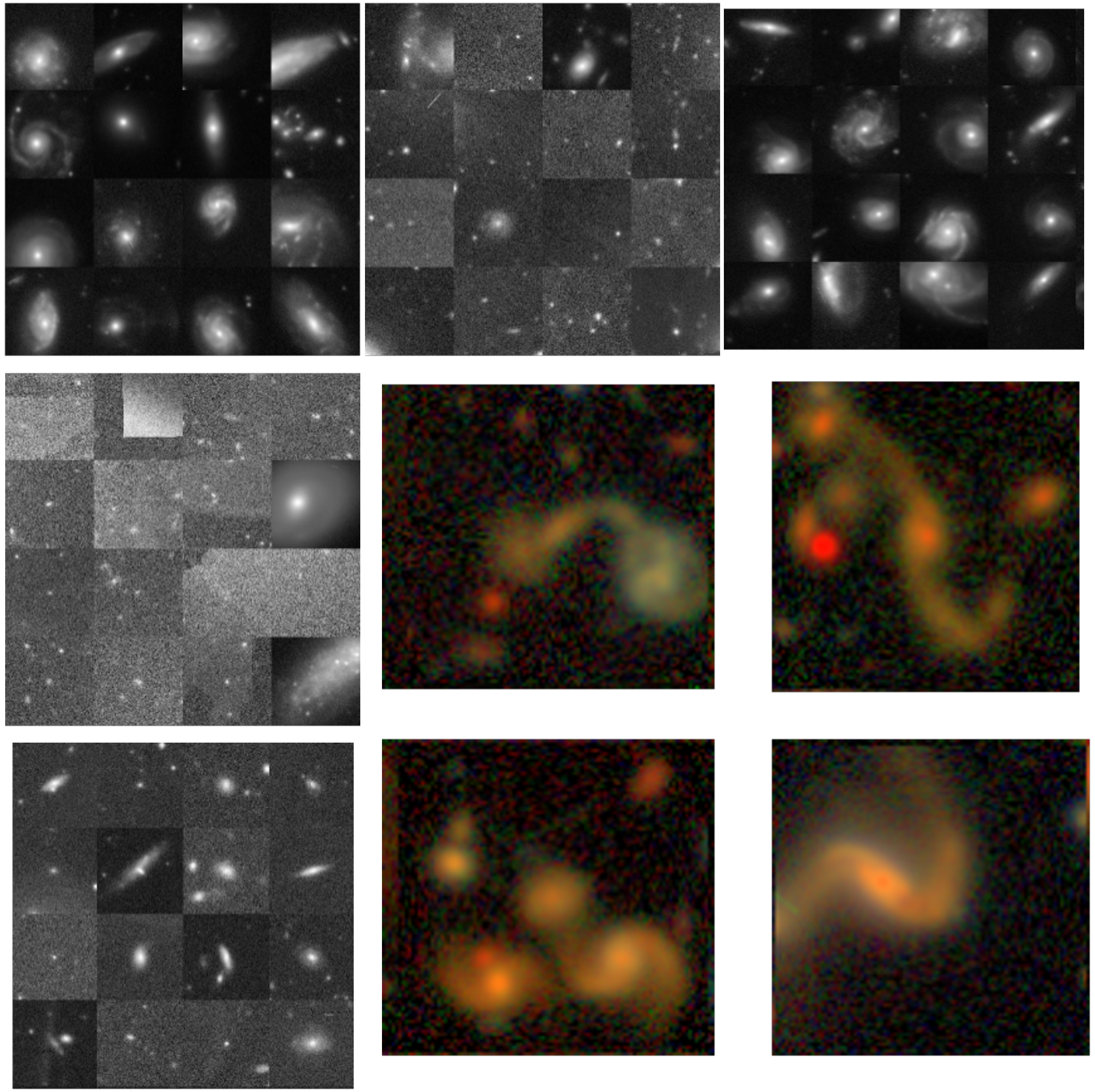
Sandeep Madireddy, **NR** et. al.: arxiv:1911.03867

# LATENT SPACE WALKS — APPLICATION 5




Finding anomalous objects in the dataset is easier in latent space. Classification of ‘unknown objects’ is possible as well.

Kate Storey-Fisher, Marc Huertas-Company, **NR** et. al.
https://arxiv.org/abs/2105.02434



# WHAT IS MISSING IN LATENT SPACE MODELING

1. **Traditional representation learning is done in unsupervised spaces**
   - Highly useful in certain datasets (natural images, texts)
   - Whereas, scientific datasets are more nuanced. Physical parameters are often associated (measured quantities, simulation settings) with datapoints.
2. **Multiple modalities, and multiple fidelities are encountered often.**
   - Parameters are measured with different levels of accuracies, biases, underlying physics may be understood at different levels of confidence.
     - **Known knowns**
     - **Known unknowns**
     - **Unknown unknowns**

Wikipedia:There are unknown unknowns

3. **Complex datasets result in more entangled representations**
   - Untractable and reduced usefulness

Disentangling Latent Spaces in Generative Models for Scientific Datasets

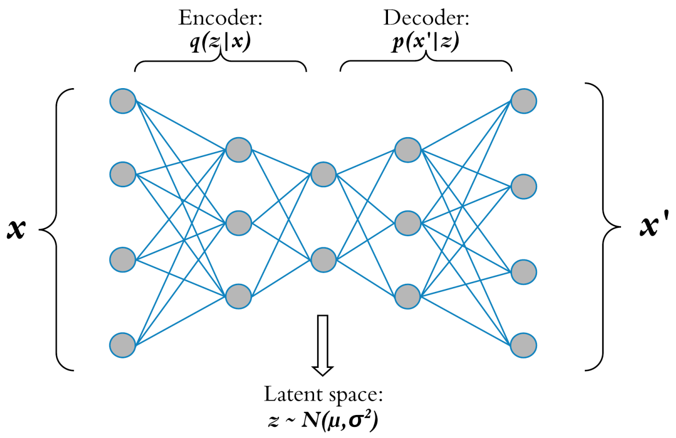
**Arkaprabha Ganguli**, Nesar Ramachandra, Julie Bessac, Emil Constantinescu

