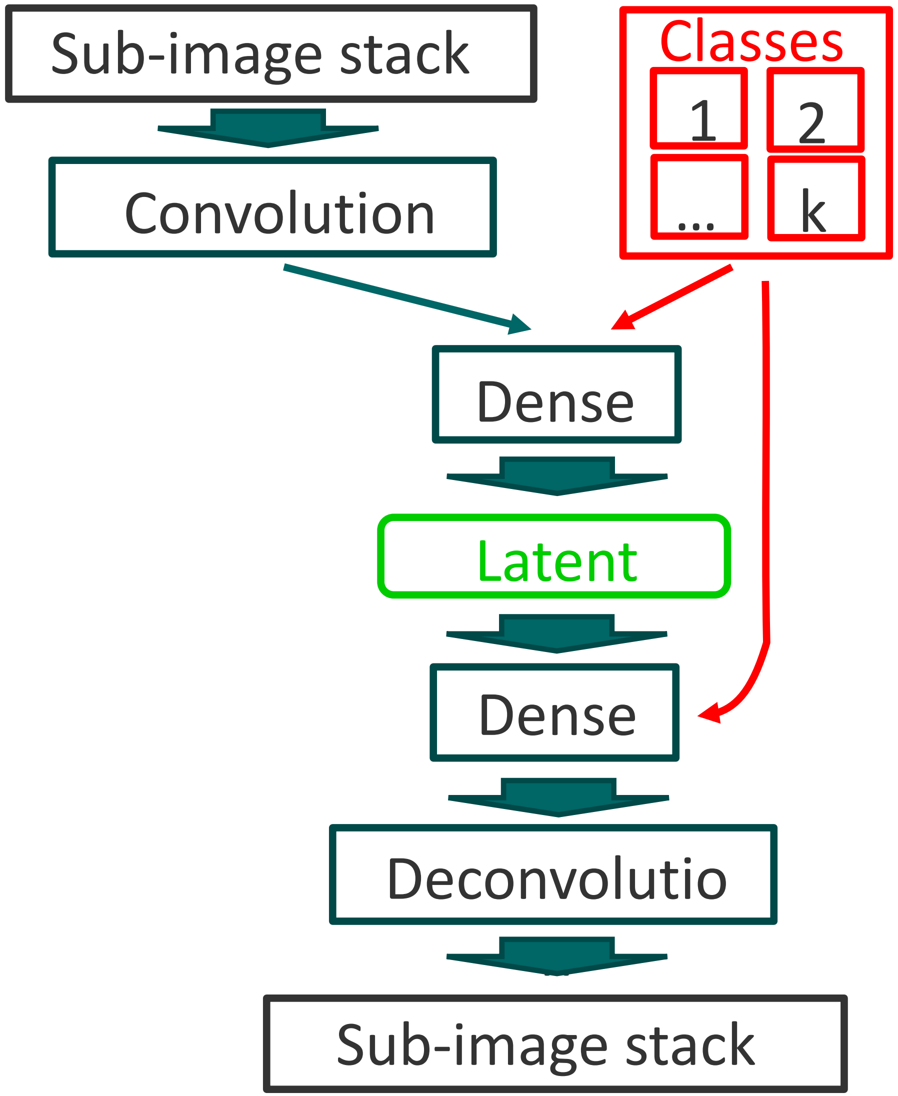
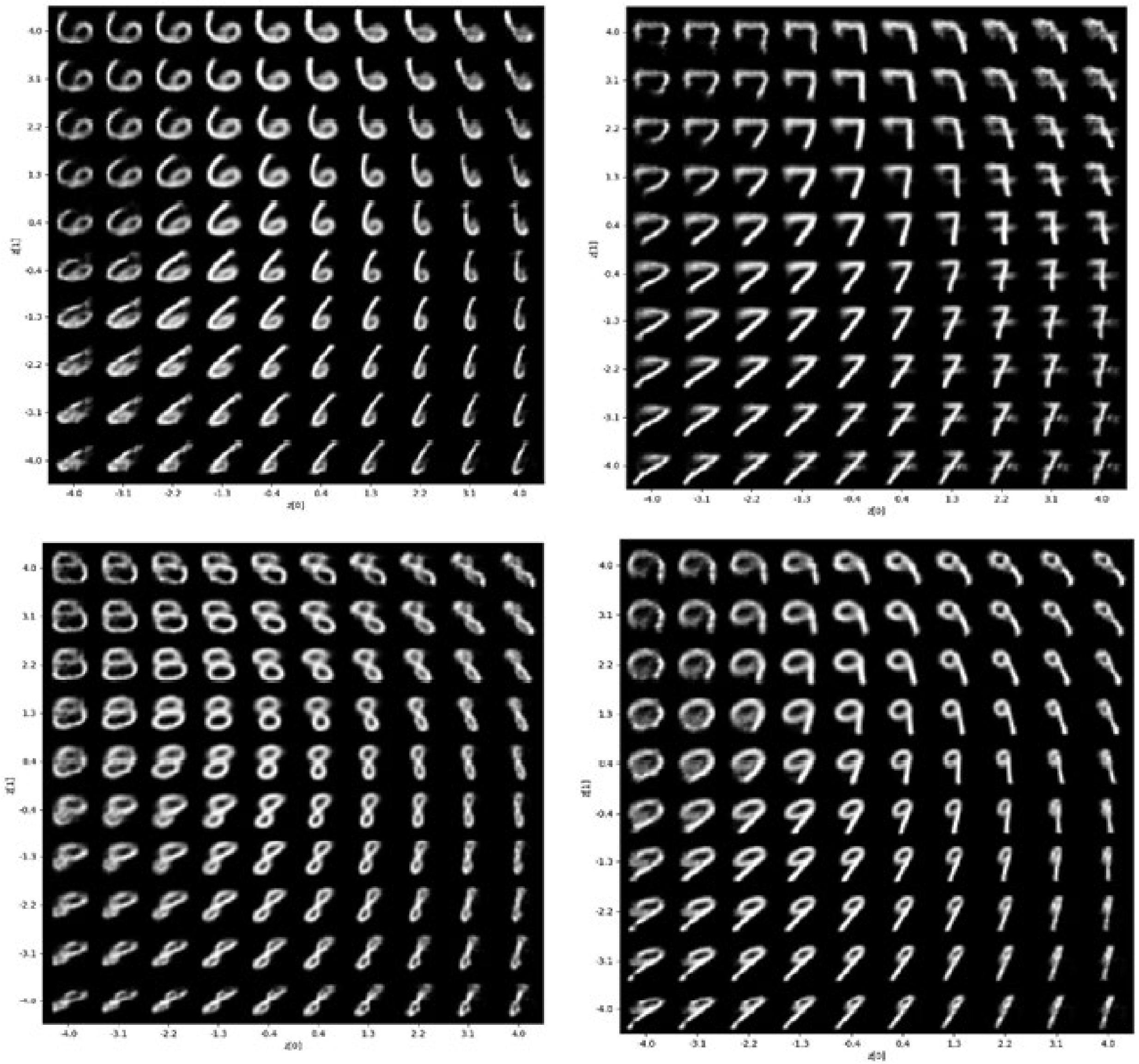
THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 30: Invariant Variational Autoencoders

Sergei V. Kalinin

Conditional VAE



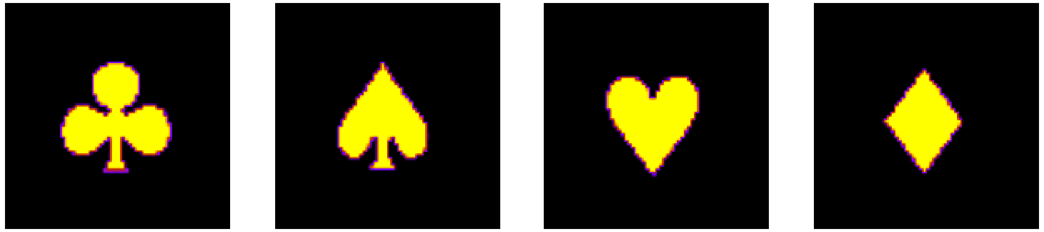


Note the trends in the latent representation for each digit: disentanglement of the representations

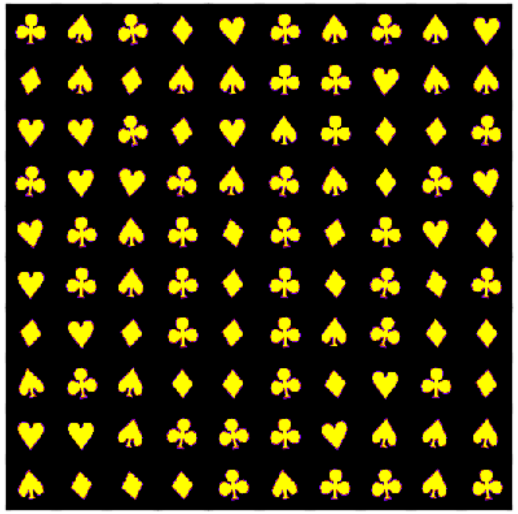
VAE on Cards

Introduce the **cards** data set:

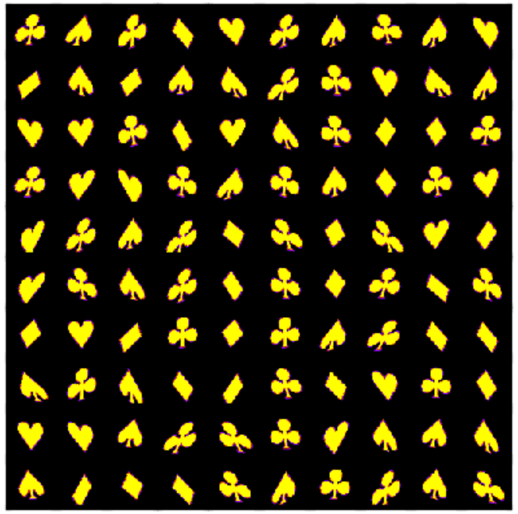
- Classical 4 hands (diamonds, clubs, pikes, hearts)
- Interesting similarities (pikes and hearts)
- And invariances on affine transforms (e.g. diamonds)



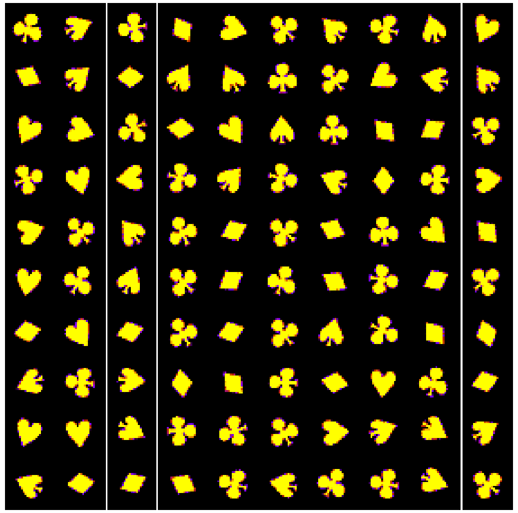
Cards 1: Low R (12 deg) and low S (1 deg)



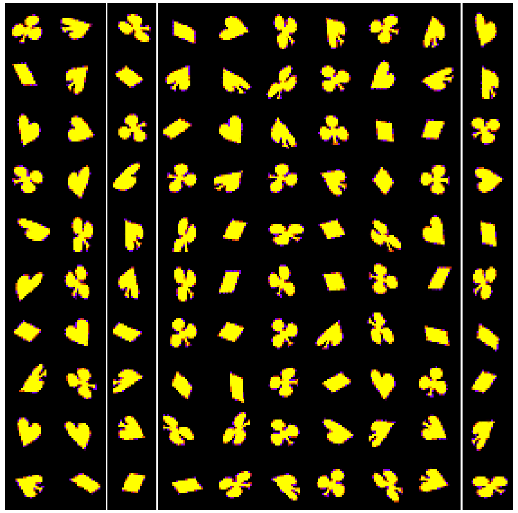
Cards 2: Low R (12 deg) and high S (20 deg)



Cards 3: High R (120 deg) and Low S (1 deg)



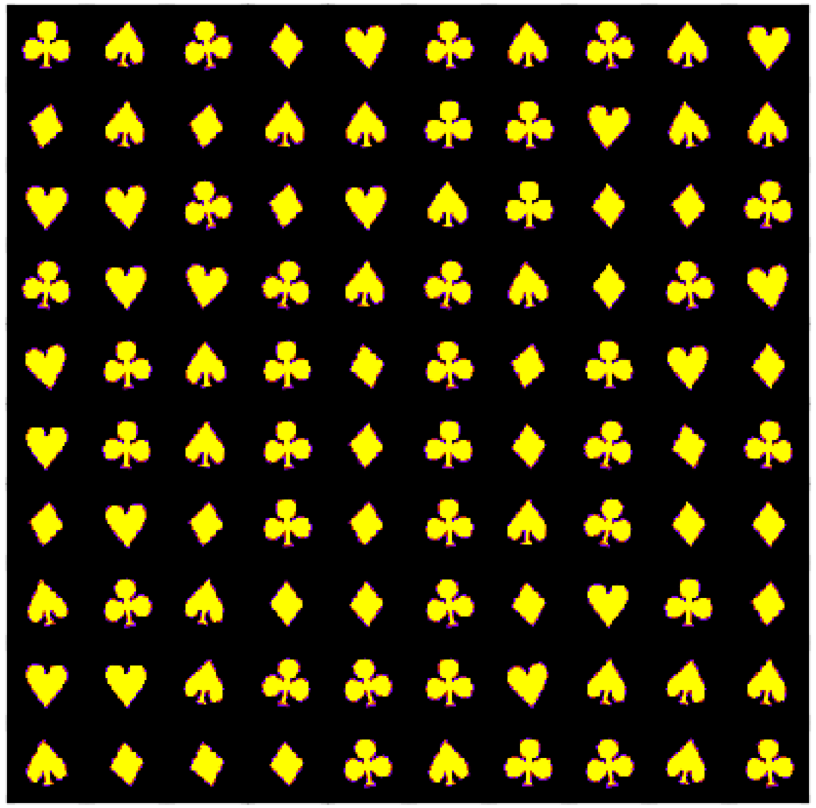
Cards 4: High R (120 deg) and high S (20 deg)



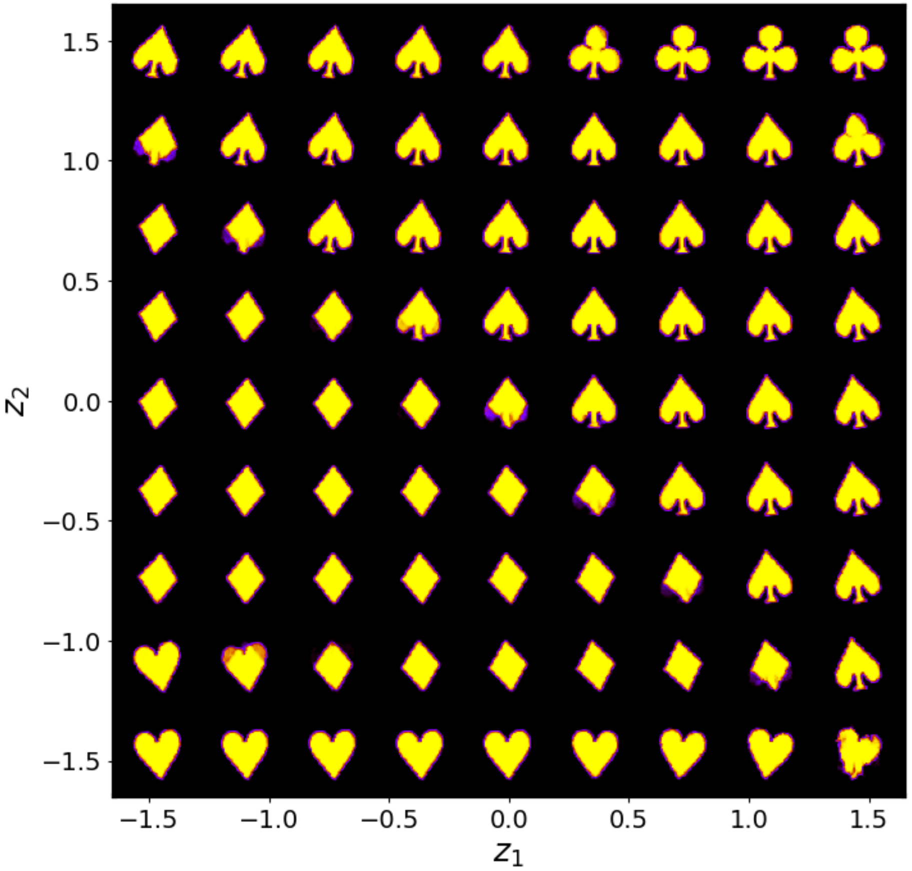
- Shear, rotations, and translations are **known** factors of variability (or traits) in data
- Can VAE disentangle representations and discover these factors of variability

VAE on Cards

Example of data

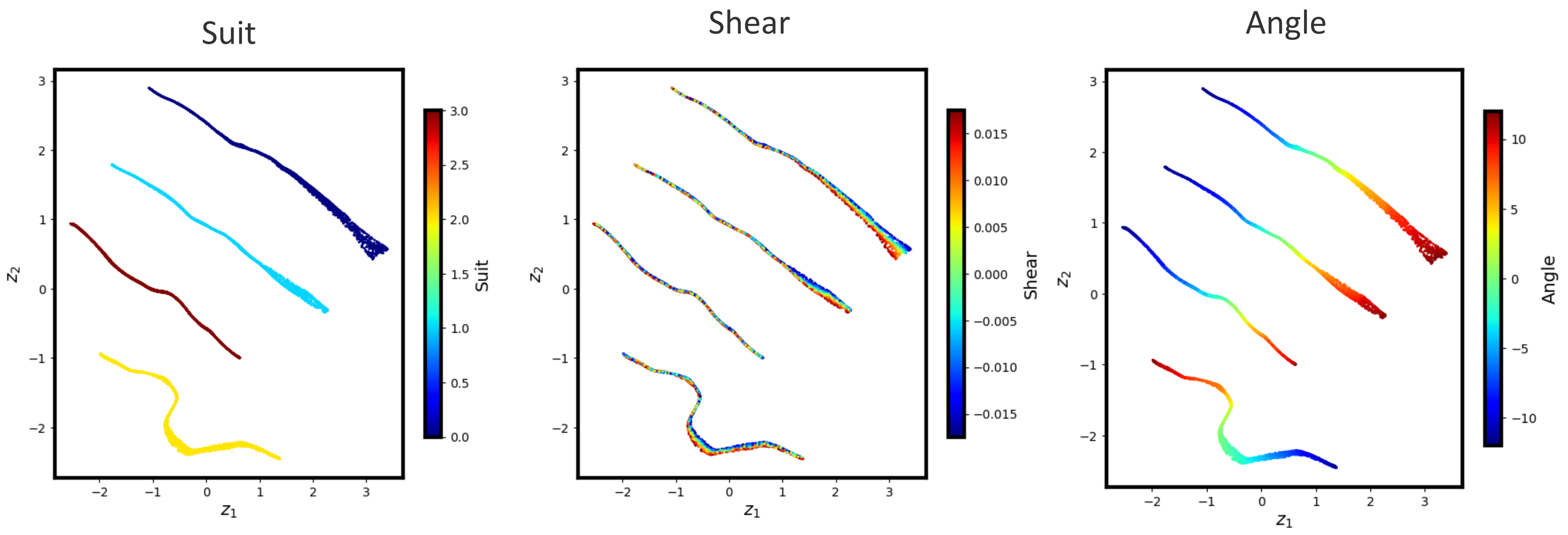


Latent representation



Cards 1: Low rotation (12 deg) and low shear (1 deg)

