Supplementary Material: Quantifying and Learning Linear Symmetry-Based Disentanglement

Anonymous Author(s)
Affiliation
Address
`email`

A   Preliminaries: group theory

In this appendix, we summarize some concepts from group theory that are important to understand the main text of the paper. Group theory provides a useful language to formalize the notion of symmetry transformations and their effects. For a more elaborate discussion we refer the reader to the book from Hall (2015) on group theory.

Group   A group is a non-empty set G together with a binary operation $\circ : G \times G \to G$ that satisfies three properties:

1. *Associativity*: For all $f, g, h \in G$, it holds that $f \circ (g \circ h) = (f \circ g) \circ h$.

2. *Identity*: There exists a unique element $e \in G$ such that for all $g \in G$ it holds that $e \circ g = g \circ e = g$.

3. *Inverse*: For all $g \in G$ there exists an element $g^{-1} \in G$ such that $g^{-1} \circ g = g \circ g^{-1} = e$.

Direct product   Let G and G' be two groups. The *direct product*, denoted by $G \times G'$, is the group with elements $(g, g') \in G \times G'$ with $g \in G$ and $g' \in G'$, and the binary operation $\circ : G \times G' \to G \times G'$ such that $(g, g') \circ (h, h') = (g \circ h, g' \circ h')$.

Lie group   A Lie group is a group where G is a smooth manifold, this means it can be described in a local scale with a set of continuous parameters and that one can interpolate continuously between elements of G.

Group action   Let A be a set and G a group. The *group action* of G on A is a function $G_A : G \times A \to A$ that has the properties [1]

1. $G_A(e, x) = x$ for all $a \in A$

2. $G_A(g, (G_A(g', a)) = G_A(g \circ g', a)$ for all $g, g' \in G$ and $a \in A$

Regular action   The action of G on A is regular if for every pair of elements $a, a' \in A$ there exists a unique $g \in G$ such that $g \cdot a = a'$.

Group representation   A *group representation* of G in the vector space V is a function $\rho : G \to GL(V)$ (where $GL(V)$ is the general linear group on V) such that for all $g, g' \in G$ $\rho(g \circ g') = \rho(g) \circ \rho(g')$ and $\rho(e) = \mathbb{I}_V$, where $\mathbb{I}_V$ is the identity matrix.

---

[1]To avoid notational clutter, we write $G_A(g, a) = g \cdot a$ where the set A on which $g \in G$ acts can be inferred from the context.

Direct sum of representations   The direct sum of two representations $\rho_1 : G \to GL(V)$ in V and $\rho_2 : G \to GL(V')$ in V' is a group representation $\rho_1 \oplus \rho_2 : G \to GL(V \oplus V')$ over the direct sum $V \oplus V'$, defined for $v \in V$ and $v' \in V'$ as:

$$(\rho_1 \oplus \rho_2)(g) \cdot (v, v') = (\rho_1(g) \cdot v, \rho_2(g) \cdot v') \tag{1}$$

B   Linear Symmetry-Based Disentanglement: Detailed Description

Higgins et al. (2018) provide a formal definition of linear disentanglement that connects symmetry transformations affecting the real world (from which data is generated) to the internal representations of a model. In the main text, we provide a definition from the perspective of a group action on the data directly, but the original definition considers an extra conceptual world state as well. Here, we describe the original setting in more detail, and explain why we choose a more direct and practical version of the definition.

The definition assumes the following setting. W is the set of possible world states, with underlying symmetry transformations that are described by a group G and its action $\cdot : G \times W \to W$ on W. In particular, G can be decomposed as the direct product of K groups $G = G_1 \times \ldots \times G_K$. Data is obtained via an *observation* function $b : W \to X$ that maps world states to observations in a *data space X*. A model's internal representation of data is modeled with the *encoding* function $h : X \to Z$ that maps data to the *embedding space Z*. Together, the observation and the encoding constitute the model's internal representation of the real world $f : W \to Z$ with $f(w) = h \circ b(w)$. The definition for Linearly Symmetry-Based Disentangled (LSBD) representations then formalizes the requirement that a model's internal representation f should reflect and disentangle the transformation properties of the real world, and that the transformation properties of the model's internal representations should be linear.