Submitted to NeurIPS-Main 2024, SUDS-2024

# AUX-VAE: GENERAL SETTING AND OBJECTIVES

1. In our database, we assume access to a **subset of ground truth factors** $S_{obs} \in \mathbb{R}^d$, represented by auxiliary variables $u \in \mathbb{R}^d$.
2. We note that, we do not assume to know the exhaustive set of true generative factors. There might be some unknown generative factors that we do not observe as the auxiliary information.
3. **Observed database** $\mathscr{D}$: $n$ independent and identically distributed pairs of data points $x$ and $u$: $\mathscr{D} = \{(x^{(1)}, u^{(1)}), (x^{(2)}, u^{(2)}), \ldots, (x^{(n)}, u^{(n)})\}$
4. **Aim**:
   - Disentangle the latent space based on identified ground truth factors $S_{obs}$ observable via $u$.
   - Ensure that each **auxiliary variable** $u_j$ **strongly associates with one specific latent factor** $z_{aux,j}$, where $Z_{aux}$ are the auxiliary-informed latent factors.
   - Capture the **remaining variability** (for the unobserved generative factors) **collectively in the remaining latent factors** $Z_{recon}$.
   - Utilize auxiliary variables to guide the learning process, enhancing interpretability and performance.

# AUX-VAE: CONDITIONING LATENT PRIORS




1. The loss function of a VAE can be expressed as:
$$\mathscr{L}_{VAE} = \mathbb{E}_{z\sim q(z|x)}\left[\log p(x\,|\,z)\right] - \mathrm{KL}\left(q(\mathbf{z}\,|\,\mathbf{x})\,||\,p(\mathbf{z})\right),$$

2. To utilize the available auxiliary information $u$, we divide the latent space as $Z = (Z_{aux}, Z_{recon})$, and enforce the following disentangled prior:
$$p_{z|u}(z) = \left(\prod_{j=1}^{d} p_{\mathcal{N}(u_j,1)}(z_j)\right) p_{\mathcal{N}(0,I_{d_Z-d})}(z_{(d+1):d_Z}) = p_{\mathcal{N}(\mu_0,I_{d_Z})}(z)$$ with
$$\mu_0 = (u_1, u_2, \ldots, u_d, 0, \ldots, 0)$$

3. The proposed disentangled prior facilitates improved interpretability and control over the latent space, enhancing the model's ability to accurately capture and separate known generative factors from the data, leading to more robust and explainable representations.

# AUX-VAE: REGULARIZATION IN LOSS FUNCTIONS

1. In addition to imposing the prior, applying posterior regularization further aids in achieving the desired disentangled structure.

2. The expected variational posterior can be defined as:

$$q_\phi(z) = \int q_\phi(z|x)p(x)dx$$

3. However, quantifying the independence in non-linear setting is non-trivial and we use polynomial regression for this purpose

   A. $$R_0^K(v,w) = \frac{1}{Km_v m_w} \sum_{k,k'=1,k\neq k'}^{K} \sum_{i=1}^{m_v} \sum_{j=1}^{m_w} |\left(Corr(v^k, w^{k'})\right)_{ij}|$$
   (Penalizing this would **reduce** the dependency)

   B. $$R_1^K(v,w) = \frac{1}{Km_v m_w} \sum_{k,k'=1,k\neq k'}^{K} \sum_{i=1}^{m_v} \left(1 - |\left(Corr(v^k, w^{k'})\right)_{ii}|\right)$$
   (Penalizing this would **increase** the dependency)

4. Hence, we can induce the desired disentangled structure in $q_\phi(z)$ via regularization:

$$\mathcal{L}_{Aux-VAE} = \mathcal{L}_{VAE} + \lambda_1 \sum_{j=1}^{d} \left(R_1^K(u_j, \mu_{\phi,aux,j}) + R_0^K(u_j, \mu_{\phi,aux,-j})\right) + \lambda_2\left(R_0^K(u, \mu_{\phi,recon})\right)$$

Intra-group dependence within $Z_{aux}$

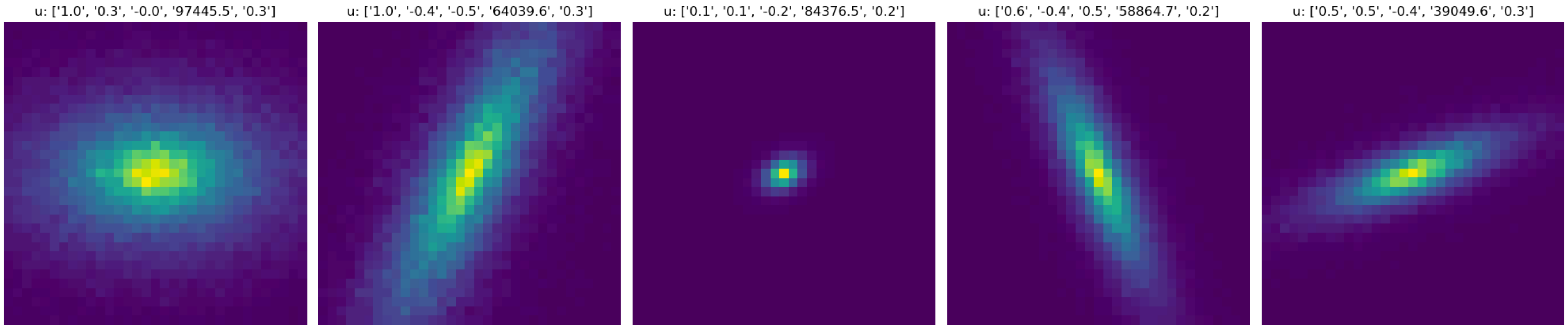
Inter-group dependence between $Z_{aux}$ and $Z_{recon}$

**Interpretations:**

- Ensure that each dimension of $Z_{aux}$ closely aligns with the auxiliary information $u$,
- Impose a penalty on the dependency between any two latent factors in $Z_{aux}$
- Reduce the dependency between $Z_{aux}$ and $Z_{recon}$.
- No restrictions on the dependency within $Z_{recon}$ to ensure good reconstruction quality.