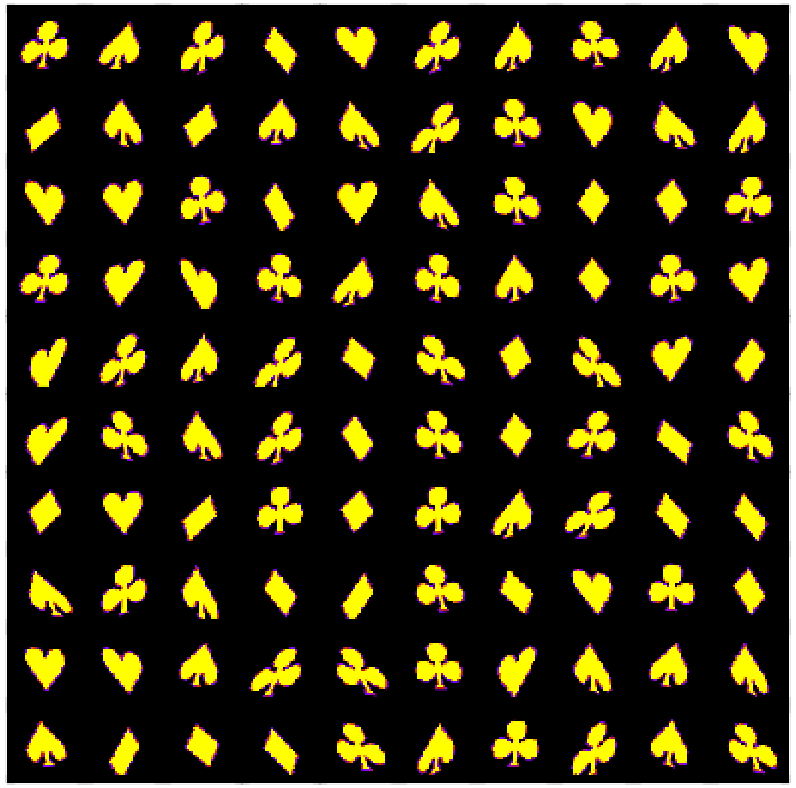
VAE on Cards



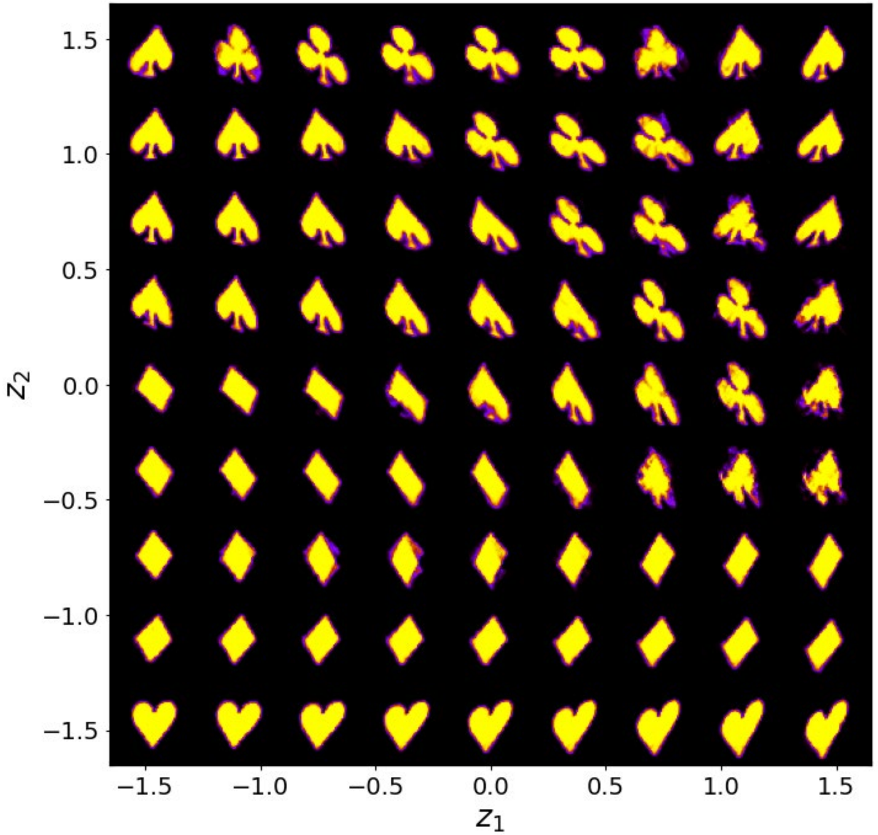
Cards 1: Low rotation (12 deg) and low shear (1 deg)

VAE on Cards

Example of data

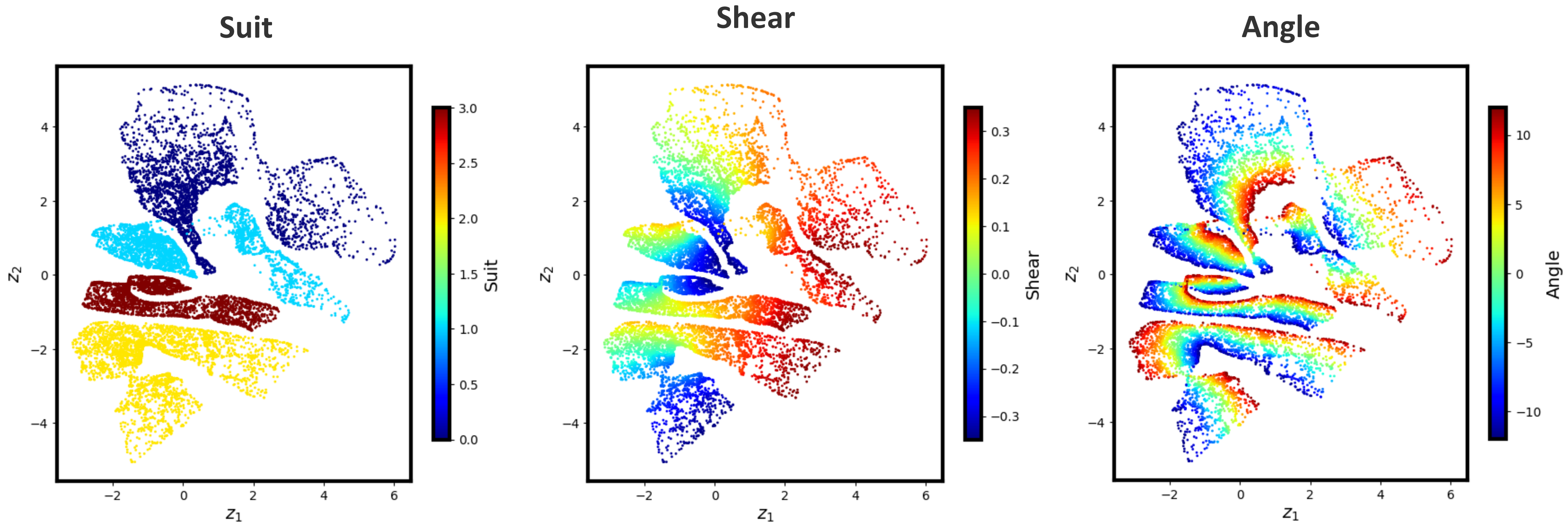


Latent representation



Cards 2: Low rotation (12 deg) and high shear (20 deg)

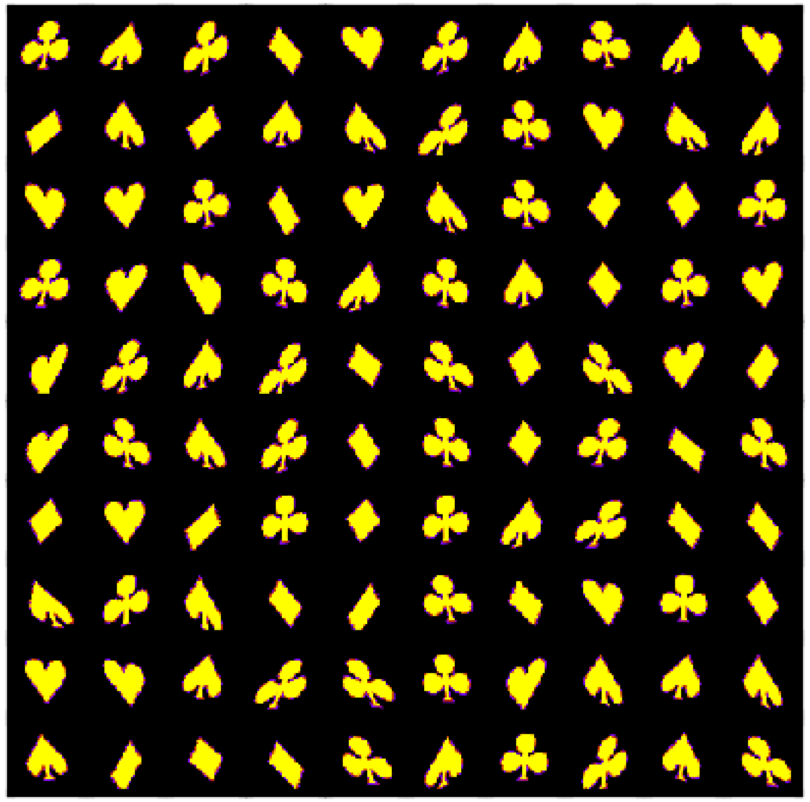
VAE on Cards



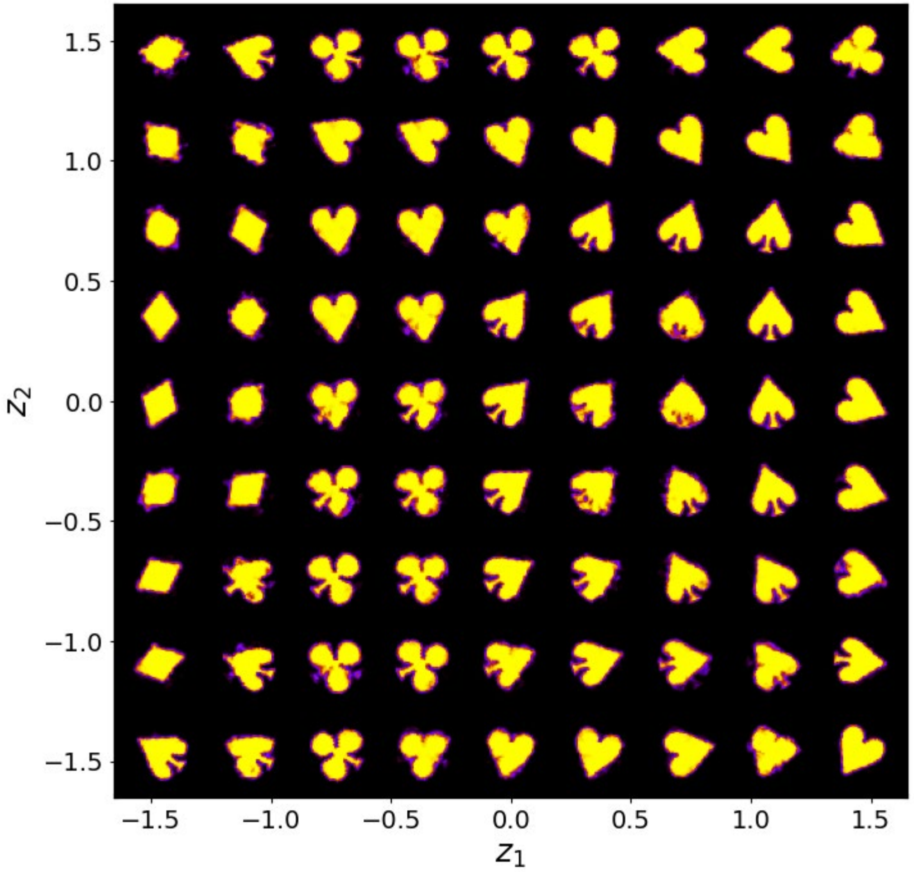
Cards 2: Low rotation (12 deg) and high shear (20 deg)

VAE on Cards

Example of data

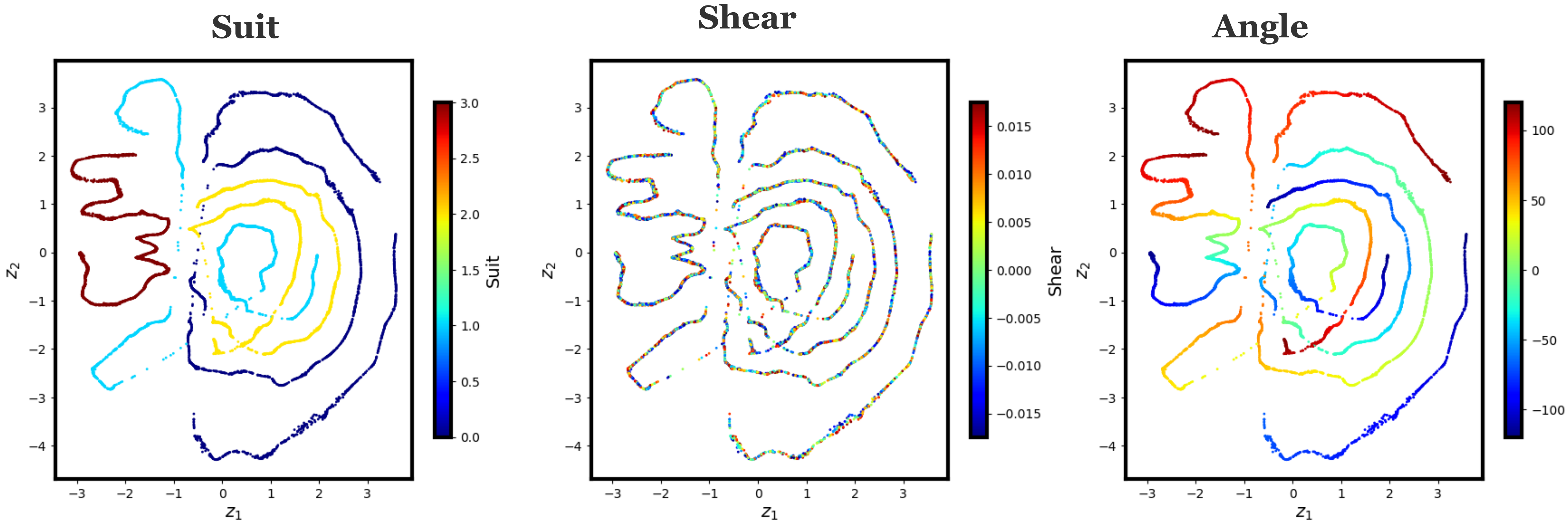


Latent representation



Cards 3: High rotation (120 deg) and low shear (1 deg)