The original definition considers G acting on W and involves the model's internal representation $f : W \to Z$, but since we do not directly observe W it is more practical to evaluate LSBD with respect to the encoding map $h : X \to Z$ instead. If the action of G on W is *regular*[2] and the observation map $b : W \to X$ is *injective*[3] though, we can instead define LSBD with respect to the action of G on X and the encoding map h, as shown in the main text.

C   Inner Product

To describe the norm $\| \cdot \|_{\rho,h,\mu}$ used in the definition of $\mathcal{D}_{\mathrm{LSBD}}$ we start with an arbitrary inner product $(\cdot, \cdot)$ on the linear latent space Z. Assume that ρ is linearly disentangled and accordingly splits in irreducible representations $\rho_k : G \to Z_k$ where $Z = Z_1 \oplus \cdots \oplus Z_K$ for some $K \in \mathbb{N}$. We will define a new inner product $\langle \cdot, \cdot \rangle_{\rho,h,\mu}$ on Z as follows. First of all we declare Z_k and Z_m to be orthogonal with respect to $\langle \cdot, \cdot \rangle_{\rho,h,\mu}$ if $k \neq m$. We denote by π_k the orthogonal projection on Z_k.

For $z, z' \in Z_i$, we set

$$\langle z, z' \rangle_{\rho,h,\mu} := \lambda_{k,h,\mu}^{-1} \int_{g \in G} (\rho(g) \cdot z, \rho(g) \cdot z') d\mathfrak{m}(g) \tag{2}$$

where $\mathfrak{m}$ is the (bi-invariant) Haar measure normalized such that $\mathfrak{m}(G) = 1$ and set

$$\lambda_{k,h,\mu} := \int_X \int_G \|\pi_k(h(x))\|^2 d\mathfrak{m}(g) d\mu(x) \tag{3}$$

if the integral on the right-hand side is strictly positive and otherwise we set $\lambda_k := 1$. This construction completely specifies the new inner product, and it has the following properties:

- the subspaces Z_k are mutually orthogonal,

- $\rho_k(g)$ is orthogonal on Z_k for every $g \in G$, in other words ρ_k maps to the orthogonal group on Z_k. Moreover, ρ maps to the orthogonal group on Z. This follows directly from the bi-invariance of the Haar measure and the definition of $\langle \cdot, \cdot \rangle_{\rho,h,\mu}$.

---

[2]This assumption holds in most practical cases with a suitable description of G.

[3]This is typically the case, but if not it can be solved through active sensing, see Soatto (2011).

- If π_k is the orthogonal projection to Z_k, then

$$\int_X \|\pi_k(h(x))\|_{\rho,h,\mu}^2 d\mu(x) = 1 \tag{4}$$

if the integral on the left is strictly positive.

For an arbitrary pair $z, z' \in Z$ the inner product $\langle \cdot, \cdot \rangle_{\rho,h,\mu}$ is given by

$$\langle z, z' \rangle_{\rho,h,\mu} = \sum_{k=1}^{K} \lambda_{k,h,\mu}^{-1} \int_{g \in G} (\rho(g) \cdot \pi_k(z), \rho(g) \cdot \pi_k(z')) d\mathfrak{m}(g) \tag{5}$$

D Evaluation of Equivariance by $\mathcal{D}_{\mathrm{LSBD}}$

We will now give an alternative expression for the disentanglement metric $\mathcal{D}_{\mathrm{LSBD}}$, since it will more visibly relate to the definition of equivariance. To avoid notational cluttering, in this section we will denote the norm $\|\cdot\|_{\rho,h,\mu}$ as $\|\cdot\|_*$. Let $\rho \in \mathcal{P}(G, Z)$ be a linear disentangled representation of G in Z. By expanding the inner product (or by using usual computation rules for expectations and variances), we first find that

$$\int_G \left\| \rho(g)^{-1} \cdot h(g \cdot x_0) - \int_G \rho(g')^{-1} \cdot h(g' \cdot x_0) d\nu(g') \right\|_*^2 d\nu(g)$$
$$= \int_G \|\rho(g)^{-1} \cdot h(g \cdot x_0)\|_*^2 d\nu(g) - \left\| \int_G \rho(g)^{-1} \cdot h(g \cdot x_0) d\nu(g) \right\|_*^2 \tag{6}$$
$$= \frac{1}{2} \int_G \int_G \|\rho(g)^{-1} \cdot h(g \cdot x_0) - \rho(g')^{-1} \cdot h(g' \cdot x_0)\|_*^2 d\nu(g) d\nu(g').$$