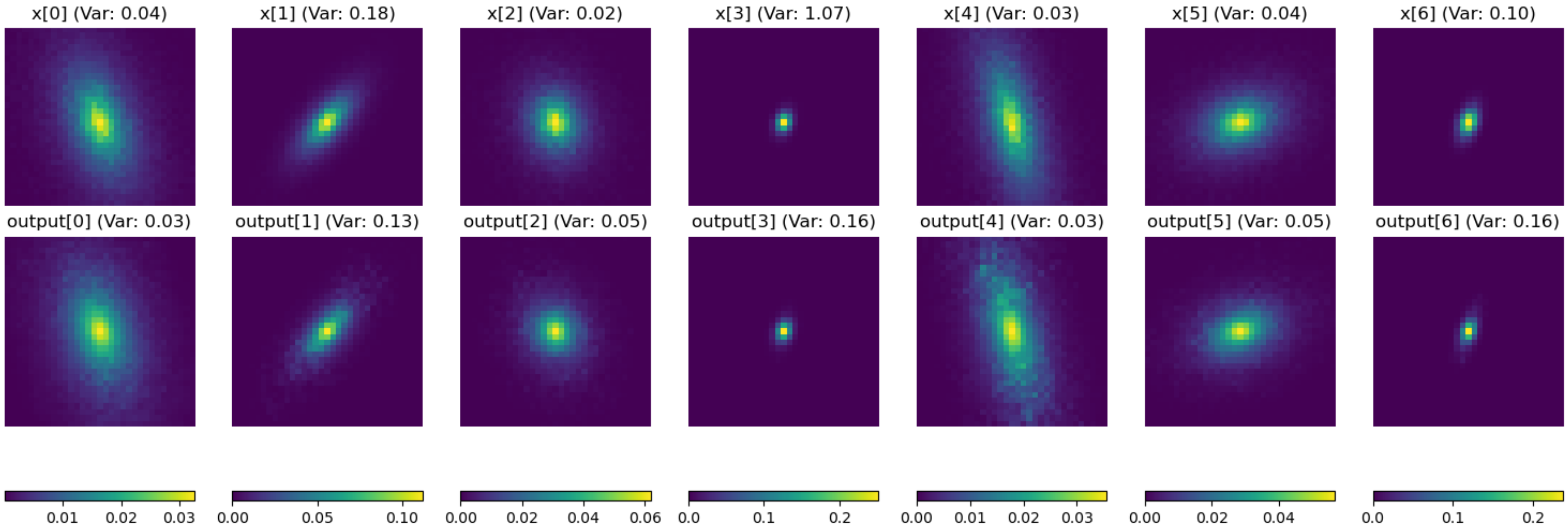
# EXAMPLE PROBLEM: SIMULATED GALAXY DATA

1. Galaxy images generated from geometric fits
   - ~160,000 galaxy images
   - 5 generating ground truth factors: **radius, g1, g2, flux, psf**
     - A few of these can be considered ‘known knowns’ (ground-truth factors) and others as ‘unknown knowns’.
     - These can be determined either from domain-knowledge or from correlation analysis.



# EXAMPLE PROBLEM: SIMULATED GALAXY DATA

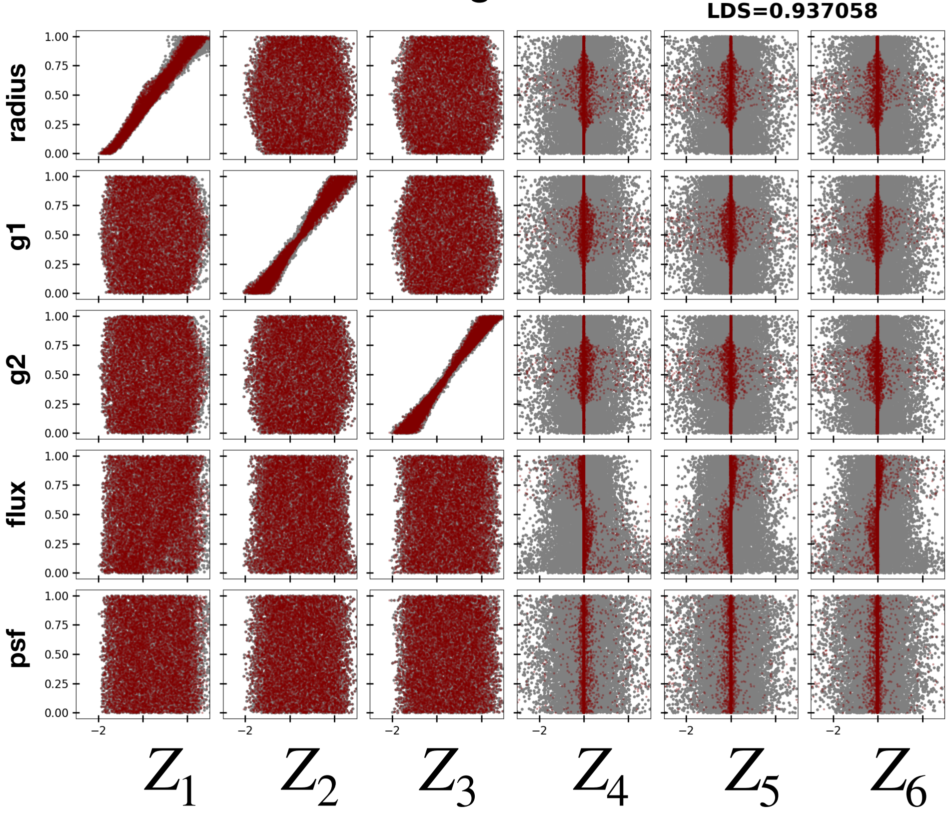
1. Realistic setting: we know/compute only some of the physical characteristics as ground truth information; but not all.
2. Only use three as auxiliary information: radius, g1, g2; and keep the remaining two factors as redundant - not use them in the training