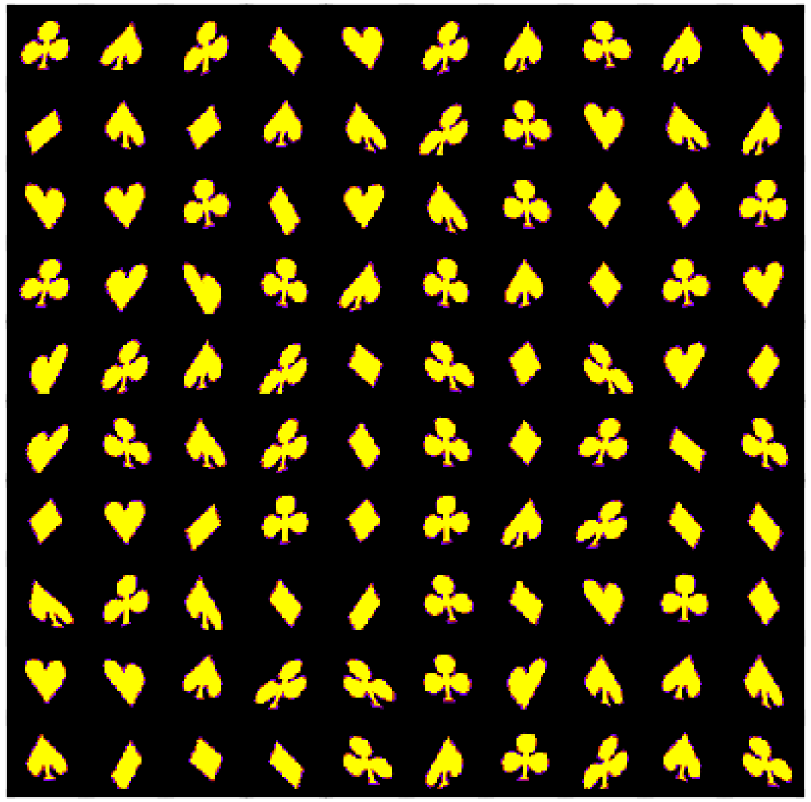
VAE on Cards



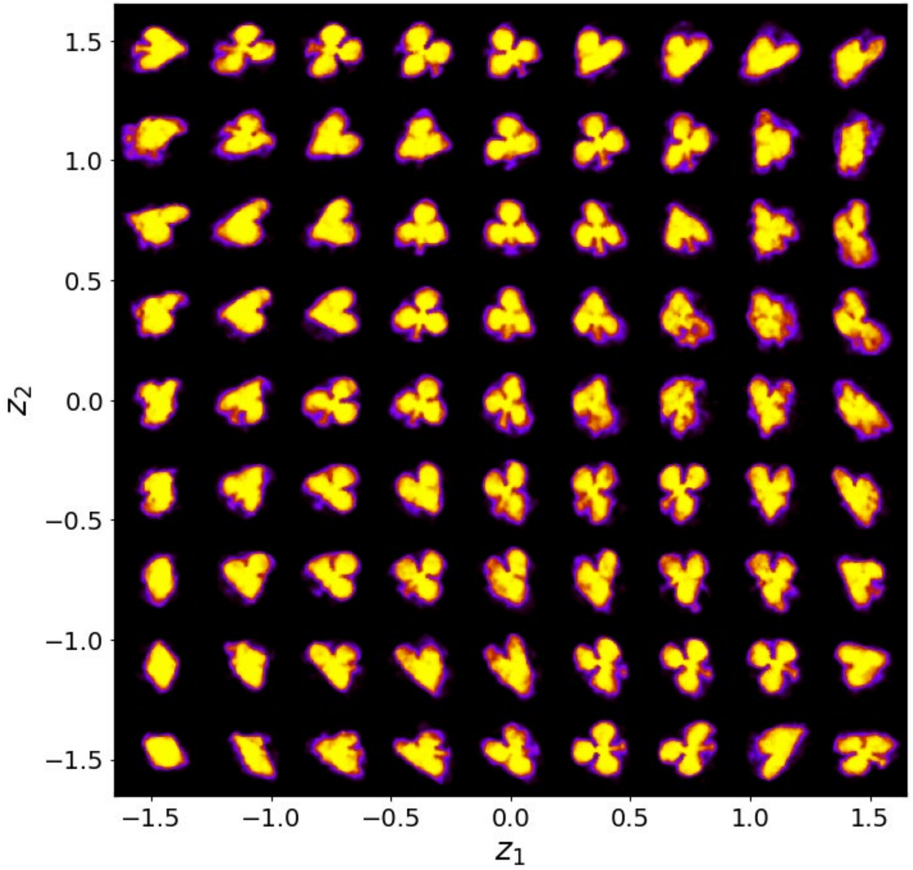
Cards 3: High rotation (120 deg) and low shear (1 deg)

VAE on Cards

Example of data

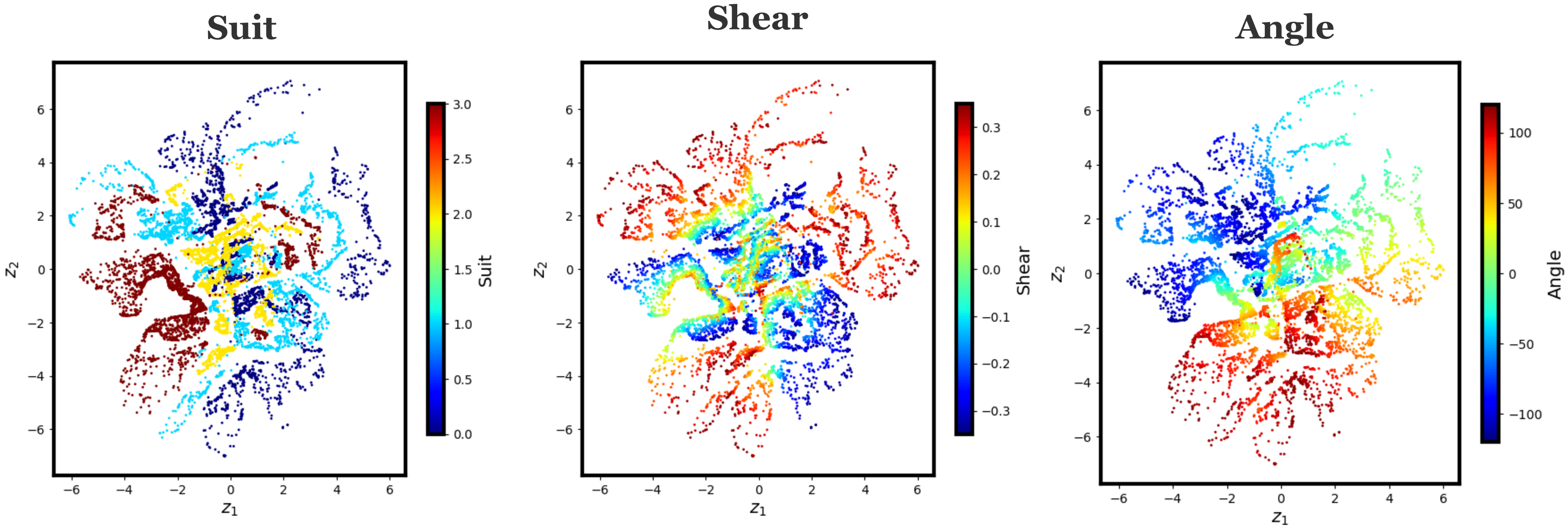


Latent representation



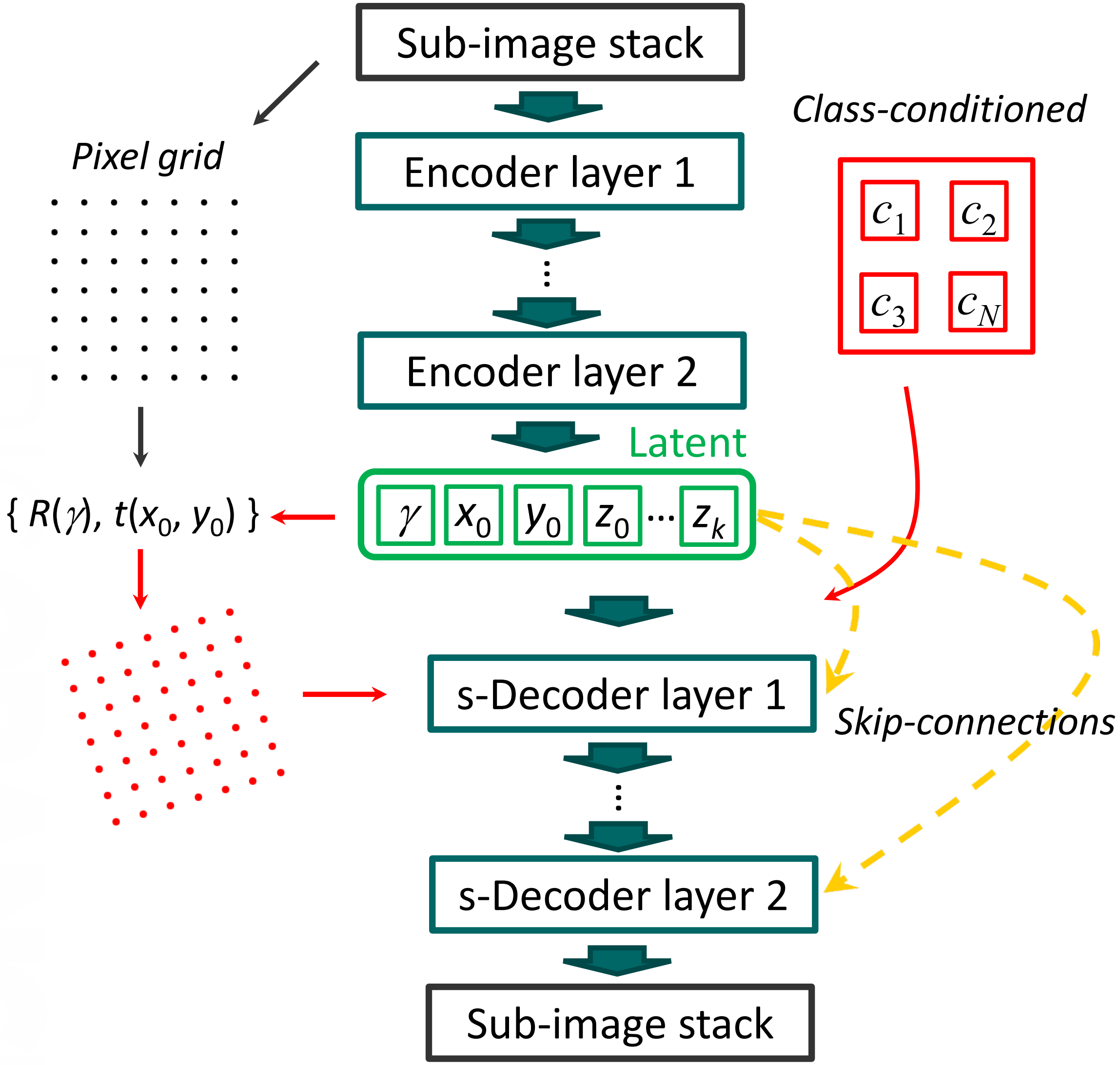
Cards 4: High rotation (120 deg) and high shear (20 deg)

VAE on Cards



Cards 4: High rotation (120 deg) and high shear (20 deg)

Conditional VAE



- Generative model is a function of spatial coordinate
- 3 additional latent variables to absorb rotations and shifts
- Disentangles rotations and translations from image content
- Ideal for analyzing microscopy sub-images on atomic level

$$
\left.\begin{aligned} ELBO &= Reconstruction\ Loss \\ &- D_{KL}\big(q(z|x)\|\mathcal{N}(0,I)\big) \end{aligned}\right\} \text{Regular VAE}
$$

$$- D_{KL}\left(q(\gamma|x)\|\mathcal{N}(0,s_\gamma^2)\right) \quad \text{Rotation}$$

$$- D_{KL}\left(q(\Delta r|x)\|\mathcal{N}(0,s_{\Delta r}^2)\right) \quad \text{Translation}$$

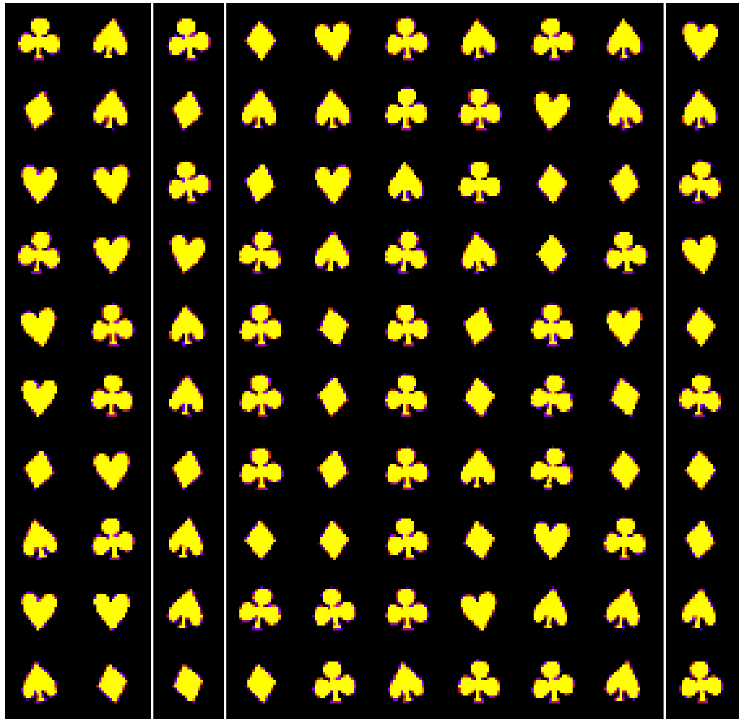
+ D_{KL}(physics-based "priors") ?

+ D (physics) ?

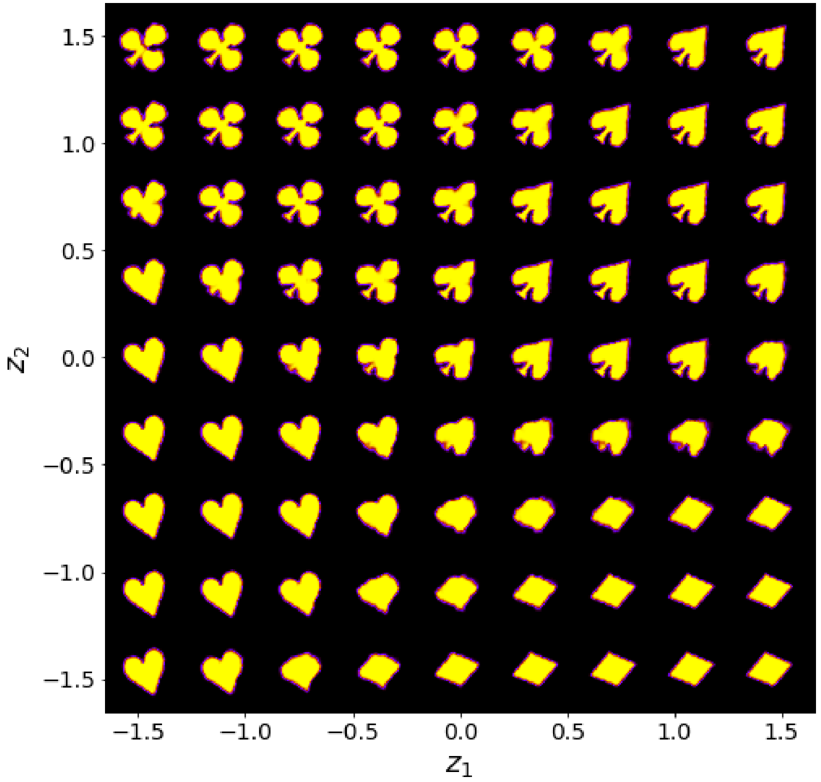
Bepler *et al.*, NeurIPS 2019

rVAE on Cards

Example of data

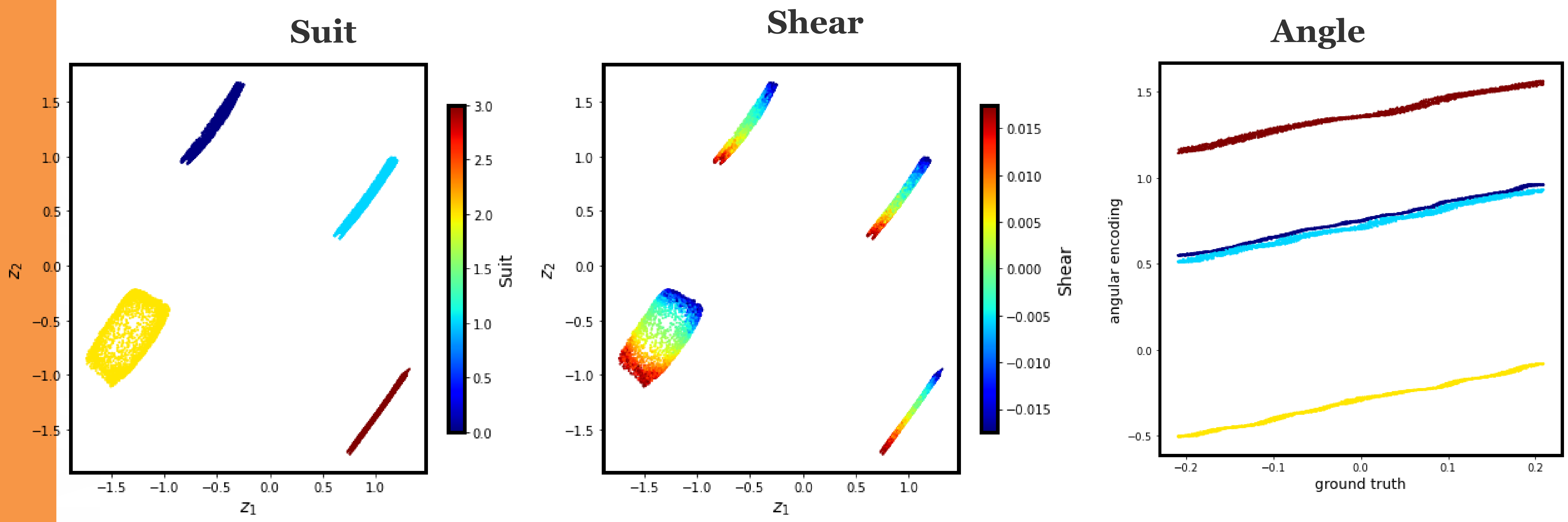


Latent representation



Cards 1: Low rotation (12 deg) and low shear (1 deg)