We now use that ρ maps to the orthogonal group for $(\cdot, \cdot)_*$, so that we can write the same expression as

$$\frac{1}{2} \int_G \int_G \|\rho(g \circ g'^{-1})^{-1} \cdot h(((g \circ g'^{-1}) \cdot g') \cdot x_0) - h(g' \cdot x_0)\|_*^2 d\nu(g) d\nu(g'). \tag{7}$$

This brings us to the alternative characterization of $\mathcal{D}_{\mathrm{LSBD}}$ as

$$\mathcal{D}_{\mathrm{LSBD}} = \inf_{\rho \in \mathcal{P}(G,Z)} \frac{1}{2} \int_G \int_G \|\rho(g \circ g'^{-1})^{-1} h(((g \circ g'^{-1}) \cdot g') \cdot x_0) - h(g' \cdot x_0)\|_*^2 d\nu(g) d\nu(g'). \tag{8}$$

In particular, if for every data point x there is a unique group element g_x such that $x = g_x \cdot x_0$, the disentanglement metric $\mathcal{D}_{\mathrm{LSBD}}$ can also be written as

$$\inf_{\rho \in \mathcal{P}(G,Z)} \frac{1}{2} \int_G \int_X \|\rho(g \circ g_x^{-1})^{-1} h((g \circ g_x^{-1}) \cdot x) - h(x)\|_*^2 d\nu(g) d\mu(x), \tag{9}$$

in which the equivariance condition appears prominently. The condition becomes even more apparent if ν is in fact the Haar measure itself, in which case the metric equals

$$\inf_{\rho \in \mathcal{P}(G,Z)} \frac{1}{2} \int_G \int_X \|\rho(g)^{-1} \circ h(g \cdot x) - h(x)\|_*^2 d\mathfrak{m}(g) d\mu(x). \tag{10}$$

E Datasets

All datasets contain 64×64 pixel images. The Square, Arrow and Airplane datasets have a known group decomposition $G = \mathrm{SO}(2) \times \mathrm{SO}(2)$ describing the underlying transformations. In these three datasets, for each subgroup a fixed number of $|\mathcal{G}_k| = 64$ with $k \in \{1, 2\}$ transformations is selected. Each image is generated from a single initial data point upon which all possible group actions are applied, resulting in datasets with $|\mathcal{G}_1| \cdot |\mathcal{G}_2| = 4096$ images. The datasets exemplify different group actions of $\mathrm{SO}(2)$: periodic translations, in-plane rotations, out-of-plane rotations, and periodic hue-shifts, see Figure 1.

The ModelNet40 and the COIL-100 datasets consist of different objects rotating with respect to a vertical axis (out-of-plane rotation). For these datasets the group $G = \mathrm{SO}(2)$ describes the underlying transformations that each object undergoes, see Figure 1. The different objects can be seen as non-symmetric variability in the data. In this particular case, each object has its own base-point x_0 from which data is generated. The metric $\mathcal{D}_{\mathrm{LSBD}}$ is then evaluated per object instance for the group $G = \mathrm{SO}(2)$, the value of $\mathcal{D}_{\mathrm{LSBD}}$ is calculated and averaged across all available objects.