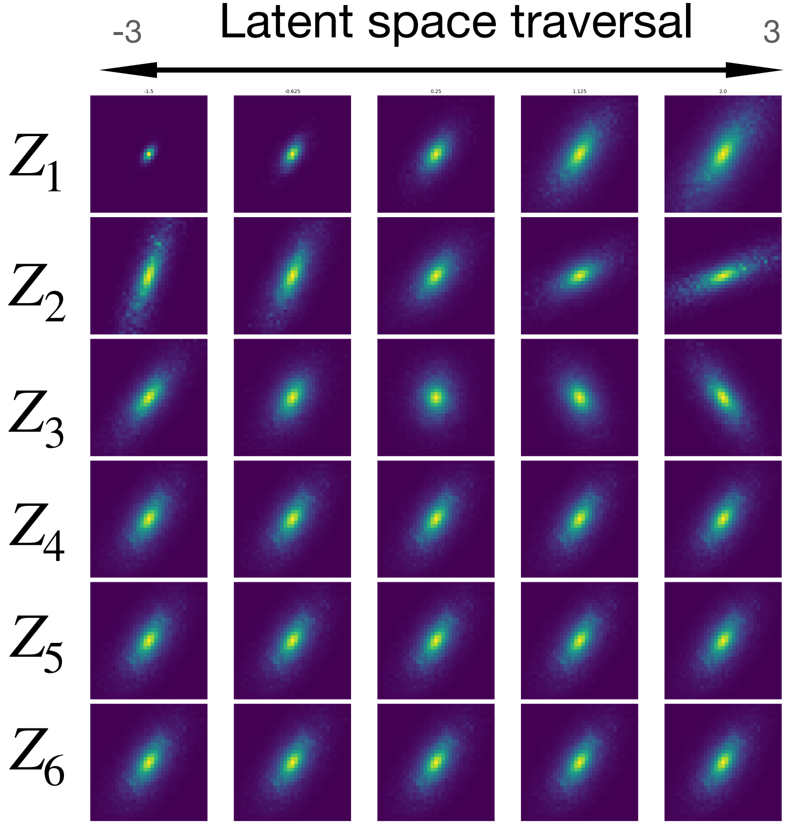


# INTERPRETABLE LATENT SPACE AND PROPER DISENTANGLEMENT

1. Latent space shows disentanglement for 3 variables, rest are entangled as usual (purely a part of reconstruction).
2. Enables a ‘principled’ generative ability — latent space traversal results in interpretable evolution.

Latent factors vs the generative factors



# APPLICATION: TOY QCF PROBLEM

**Basic setup:** Consider, the 1-D QCFs are represented as 2 beta-distributions $u() \equiv \beta(a_u, b_u), d() \equiv \beta(a_d, b_d)$.

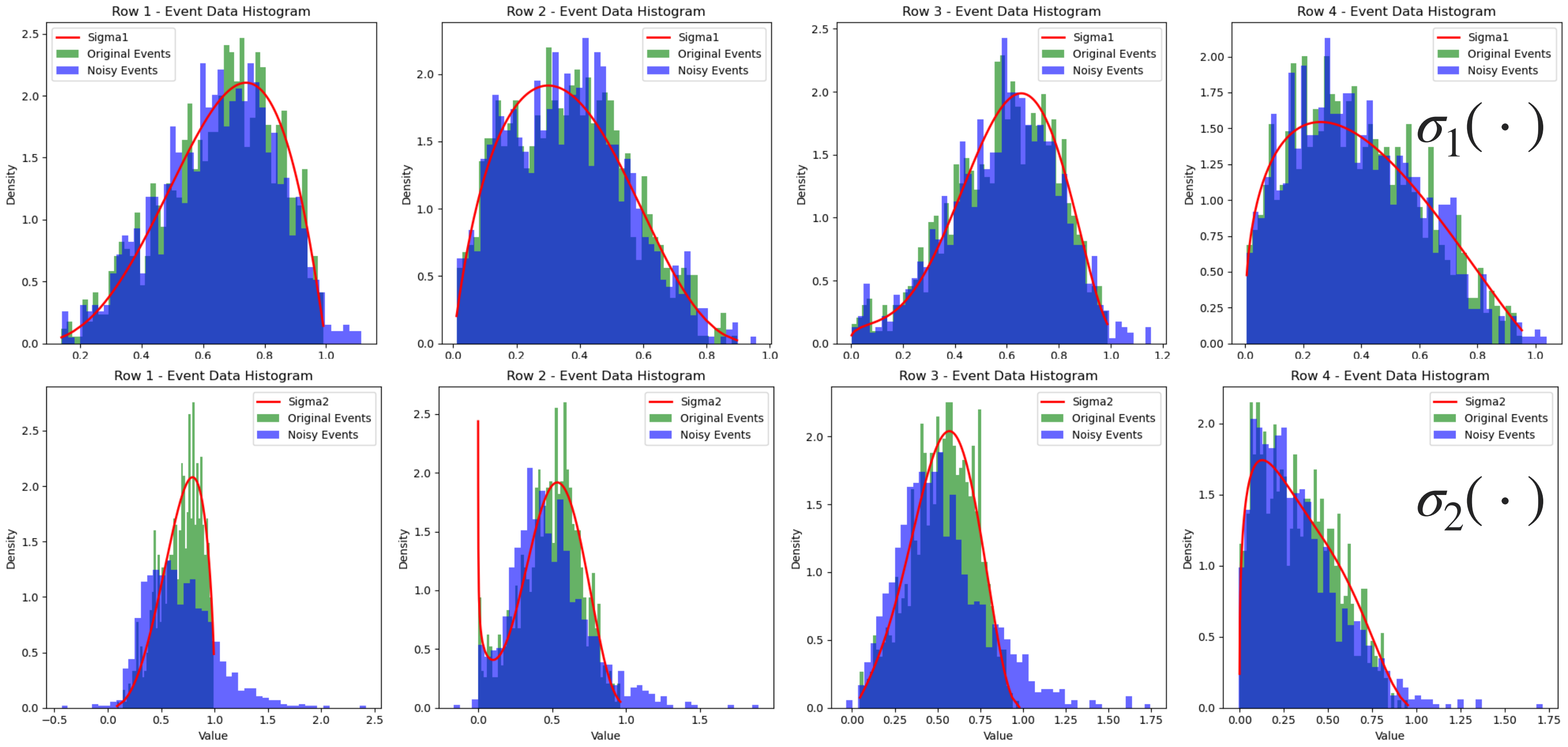
- The associated 1-D cross-sections are
  $\sigma_1(\cdot) = \frac{4}{5}u(\cdot) + \frac{1}{5}d(\cdot)$ and
  $\sigma_2(\cdot) = \frac{1}{5}u(\cdot) + \frac{4}{5}d(\cdot)$.
- The **theory generated events are passed through a detector model**: $x \mapsto x_d$, such that $x_d \sim N(0, cx^2)$, c=0.1, 0.5, 1,….
- Our observed database is the events coming out of this detector phase. Based on this, our goal is to estimate the parameters in the QCF model.