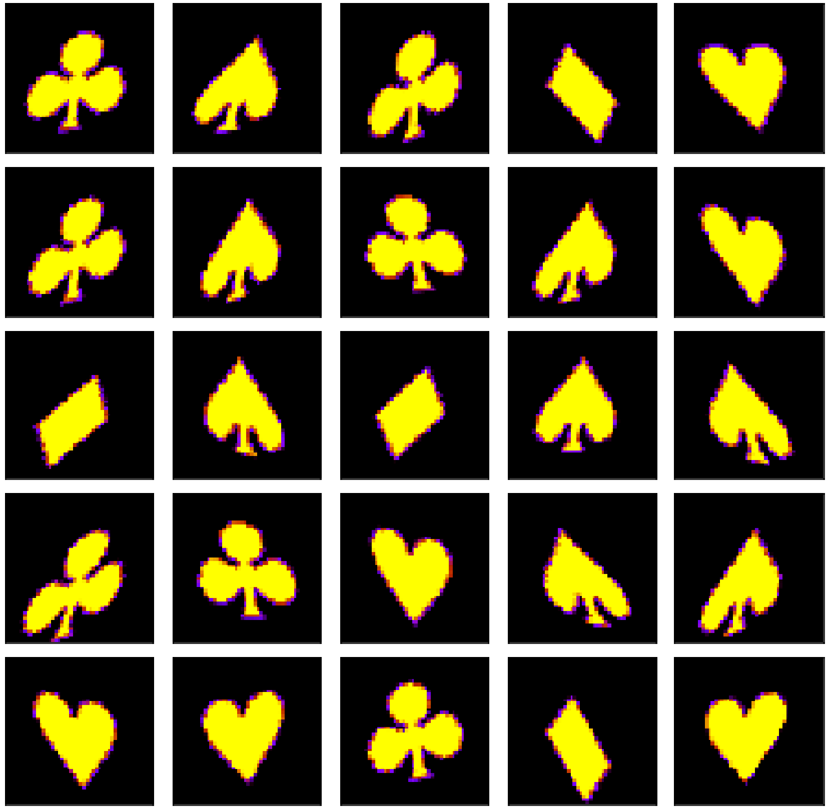
rVAE on Cards



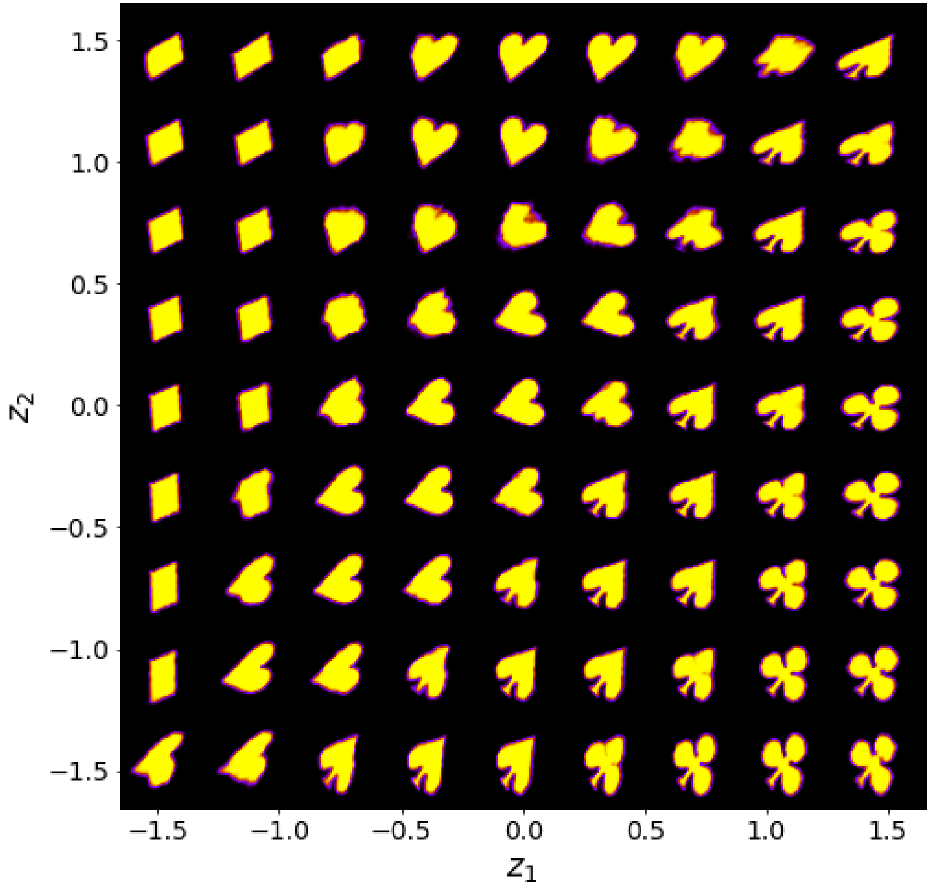
Cards 1: Low rotation (12 deg) and low shear (1 deg)

rVAE on Cards

Example of data

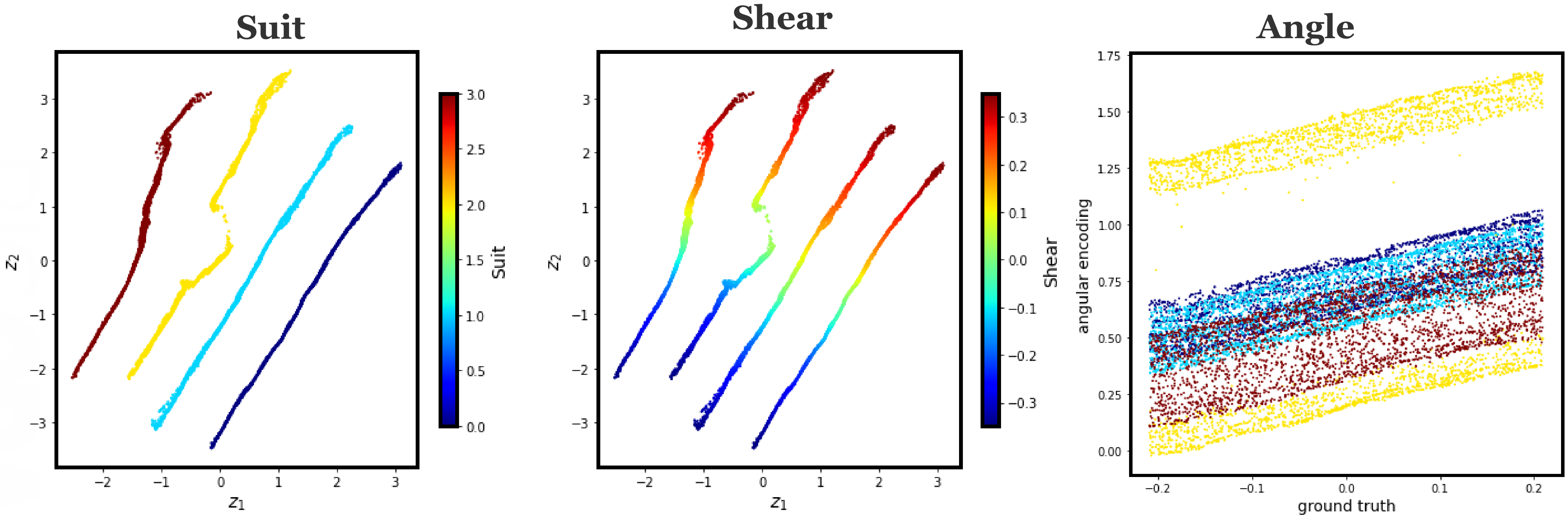


Latent representation



Cards 2: Low rotation (12 deg) and high shear (20 deg)

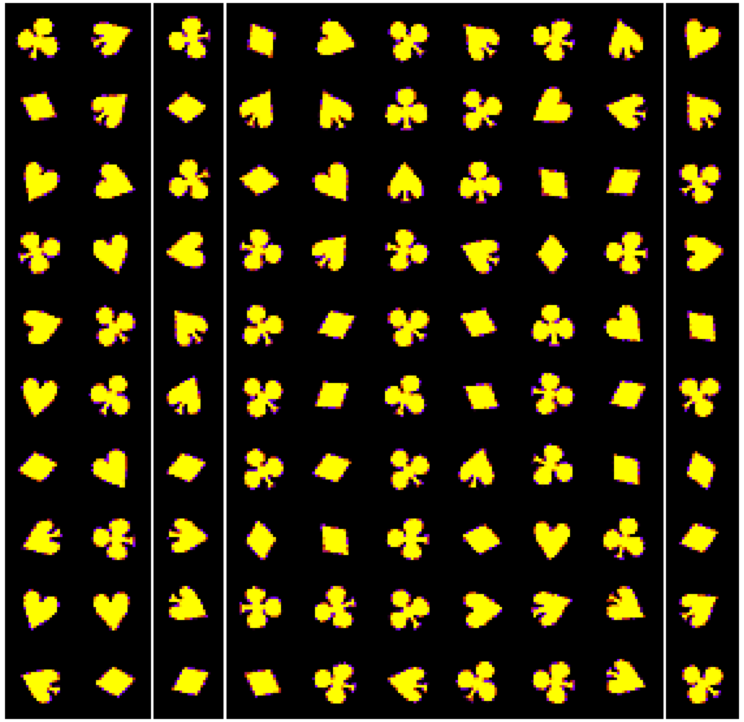
rVAE on Cards



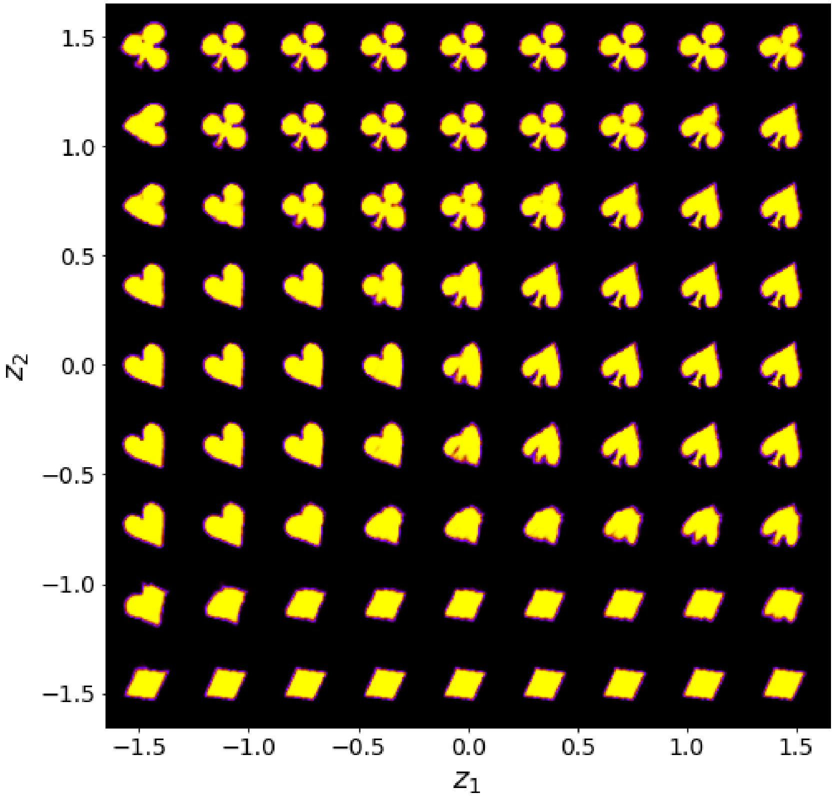
Cards 2: Low rotation (12 deg) and high shear (20 deg)

rVAE on Cards

Example of data

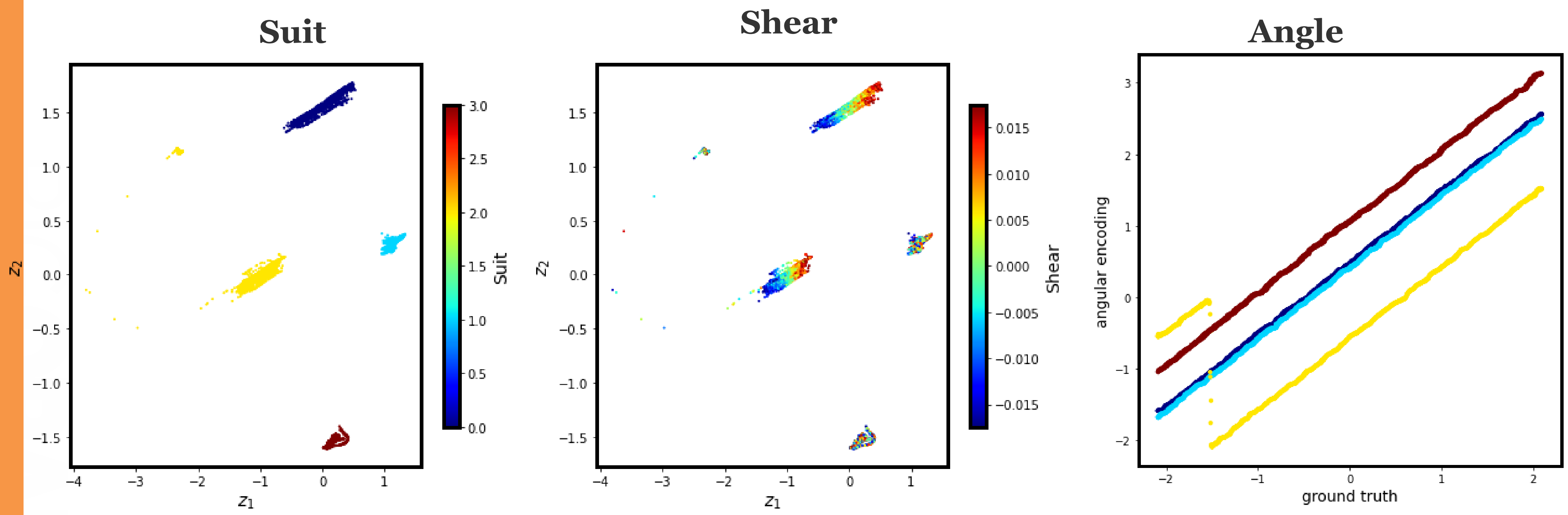


Latent representation



Cards 3: High rotation (120 deg) and low shear (1 deg)

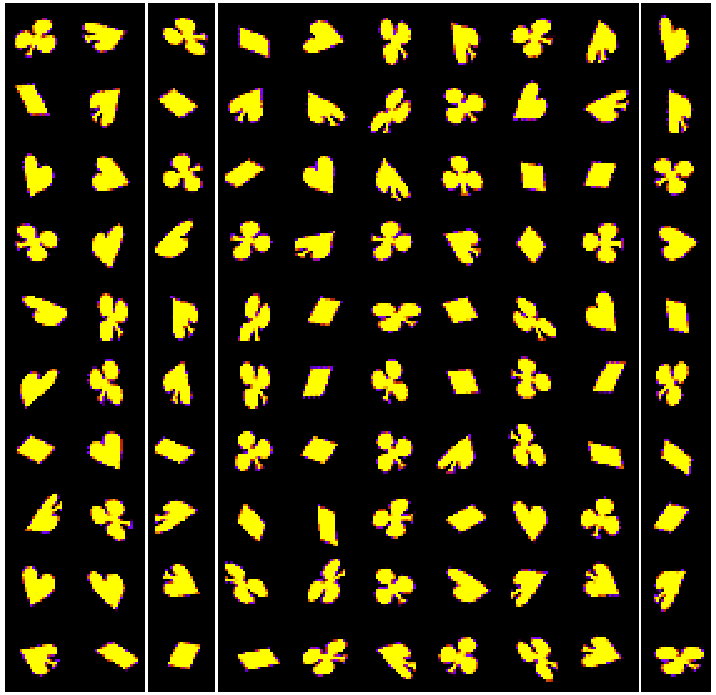
rVAE on Cards



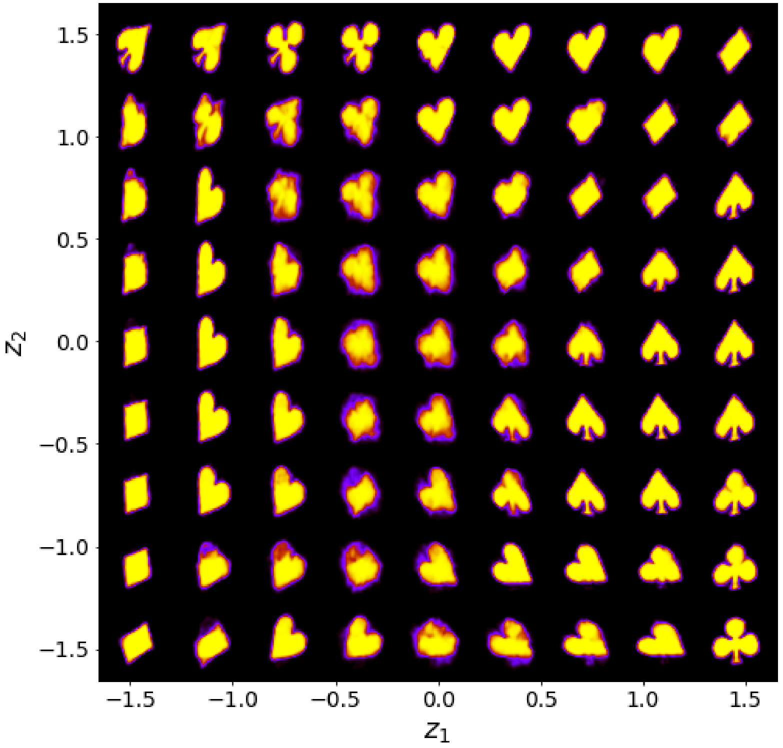
Cards 3: High rotation (120 deg) and low shear (1 deg)

rVAE on Cards

Example of data

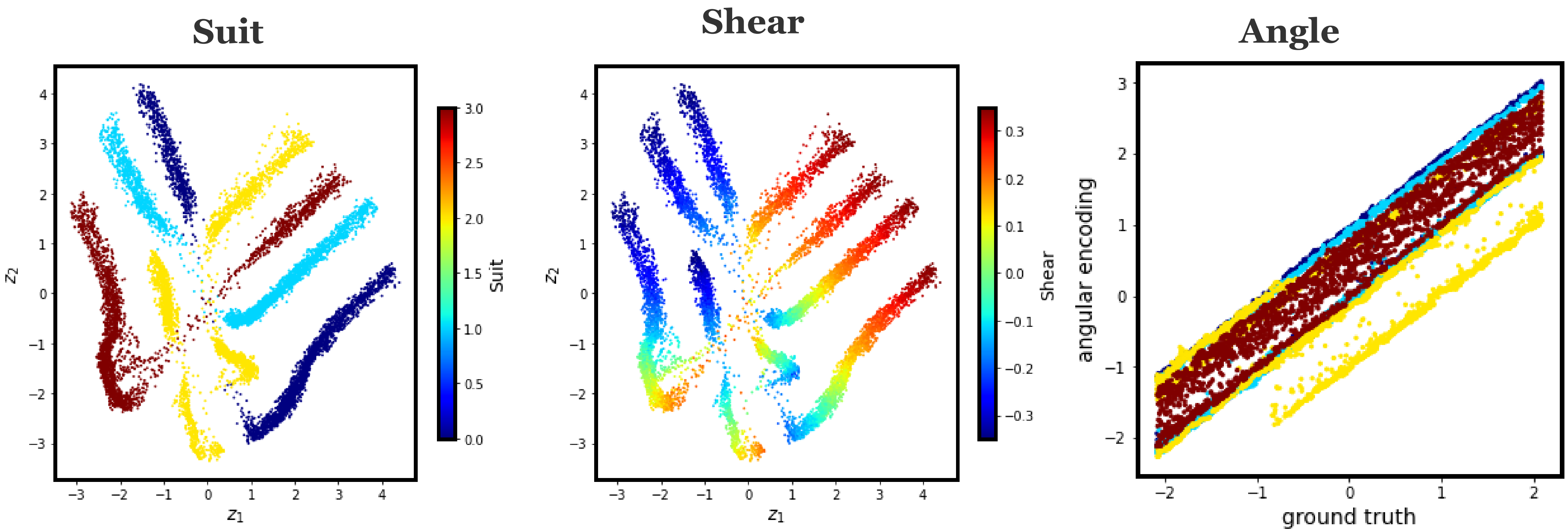


Latent representation



Cards 4: High rotation (120 deg) and high shear (20 deg)

rVAE on Cards



Cards 4: High rotation (120 deg) and high shear (20 deg)

VAE applications for real problems

- We often deal with complex data sets containing the information on physics of objects we seek to understand
- This can be spectral data sets (EELS in STEM, CITS in STM, complex spectroscopies in PFM) or single, multimodal, or hyperspectral images
- Often, we seek to reduce dimensionality and explore similarities in these data sets.