3

(a) Square	(b) Arrow	(c) Airplane

|	
(d) ModelNet40	(e) COIL-100
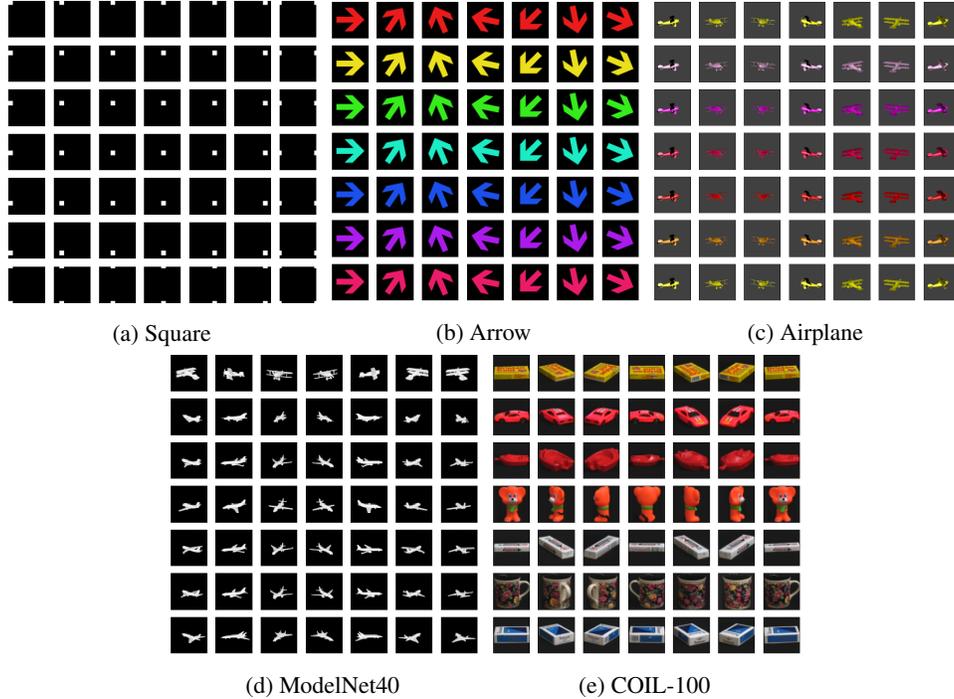
Figure 1: Example images from each of the datasets used. Each image corresponds to an example data point for a combination of two factors, e.g. color and orientation. The factors change horizontally and vertically. The boundaries for the Square, Arrow and Airplane dataset are periodic. For the ModelNet40 and COIL-100 dataset, the vertical direction represents different object instances and the horizontal direction represents the rotation of the corresponding object.

Square   This dataset consists of a set of images of a black background with a square of 4×4 white pixels. The dataset is generated applying vertical and horizontal translations of the white square considering periodic boundaries.

Arrow   This dataset consists of a set of images depicting a colored arrow at a given orientation. The dataset is generated by applying cyclic shifts of its color and in-plane rotations. The cyclic color shifts were obtained by preselecting a fixed set of 64 colors from a circular hue axis. The in-plane rotations were obtained by rotating the arrow along an axis perpendicular to the picture plane over 64 predefined positions.

Airplane   This dataset consists of renders obtained using Blender v2.7 (Community, 2020) from a 3D model of an airplane within the ModelNet40 dataset (Wu et al., 2014) (this dataset is provided for the convenience of academic research only). We created each image by varying two properties: the airplane's color and its orientation with respect to the render camera. The orientation was changed via rotation with respect to a vertical axis (out-of-plane rotation). The colors of the model were selected from a predefined cyclic set of colors similar to the arrow rotation dataset.

ModelNet40   This dataset also consists of a dataset of renders obtained using Blender v2.7 (Community, 2020) from all the 3D models within the airplane category of the ModelNet40 dataset (Wu et al., 2014). We created each image by varying each airplane's orientation with respect to the render camera, via rotation with respect to a vertical axis (out-of-plane rotation). In this case we used 64 orientations for each object, i.e. $|\mathcal{G}| = 64$.

COIL-100   This dataset (Nene et al., 1996) consists of images from 100 objects placed on a turntable against a black background. For each object, 72 views of the rotated object are provided. The original images have a resolution of 128×128 and were re-scaled to 64×64 to match our other datasets. In

4

119 this case for each object $|\mathcal{G}| = 72$. This dataset is intended for non-commercial research purposes
120 only. This dataset was obtained by using TensorFlowDatasets (2021).