***Data* setup:** Creating a synthetic dataset to learn the underlying functional structure using a Variation AutoEncoder (VAE):

- Generate B different parameter configuration covering the whole parameter range:
  $\theta^i = (a_u^i, b_u^i, a_d^i, b_d^i) \sim Uniform(L_b, U_b)^4$
  (e.g. in our experiment, we set
  $B = 25000, L_b = -0.5, U_b = 5$
- For each parameter setting, we generate $n_{sim} = 1000$ events by sampling from the corresponding cross-section and then passing it through the detector model.
- This creates our dataset which consists of the pairs $(\theta^i, x^i)_{i=1}^B$

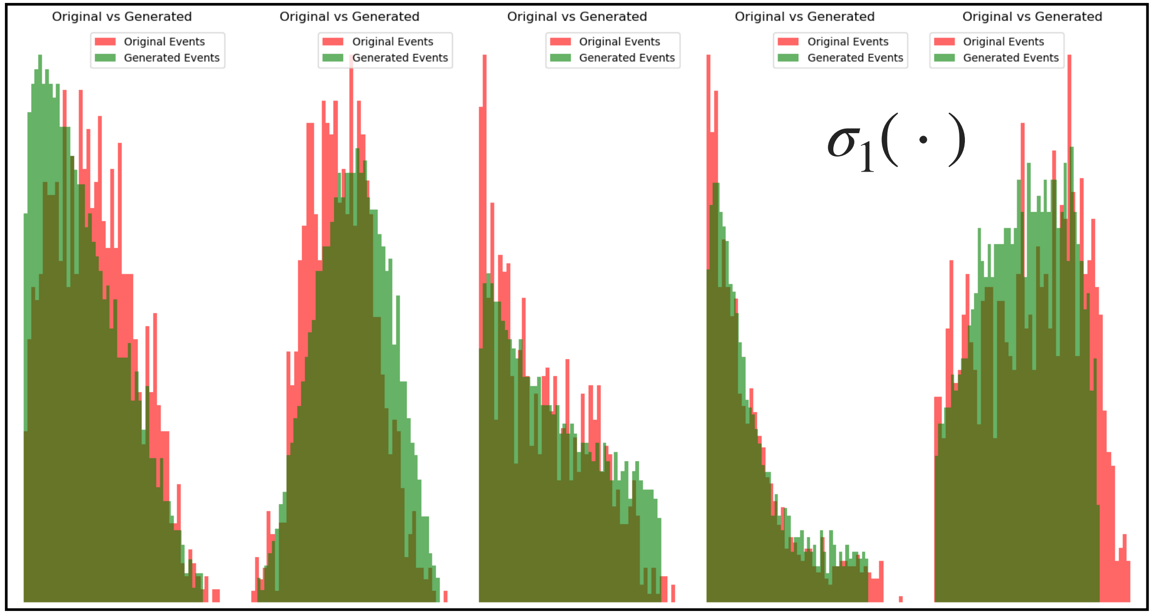
# APPLICATION: TOY QCF PROBLEM

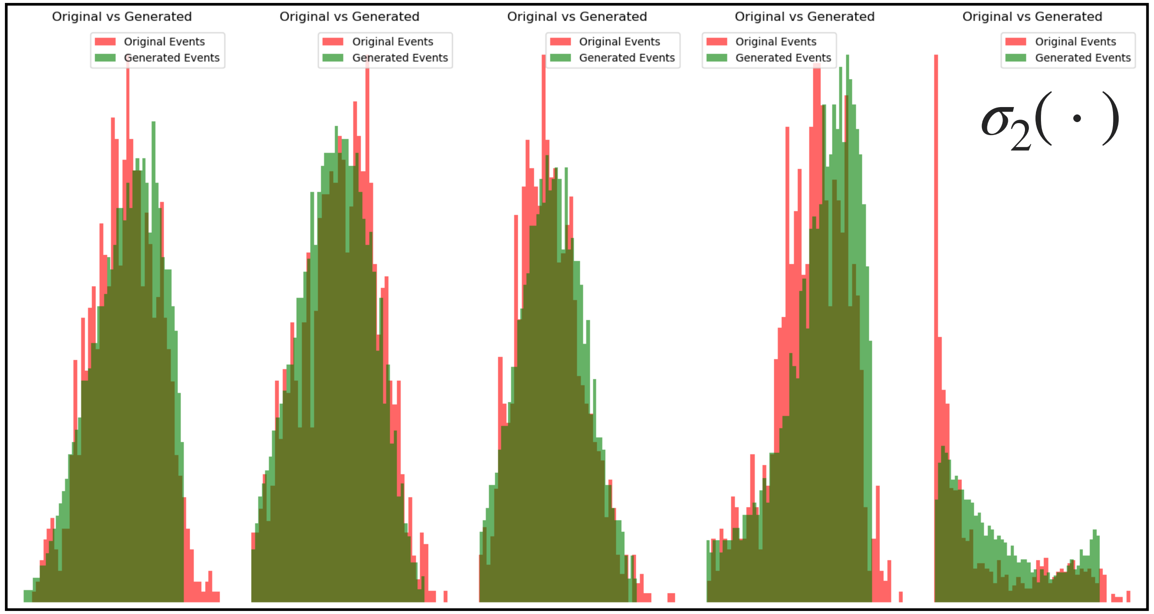
***Data* setup:** Events corresponding to 2 cross-sections are shown for pre-detector and post-detector stages.