Two things matter: descriptors and ML method

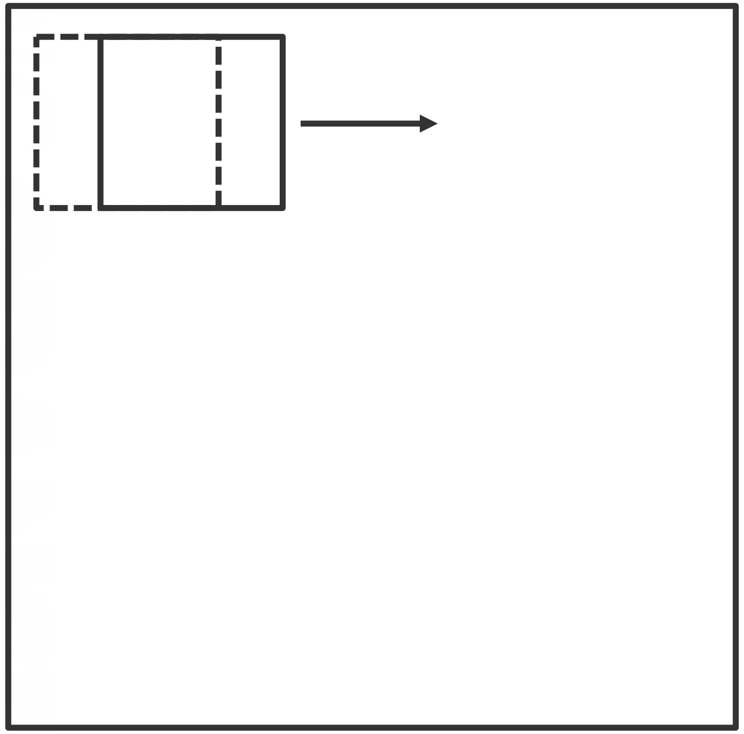
- In analysis of EELS or CITS data, very often our descriptor is the spectrum at each pixel. Typical analysis will be either linear or non-linear dimensionality reduction or clustering:
  - Linear dimensionality reduction: PCA, NMF, BLU
  - Clustering: k-means, GMM
  - Manifold learning: ISO, UMAP, tSNE, DBSCAN
  - Neural nets: SOFM, AEs, VAEs

- Typical result will be the components (representing behavior), and loading maps representing spatial variability of these behaviors. By construct, components will not depend on the relative spatial positions of pixel.

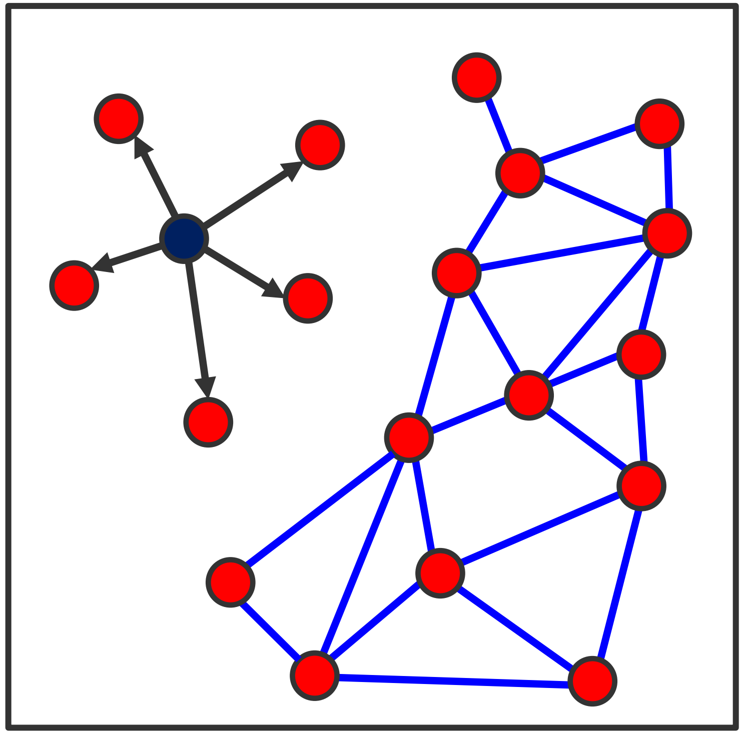
Describing the building blocks

- The classical physical descriptions (symmetry, etc) can be defined locally only in Bayesian sens
- We can argue that local descriptors are simple, if not necessarily known
- And the rules that guide their emergence are also simple, if not known