F Experimental Settings and Hyperparameters

F.1 Architectures

123 Table 1 shows the encoder and decoder architectures used for almost all methods and datasets. The
124 encoder's last layer depends on the method. For VAE, cc-VAE, FactorVAE, DIP-I, DIP-II, two dense
125 layers with 4 units each were used. For LSBD-VAE and ∆VAE two dense layers with 4 and 2 units
126 each were used. For Quessard a single dense layer with 4 units was used. The only model that was
127 not trained with this architectures was LSBD-VAE/0 method for the ModelNet40 dataset the reason
128 for this choice was that during training the loss was getting NaN values, in this case the architecture
129 used was that of Table 2.

Table 1: Encoder and decoder architectures used in most methods.

ENCODER	
INPUT	SIZE (64,64, NUMBER CHANNELS)
CONV	FILTERS 32, KERNEL 4, STRIDE 2, ReLU
CONV	FILTERS 32, KERNEL 4, STRIDE 2, ReLU
CONV	FILTERS 64, KERNEL 4, STRIDE 2, ReLU
CONV	FILTERS 64, KERNEL 4, STRIDE 2, ReLU
DENSE	UNITS 256, ReLU
DENSE(X2)	UNITS DEPEND ON METHOD
DECODER	
INPUT	SIZE (NUMBER OF LATENT DIMENSIONS)
DENSE	UNITS 256, ReLU
DENSE	UNITS 4*4*64, ReLU
RESHAPE	(4,4,64)
CONVT	FILTERS 64, KERNEL 4, STRIDE 2, ReLU
CONVT	FILTERS 32, KERNEL 4, STRIDE 2, ReLU
CONVT	FILTERS 32, KERNEL 4, STRIDE 2, ReLU
CONVT	FILTERS (NUMBER CHANNELS), KERNEL 4, STRIDE 2, SIGMOID

Table 2: Encoder and decoder architecture used to train LSBD-VAE/0 for ModelNet40 dataset.

ENCODER	
INPUT	SIZE (64, 64, NUMBER CHANNELS)
DENSE	UNITS 512, ReLU, BATCH NORMALIZATION
DENSE	UNITS 256, ReLU, BATCH NORMALIZATION
DENSE(X2)	UNITS DEPEND ON METHOD
DECODER	
INPUT	SIZE (NUMBER OF LATENT DIMENSIONS)
DENSE	UNITS 256, ReLU, BATCH NORMALIZATION
DENSE	UNITS 512, ReLU, BATCH NORMALIZATION
DENSE	UNITS 64*64*NUMBER OF CHANNELS, SIGMOID
RESHAPE	(64, 64, NUMBER OF CHANNELS)

F.2 Hyperparameters

131 Table 3 shows the hyperparameters used to train each model for all datasets. Table 4 shows the
132 hyperparameters used to train the LSBD-VAE models for each dataset. In the latter case, the number
133 of epochs for the LSBD-VAE model were increased. The range of values used for the scale parameter
134 t were increased for ModelNet40 and COIL-100 datasets since it was noticed that this provided

5

better results in terms of data reconstruction and disentanglement. For the Arrow dataset, a value of $\gamma = 1$ was producing unstable results. However, the values 10, 100, 1000 or even 10000 were producing good results without significant changes among them. Therefore the value 100 was used for the datasets with the same structure (Square, Arrow and Airplane). For the ModelNet40 and COIL-100 the experiments showed that this hyperparameter for values as high as 10000 could affect the reconstructions, thus a lower value $\gamma = 1$ was chosen.