# TOY-QCF EVENT RECONSTRUCTION

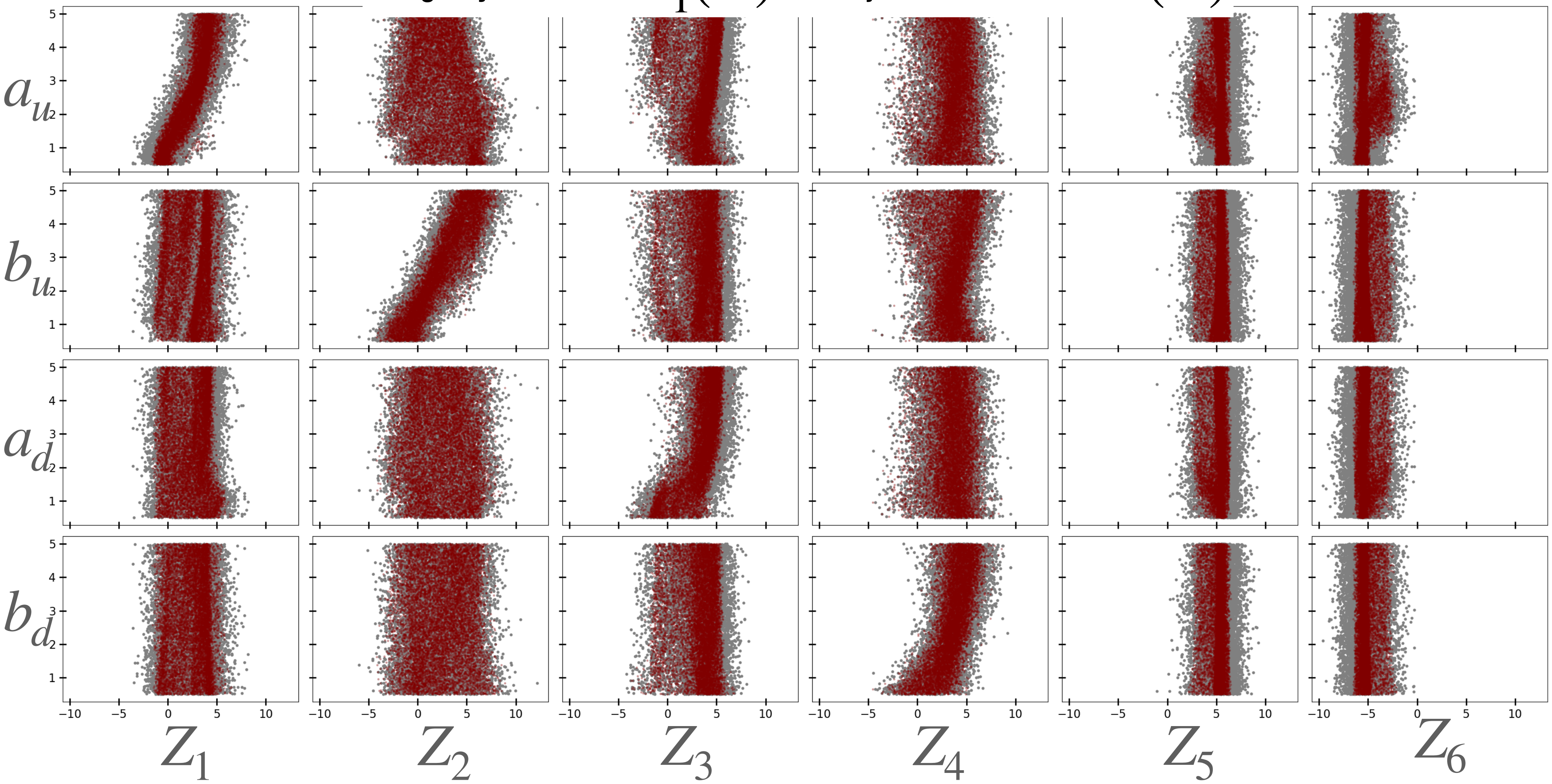
- **Aux-VAE** input/output dimensions are $d_{in} = d_{out} = n_{sim} = 1000$
- Bottleneck of the ***Aux-VAE*** to with six latent dimensions:
  - First four latent factors are associated with the QCF parameters.
  - Other 2 latent factors are associated with reconstruction-only (detector noise or any other effects — collectively but unspecified)



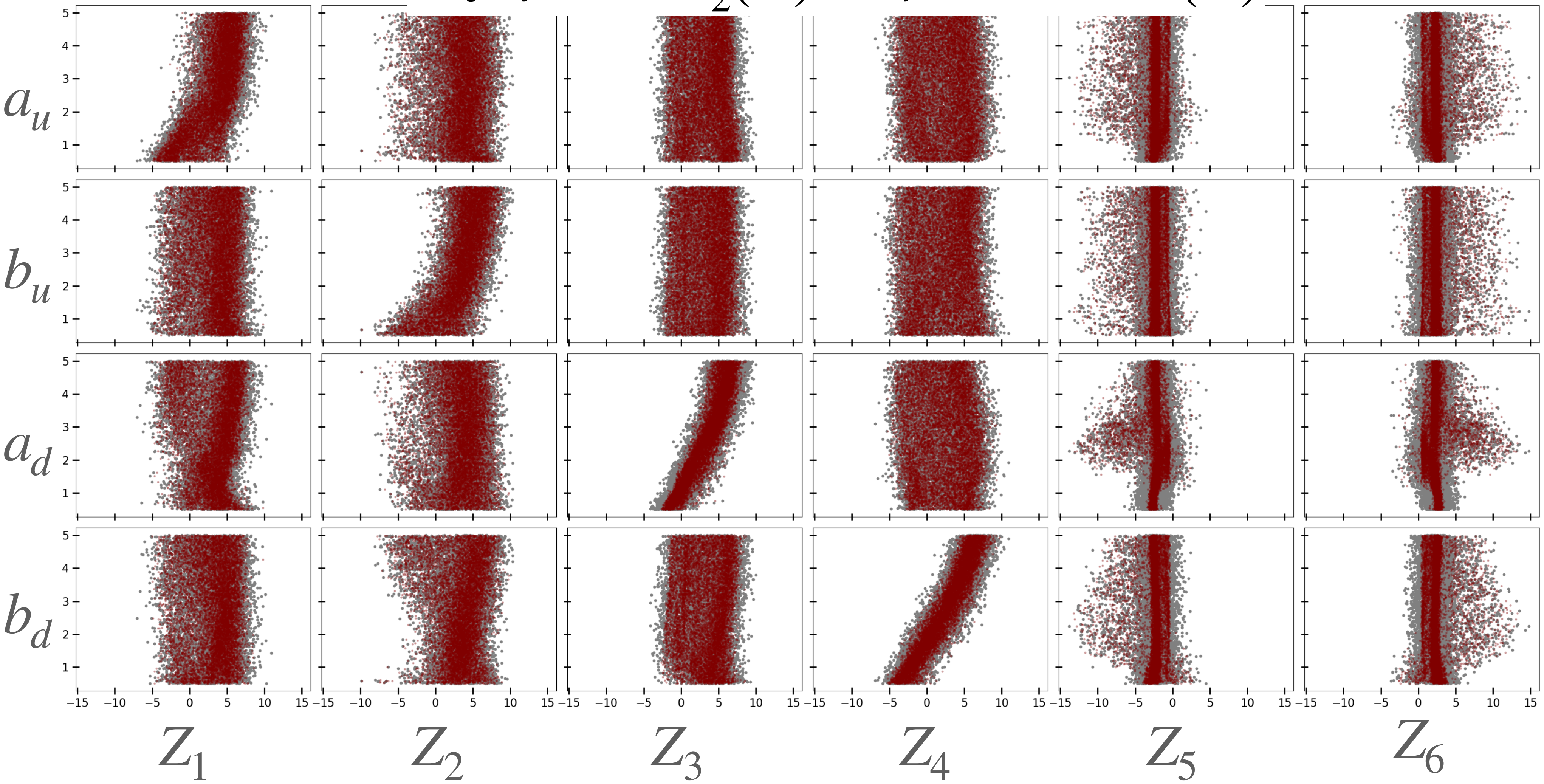
# AUX-VAE'S LATENT FACTORS VS QCF PARAMETERS ON TEST DATA