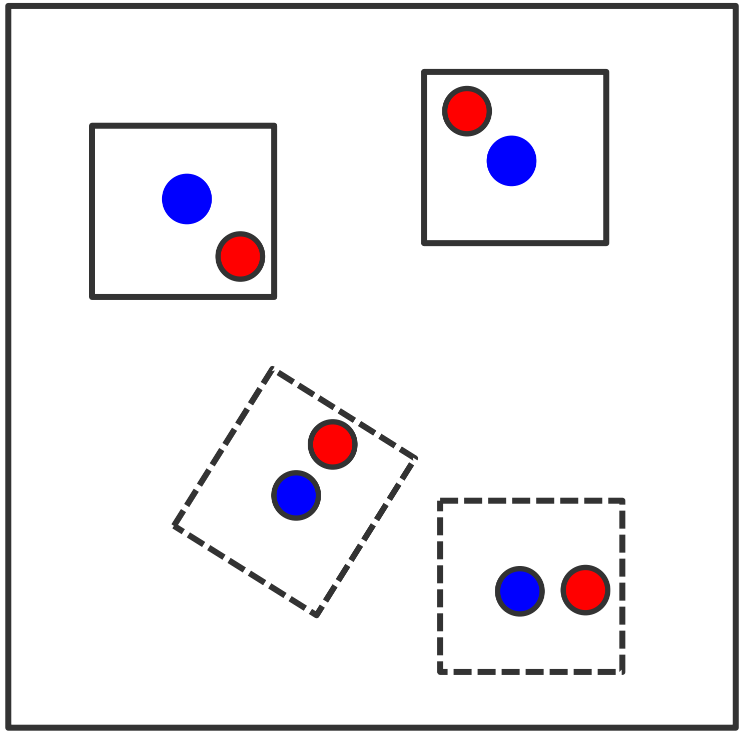
Continuous translational symmetry



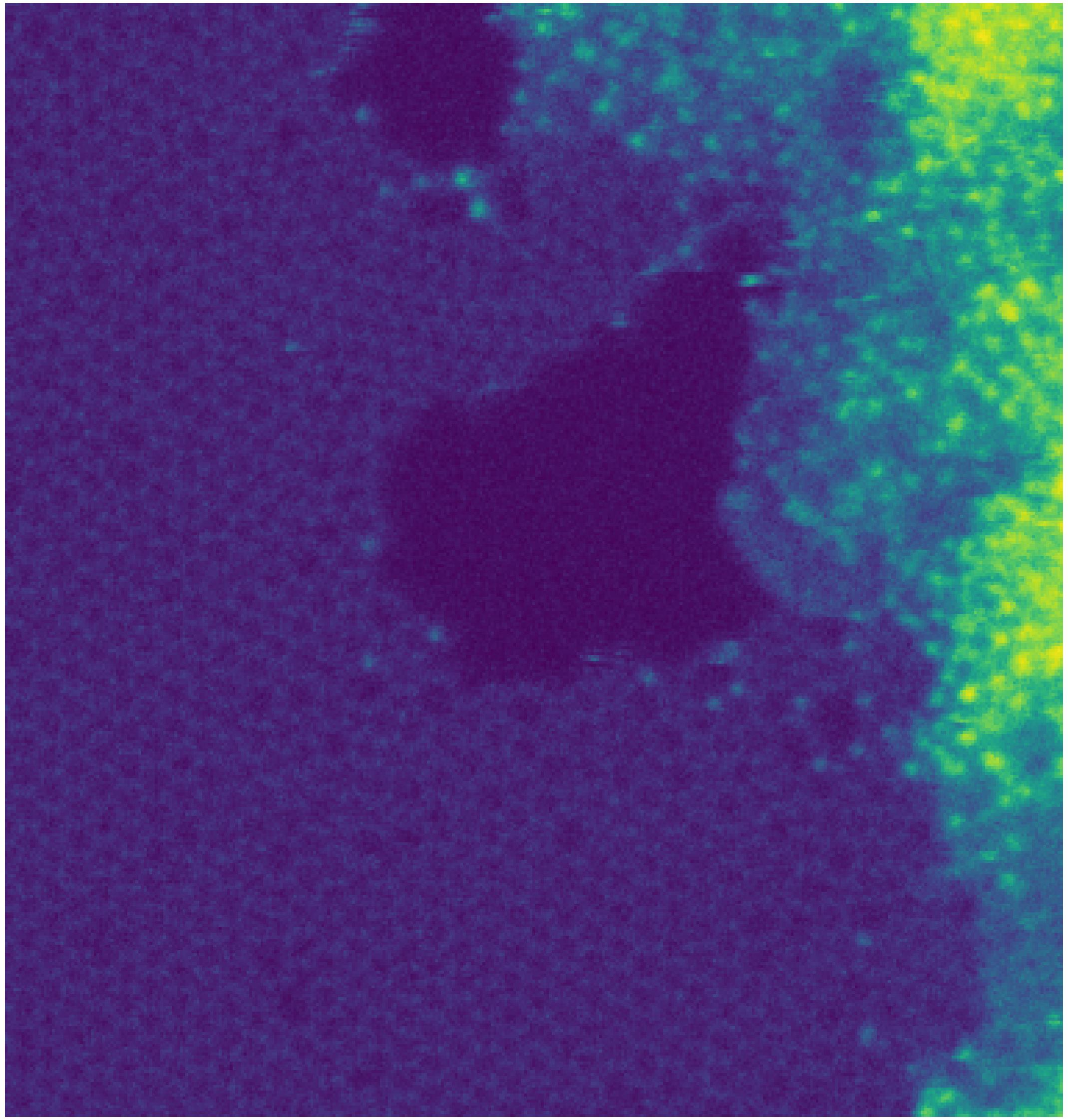
Atom based descriptions

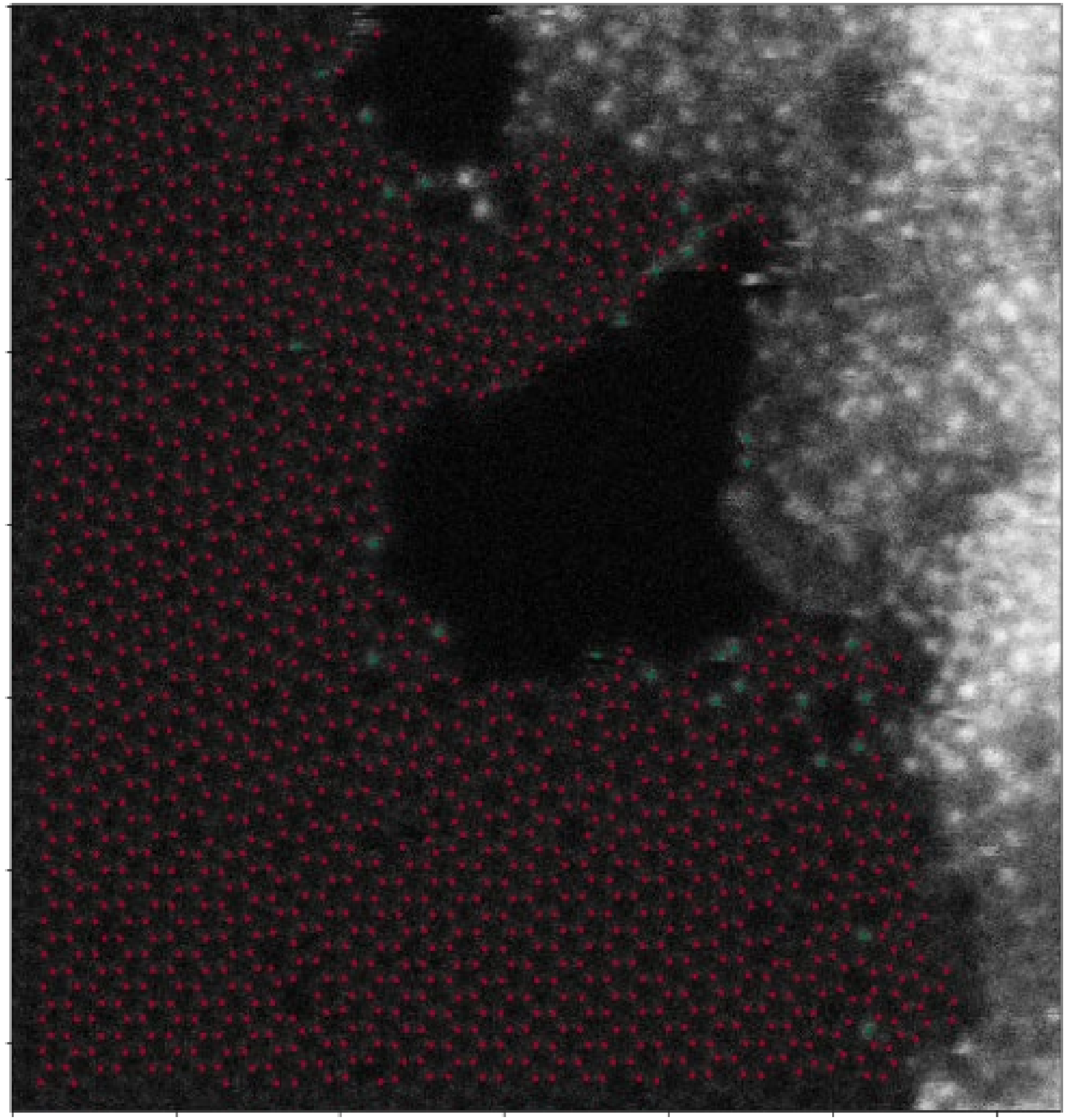


Localized sub-images

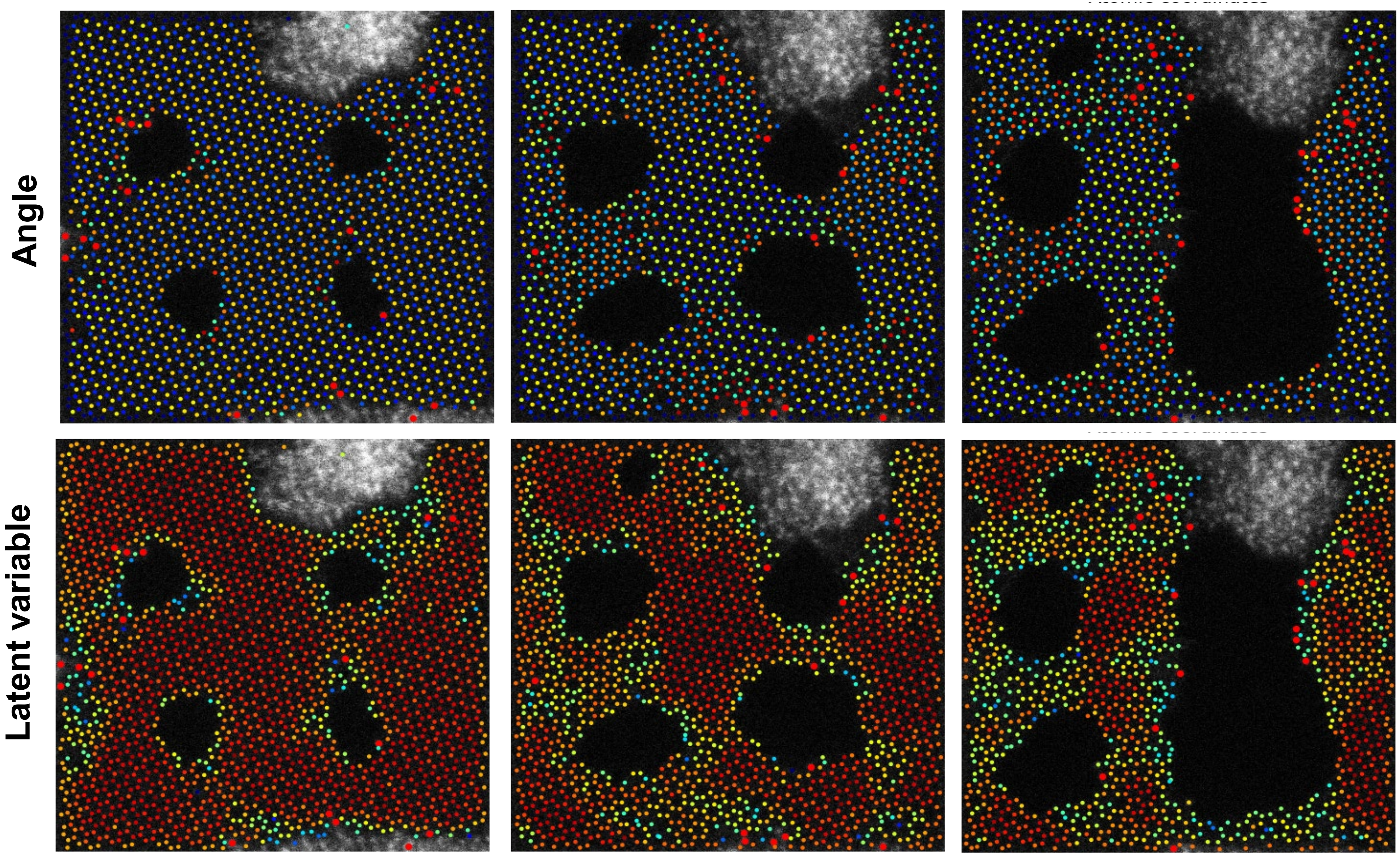


Off to chemically-disordered systems

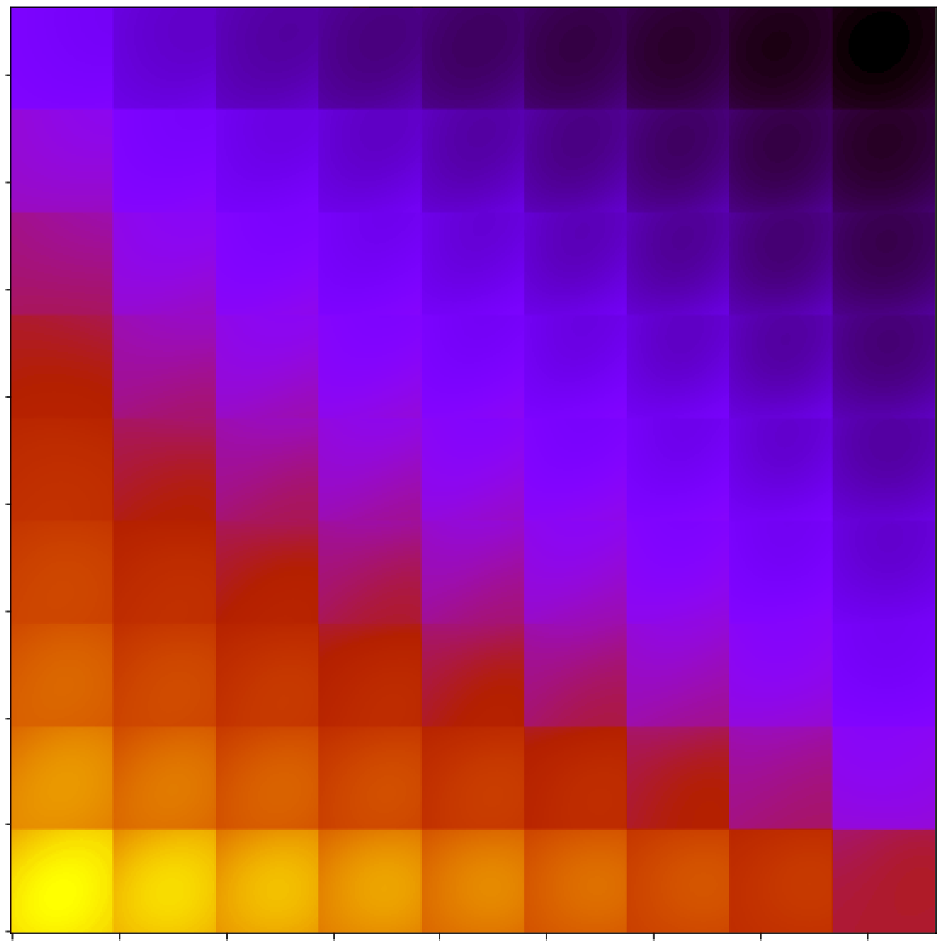


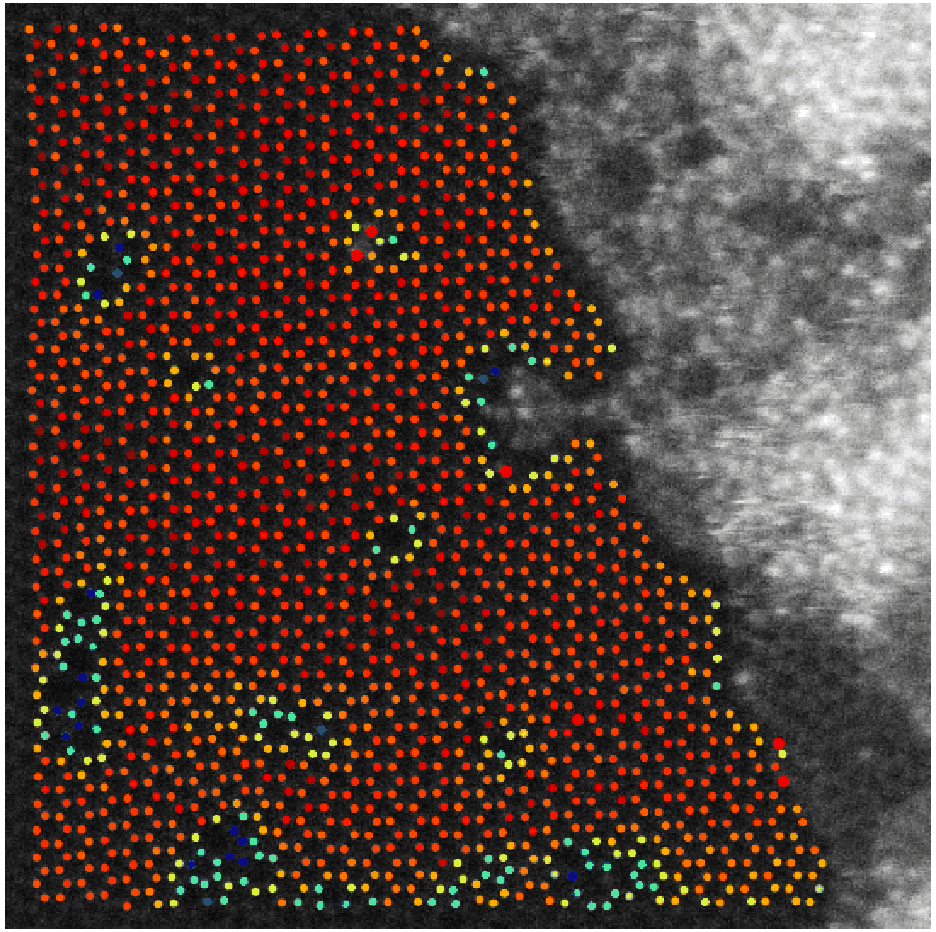


rVAE analysis at different time steps



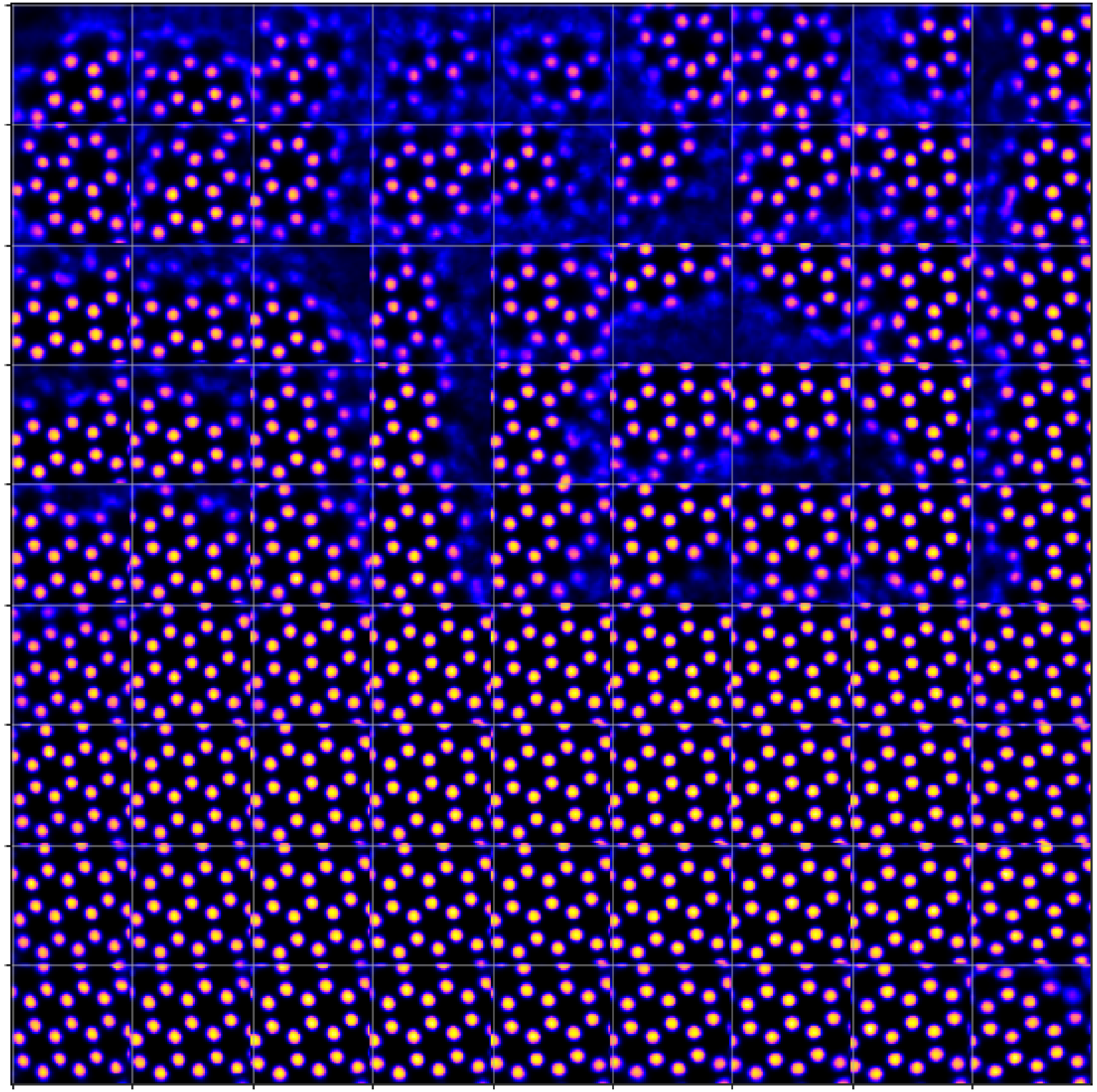
There is nothing as beautiful as training VAE



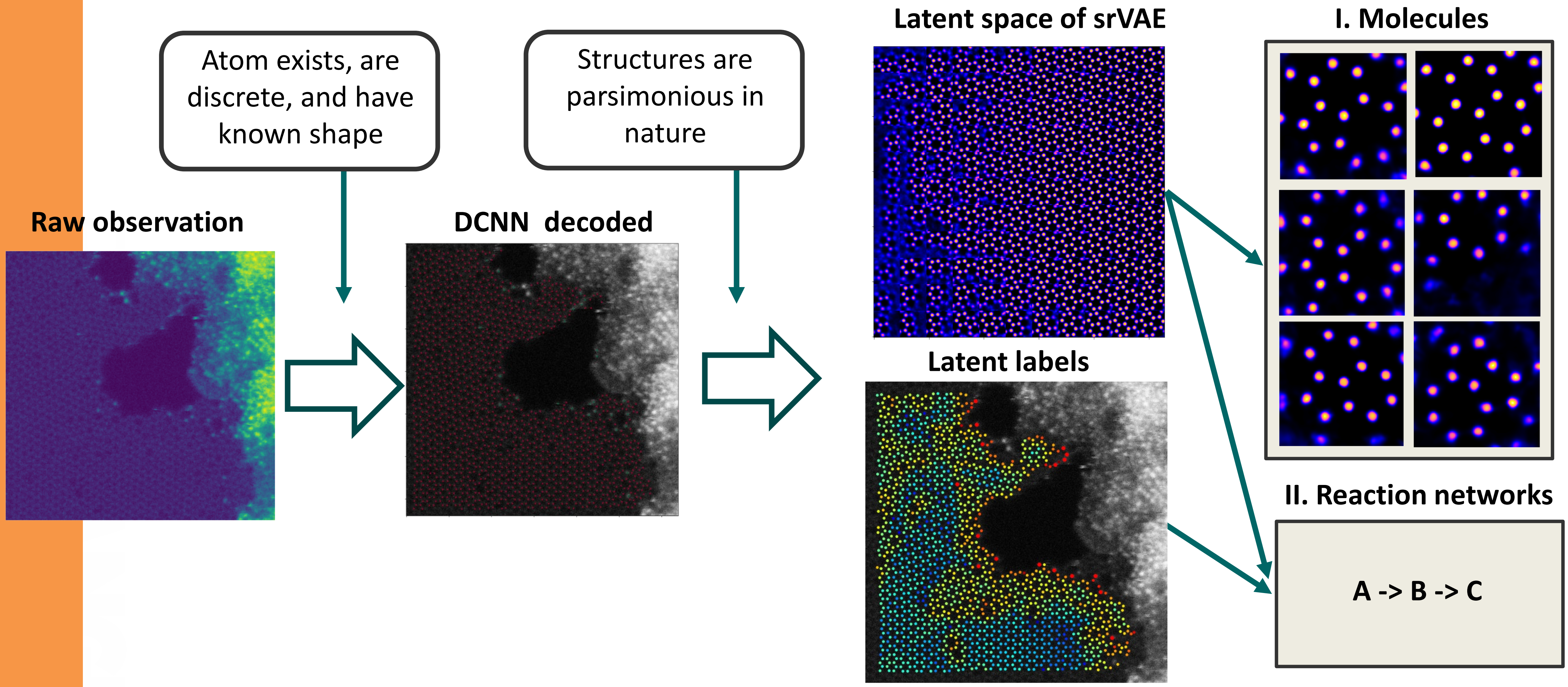


Next step: skip-rVAE

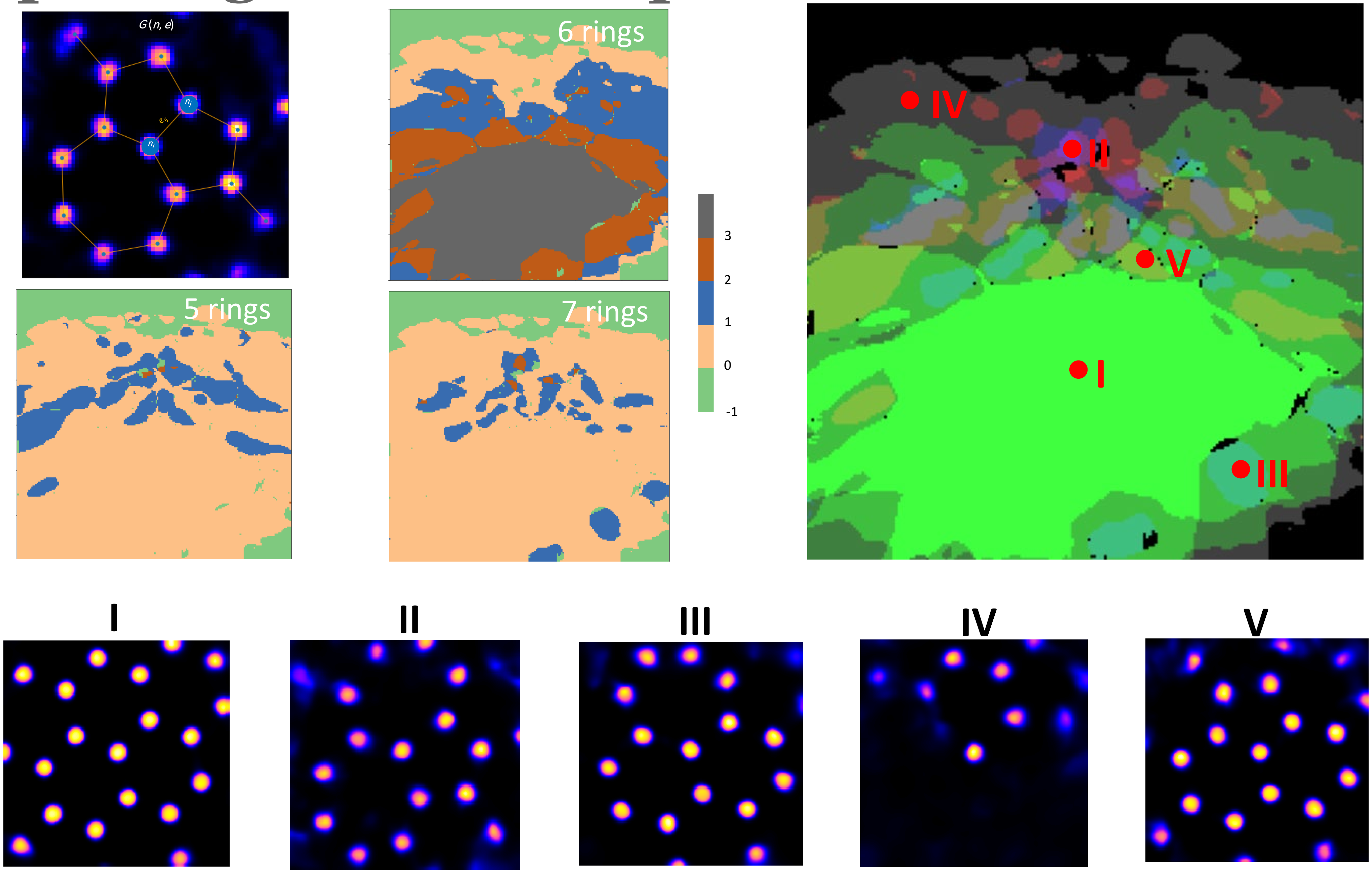
THE UNIVERSITY of TENNESSEE KNOXVILLE



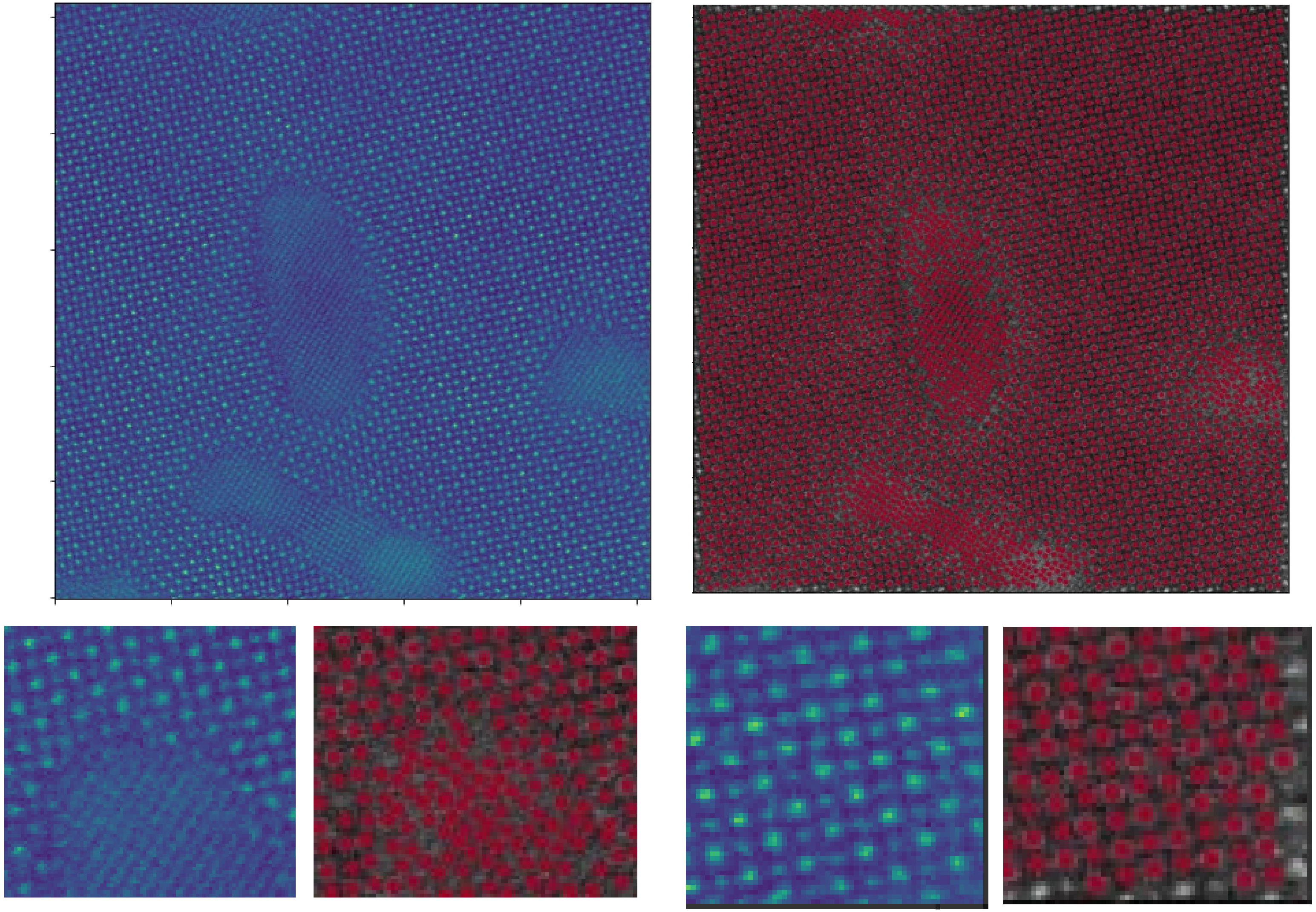
Unsupervised discovery of molecules



Exploring the latent space structure

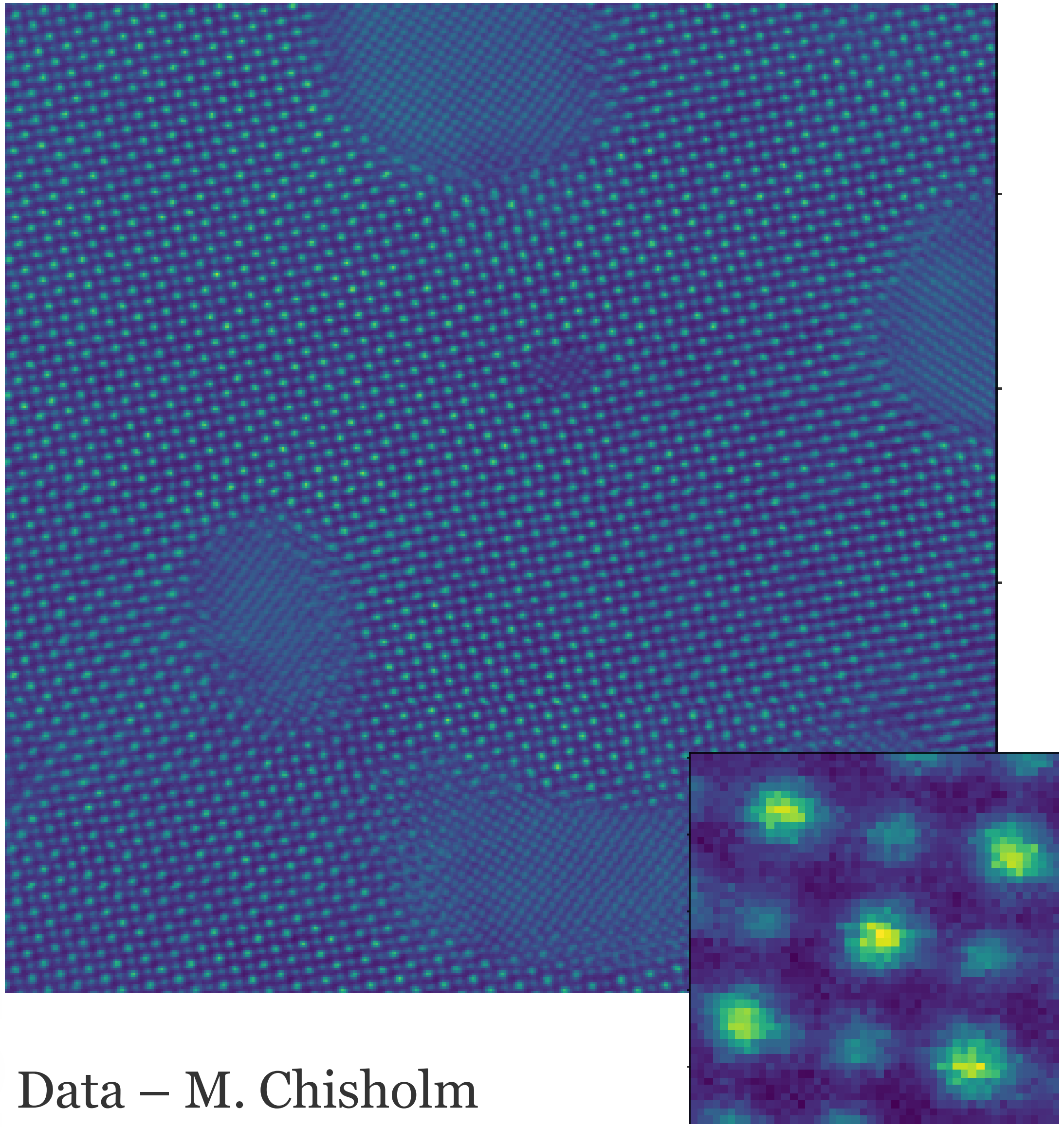


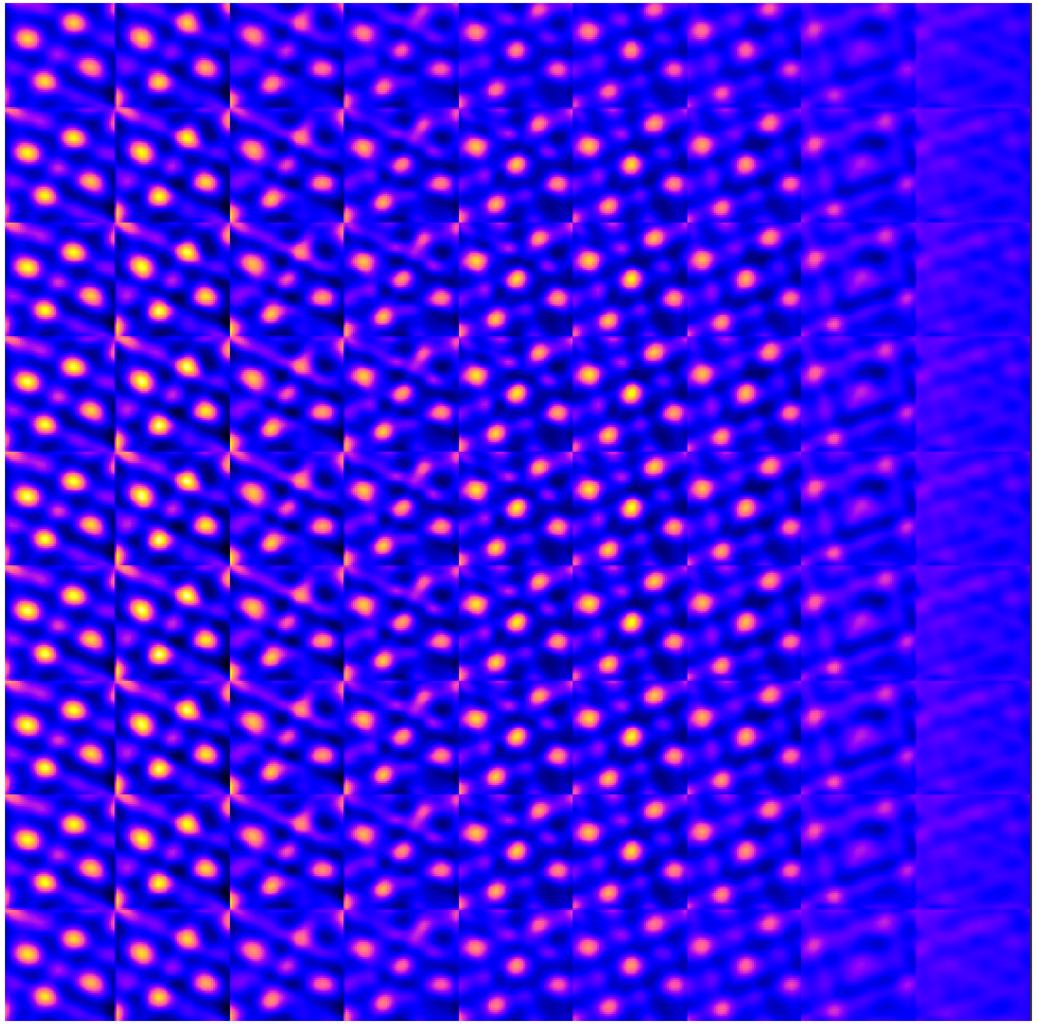
Let's put it all together!



Step 1: Find all atoms (or all that you can) – use maximum finders, blob-log, or DCNNs
Step 2: Create descriptors – patches centered on atoms. Keep track on what part of image (or stack) it has came from