Using only events from $\sigma_1(\cdot)$ with major contribution from $u(\cdot)$

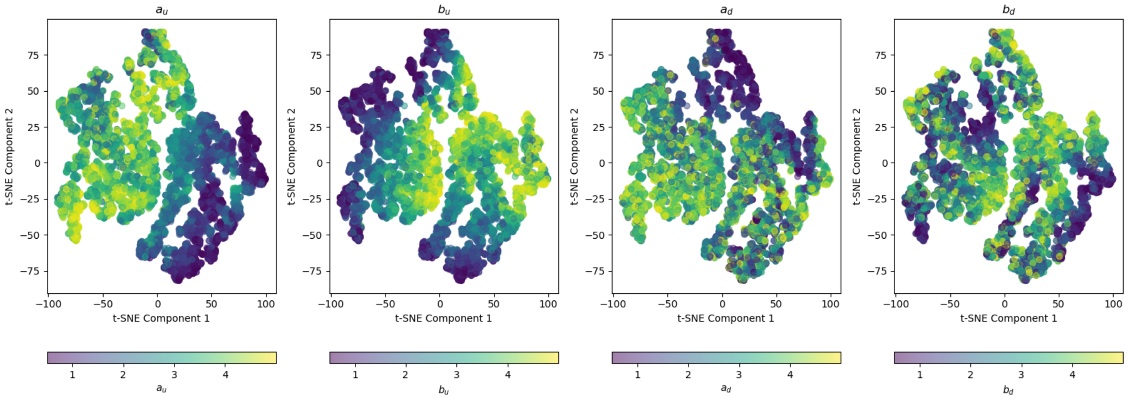


# AUX-VAE'S LATENT FACTORS VS QCF PARAMETERS ON TEST DATA

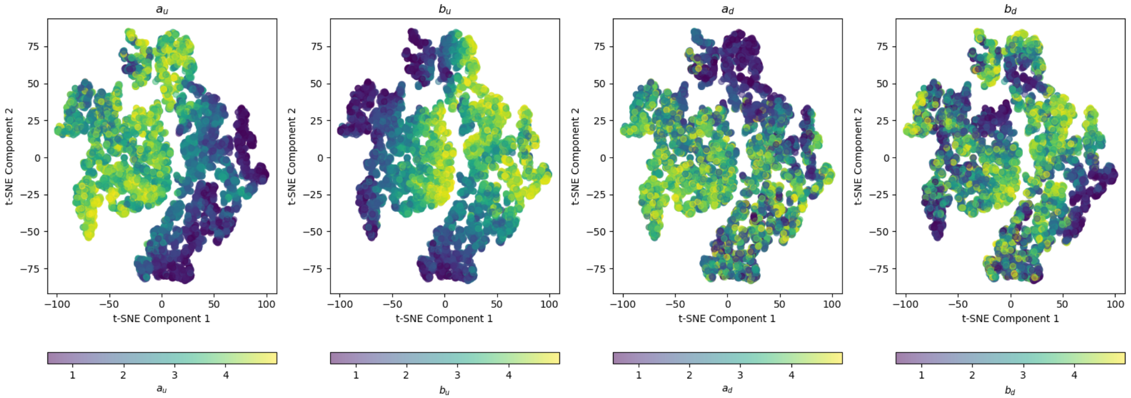
Using only events from $\sigma_2(\cdot)$ with major contribution from $d(\cdot)$



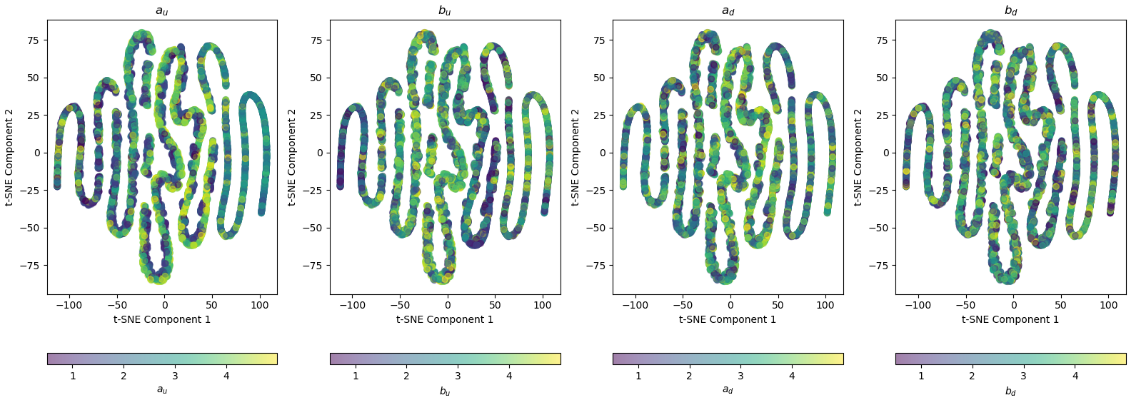
# VISUALIZING THE LATENT SPACE VIA T-SNE