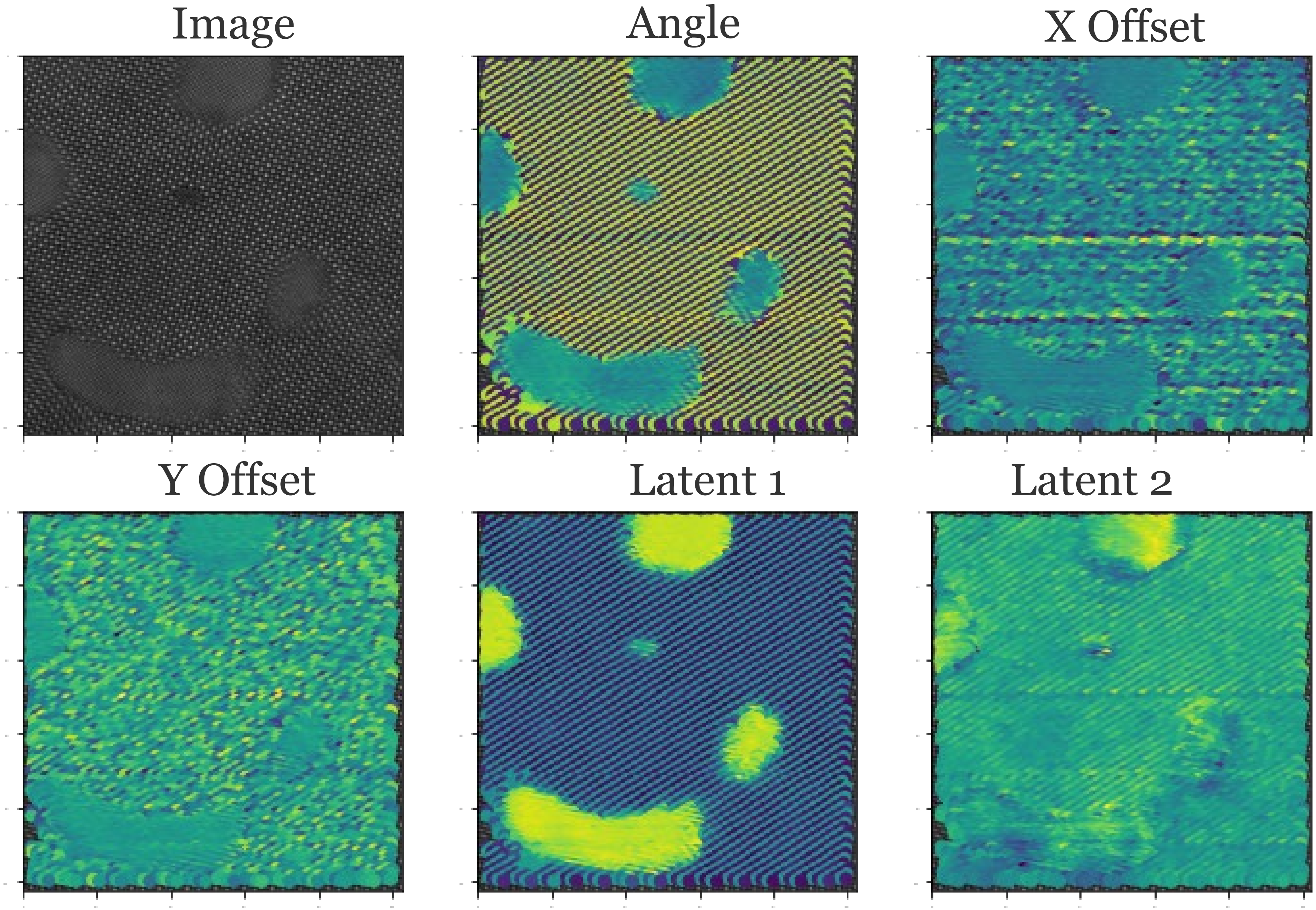
Analysis of the NiO-LSMO



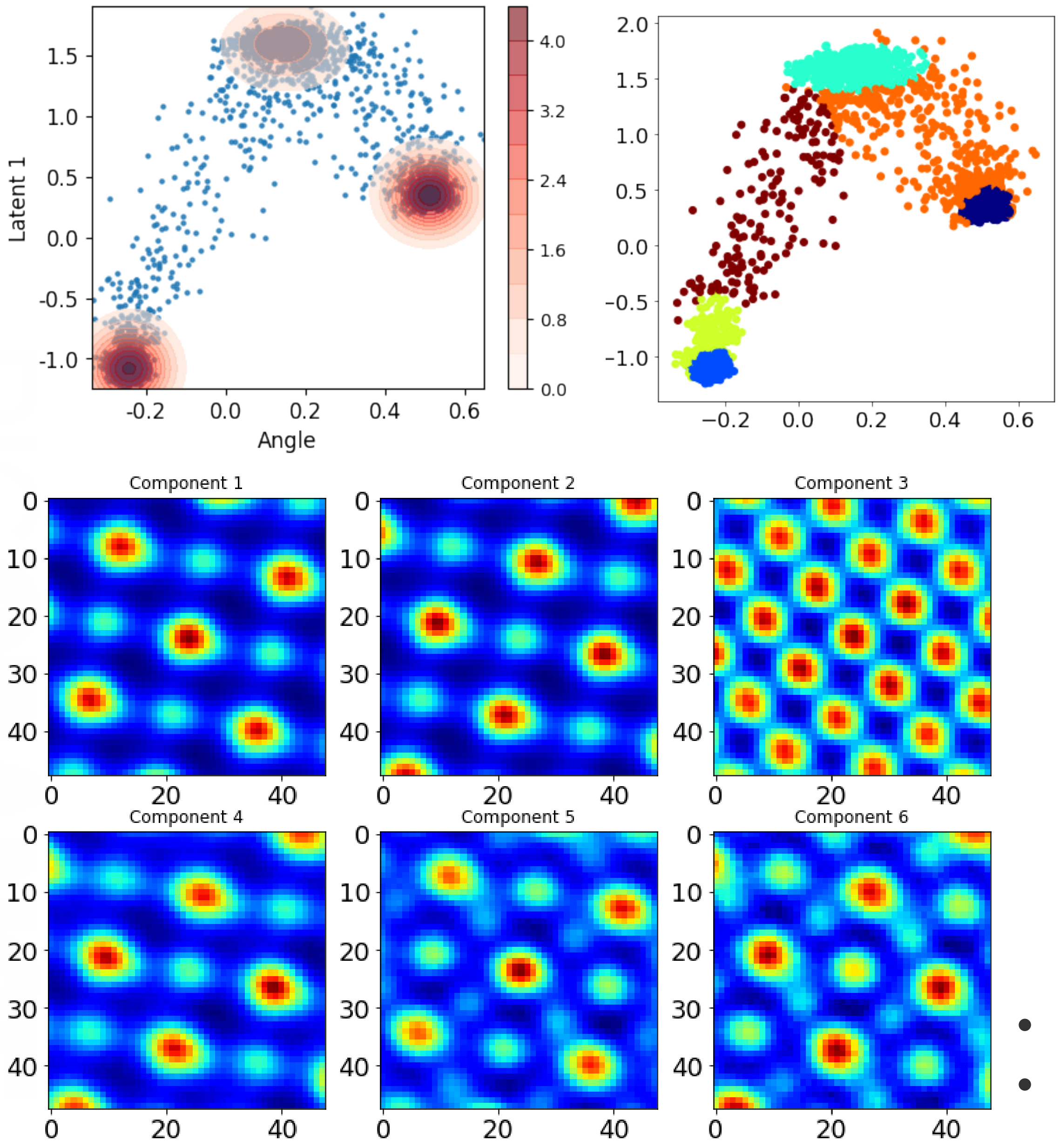


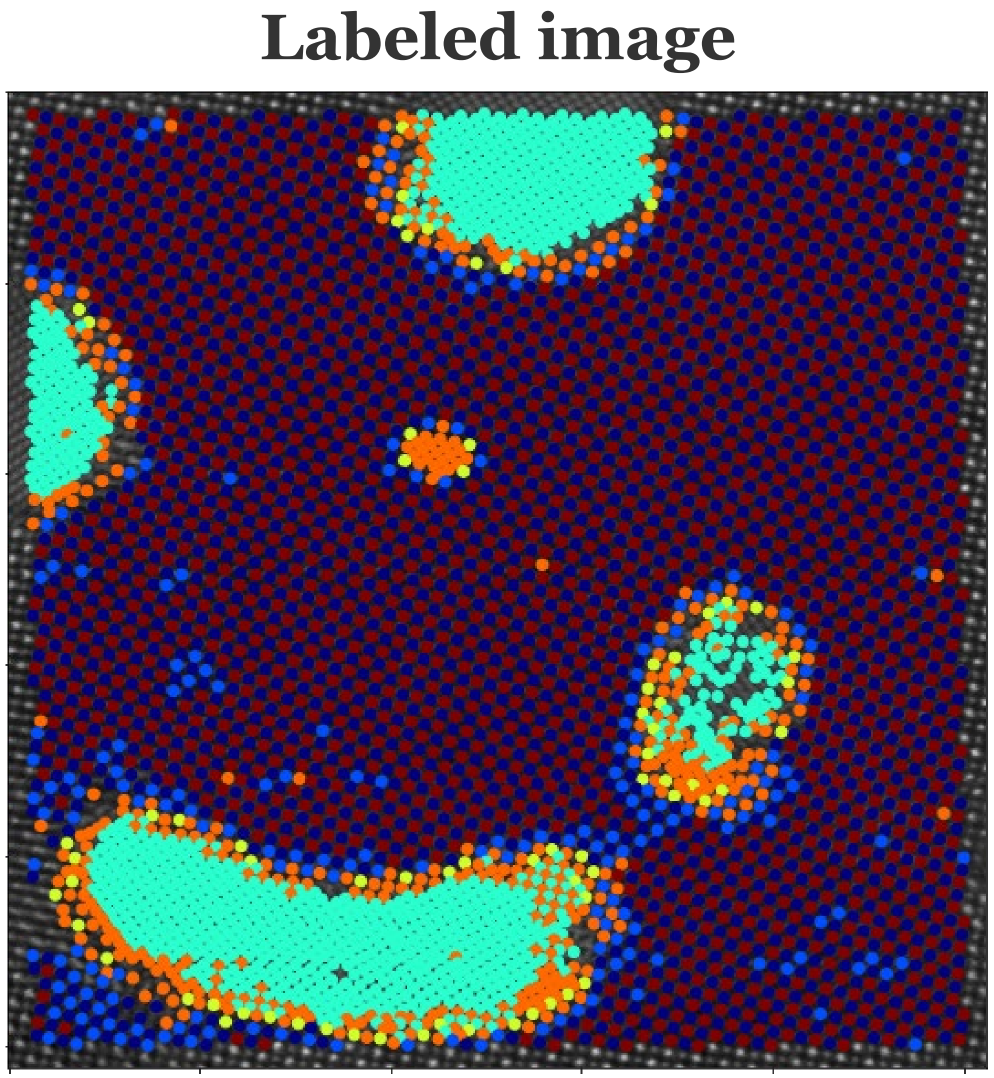
Data – M. Chisholm

Let's look at latent space



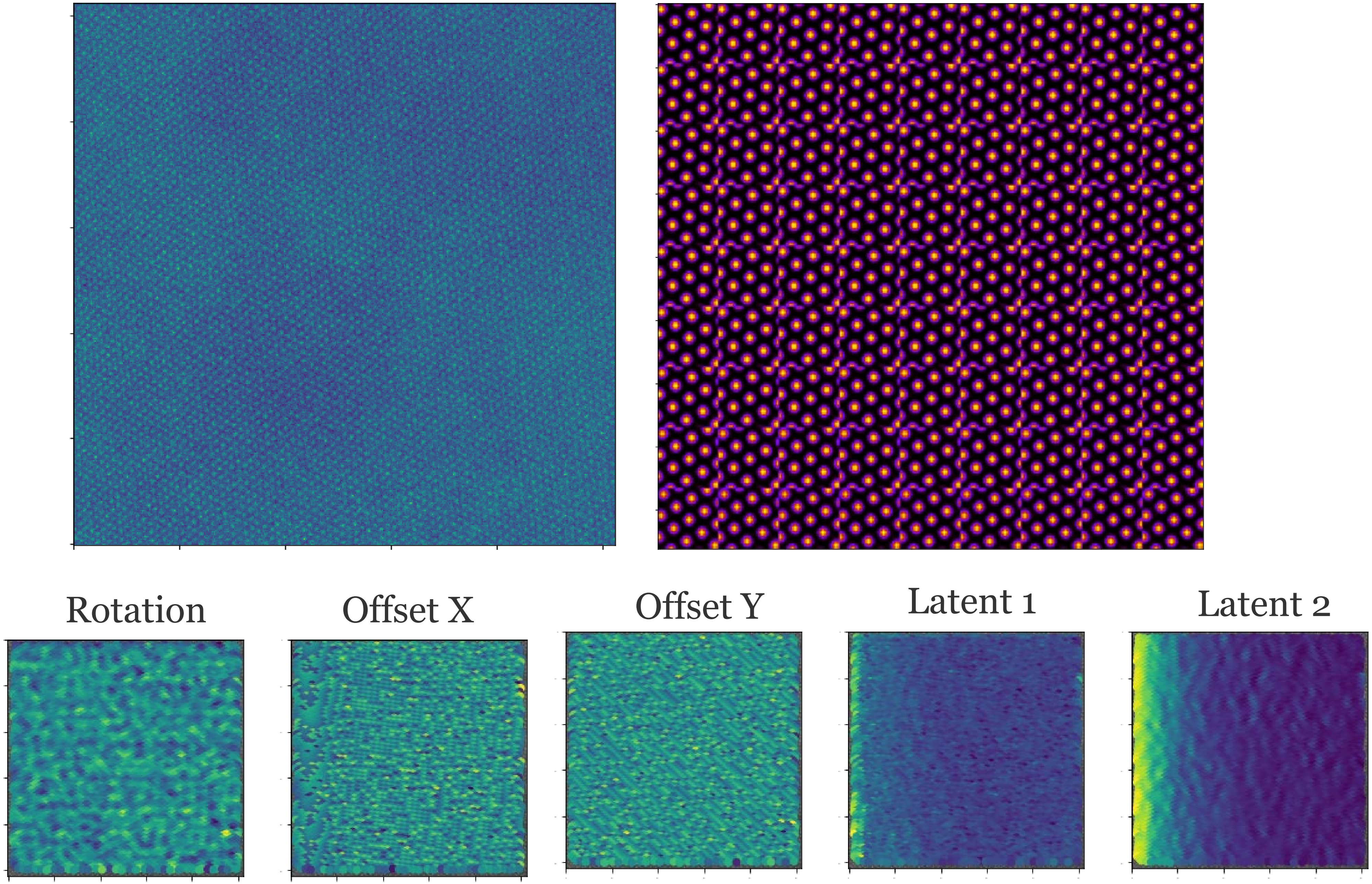
Exploring latent distributions





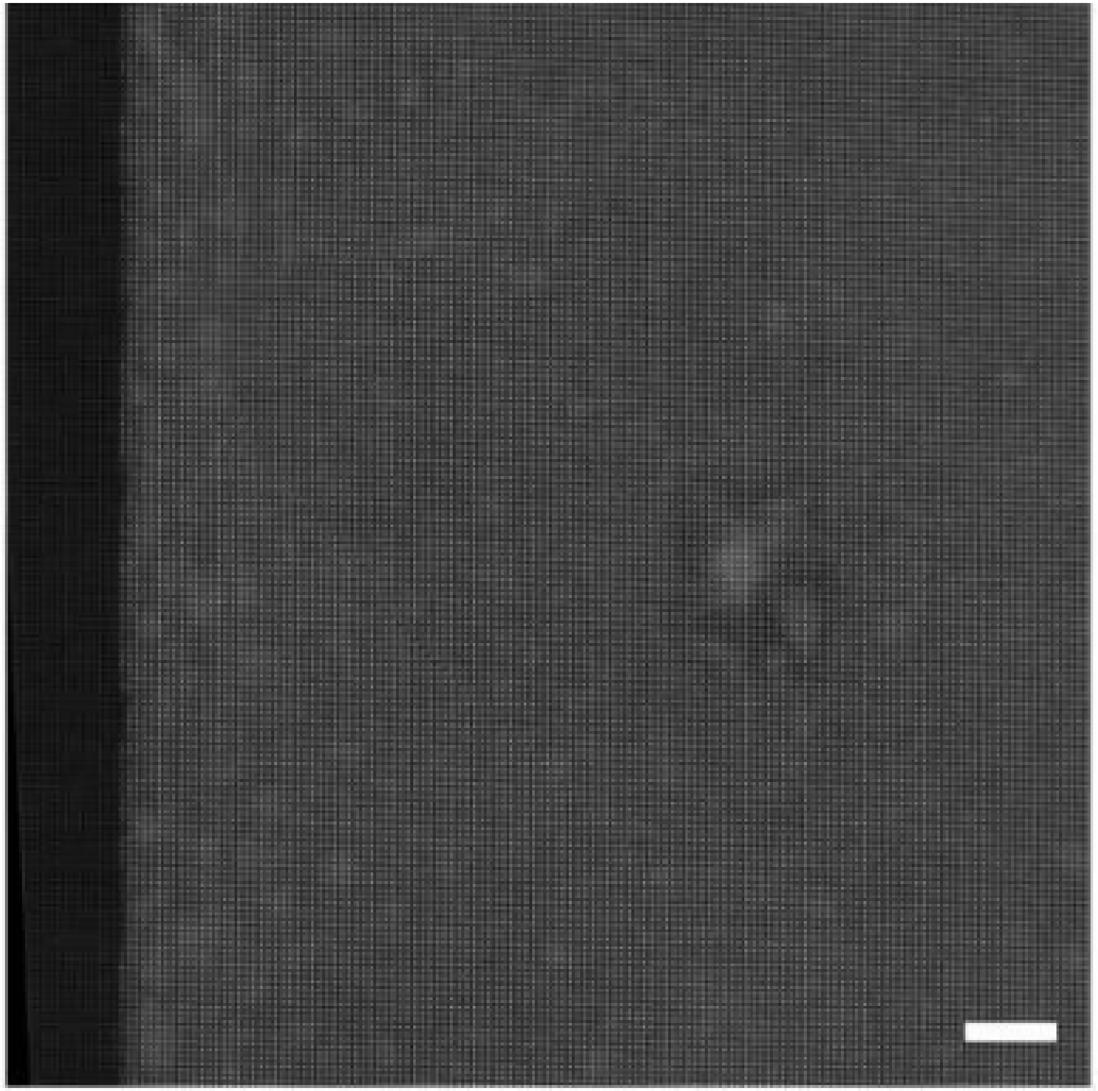
- Classes and variability are mixed in latent space
- Disentangling of representation

Out of curiosity: single crystal?

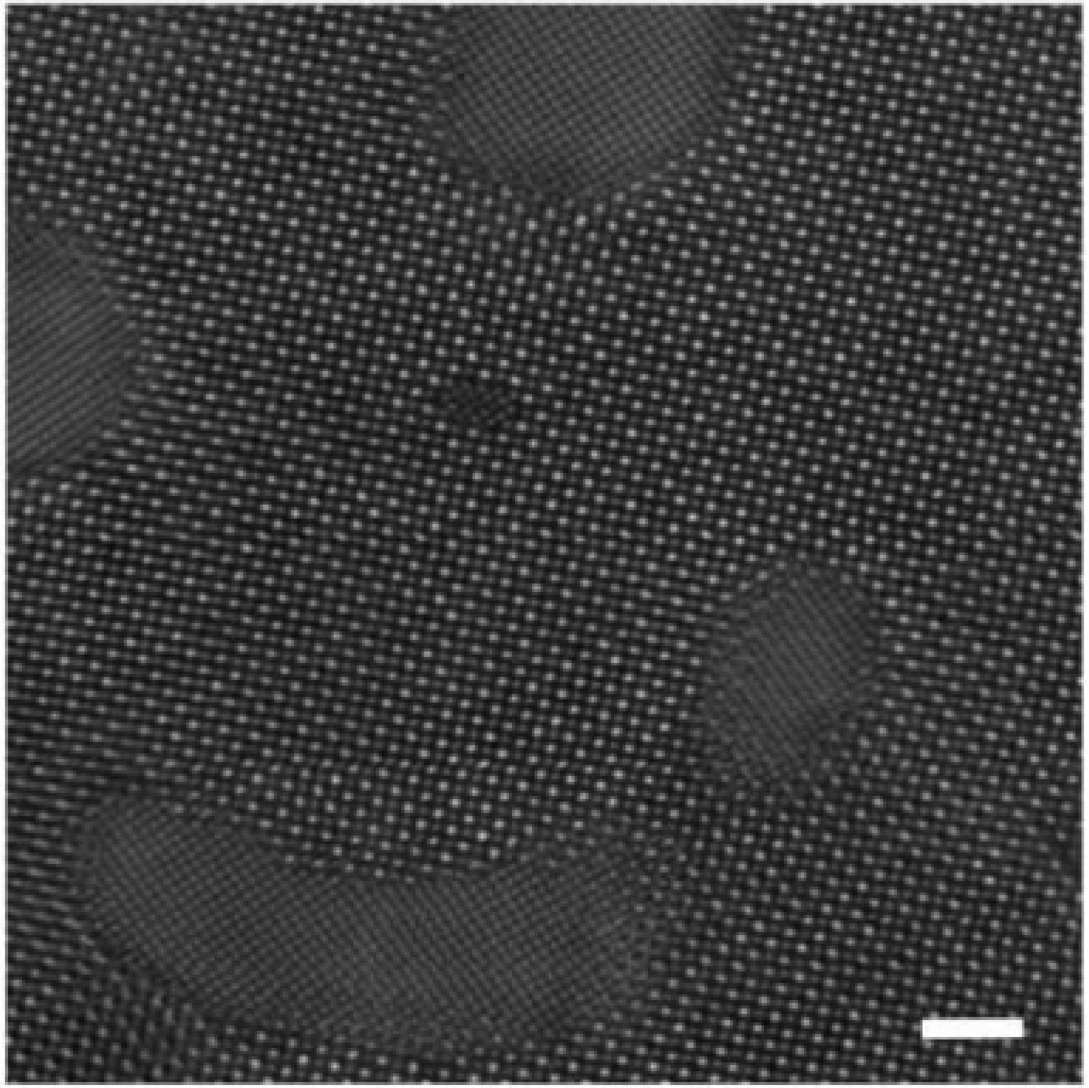


VAE without Atom Finding

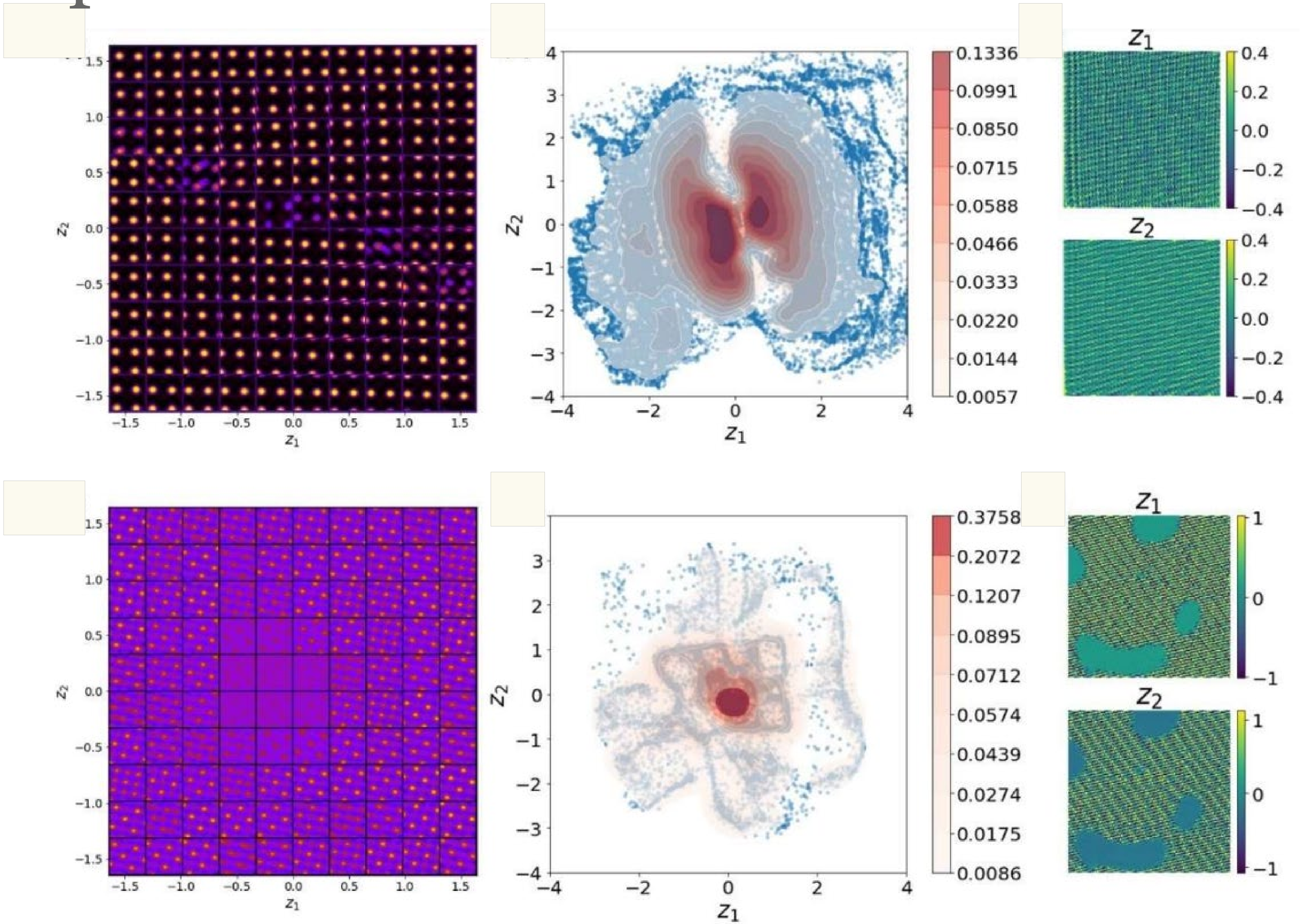
Ferroelectric $BiFeO_3$



$NiO – La_xSr_{1-x}MnO_3$



Simple VAE



0.1336
0.0991
0.0850
0.0715
0.0588
0.0466
0.0333
0.0220
0.0144
0.0057
0.3758
0.2072
0.1207
0.0895
0.0712
0.0574
0.0439
0.0274
0.0175
0.0086

Shift VAE: Translational Invariance