**Using all six latent factors**




**Using only the first four latent factors**



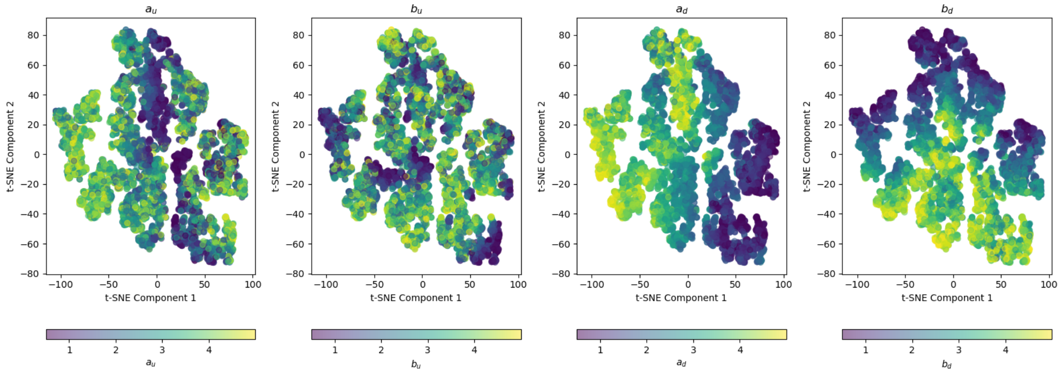

**Using only the last two latent factors**

Using only events from $\sigma_1(\cdot)$ with major contribution from $u(\cdot)$

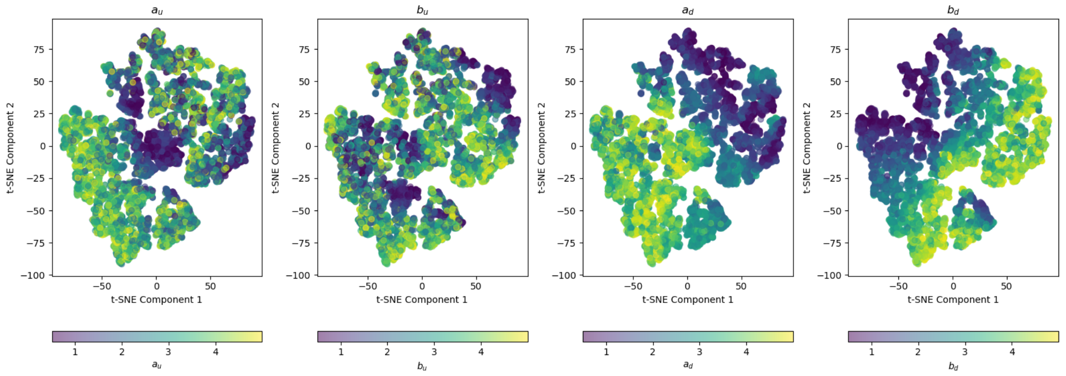
**Interpretation:**

- For $\sigma_1(\cdot)$, $a_u$ and $b_u$ show highest correlation
- The first 4 latent factors are nicely capturing the individual QCF parameters.
- Last 2 latent factors (corresponding to reconstructions do not capture much of the parameter variation) — mostly just the stochasticity.

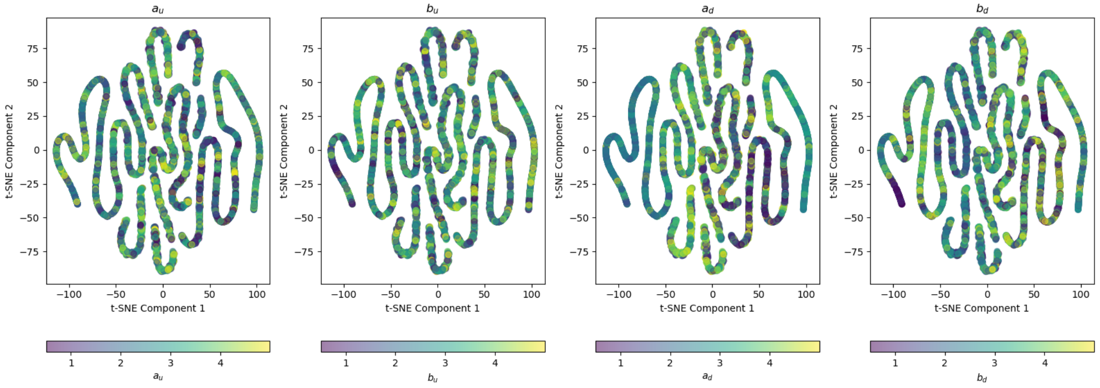
# VISUALIZING THE LATENT SPACE VIA T-SNE




**Using all six latent factors**




**Using only the first four latent factors**




**Using only the last two latent factors**

Using only events from $\sigma_2(\cdot)$ with major contribution from $d(\cdot)$

**Interpretation:**

- For $\sigma_2(\cdot)$, $a_d$ and $b_d$ show highest correlation
- The first 4 latent factors are nicely capturing the individual QCF parameters.
- Last 2 latent factors (corresponding to reconstructions do not capture much of the parameter variation) — mostly just the stochasticity.

# CONCLUSIONS AND FUTURE OUTLOOK

Argonne NATIONAL LABORATORY

- Representation learning can be powerful if the architectures/frameworks are designed with scientific applications in mind.
- A set of carefully curated changes have enabled us to disentangle latent space, while capturing ground truth parameters well.
  - Parameters such as noise/detector parameters can be treated separately, and focus of inference/sensitivity studies can be solely on physics parameters of interest.
  - Next trials will include higher dimensional cross-sections, more realistic detector models and possible parameterizations — **need inputs from the experts!**

  - Questions?