Table 3: Model hyperparameters for all datasets

MODEL	PARAMETERS
VAE	TRAINING STEPS 30000
β-VAE	$\beta = 5$, TRAINING STEPS 30000
CC-VAE	$\beta = 5, \gamma = 1000, c_{max} = 15$, ITERATION THRESHOLD 3500, TRAINING STEPS 30000
FACTOR	$\gamma = 1$, EPOCHS 30000
DIP-I	$\lambda_{od} = 1, \lambda_d = 10$, TRAINING STEPS 30000
DIP-II	$\lambda_{od} = 1, \lambda_d = 1$, TRAINING STEPS 30000
QUESSARD	$\lambda = 0.01$, TRAJECTORIES 3000

Table 4: LSBD-VAE hyperparameters for all datasets

DATASETS	PARAMETERS
SQUARE, ARROW, AIRPLANE	$t \in [10^{-10}, 10^{-9}], \gamma = 100.0$, EPOCHS 1500
MODELNET40	$t \in [10^{-10}, 10^{-5}], \gamma = 1.0$, EPOCHS 1500
COIL-100	$t \in [10^{-10}, 10^{-5}], \gamma = 1.0$, EPOCHS 6000

F.3 Hardware & Running Time

The hardware used across all experiments was a DGX station with 4 NVIDIA GPUs V100 and 32GB . Only one GPU was used per experiment. The running time for the LSBD-VAE across all 9 degrees of supervision $L \in \{0, 256, 768, 1024, 1280, 1536, 1792, 2048\}$ and all 10 runs (total $9 \cdot 10$ repetitions) for the datasets were: Arrow 33 ± 4 minutes Airplane 29 ± 4 minutes and Square 28 ± 4 minutes. The running time for the LSBD-VAE across 2 degrees of supervision and 10 runs (total $2 \cdot 10$ repetitions) for ModelNet40 was 136 ± 10 minutes and for COIL-100 90 ± 6 minutes. For the method from (Quessard et al., 2020) the training times were approximately 30 minutes across all datasets. The training times for the methods from `disentanglement_lib` (Locatello et al., 2019) were not measured.

F.4 Code Licenses

The `disentanglement_lib` (Locatello et al., 2019) code is registered with an Apache 2.0 License while the code used to reproduce the method by Quessard et al. (2020) is registered with an MIT license.

G Full results

The full results for all experiments on all datasets are given in Tables 5, 6, 7, 8, and 9. We report the mean and standard deviation over 10 runs for each experiment.

G.1 Limited Supervision Suffices to Learn LSBD Representations

The results obtained from Tables 5, 6, 7 show that we do not need transformation-labels for all data points, only a subset of labeled pairs is sufficient to learn LSBD representations. To further highlight this, Figure 2 shows $\mathcal{D}_{\text{LSBD}}$ scores for LSBD-VAE trained on the Square, Arrow, and Airplane datasets respectively, for various values for the number of labeled pairs L. For each L and each dataset, we trained 10 models so we can report box plots of the $\mathcal{D}_{\text{LSBD}}$ scores.
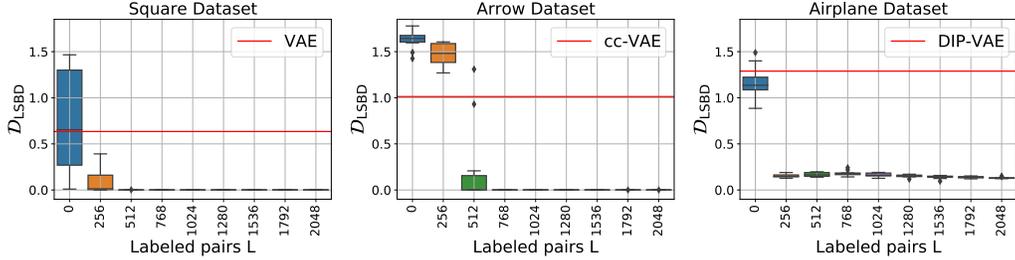
Figure 2: Box plots for $\mathcal{D}_{\mathrm{LSBD}}$ scores over 10 training repetitions for different numbers of labeled pairs L, for all datasets. The red line indicates the best-performing traditional disentanglement method.

For low values of L we see worse scores and high variability. But for slightly higher L, scores are consistently good, starting already at $L = 512$ for the Square, $L = 768$ for the Arrow, and $L = 256$ for the Airplane. This corresponds to respectively 25%, 37.5%, and 12.5% of the data being involved in a labeled pair. Moreover, we see that with just a little supervision we outperform the best traditional method on $\mathcal{D}_{\mathrm{LSBD}}$. Overall, these results suggest that with some expert knowledge (about the underlying group and a suitable representation) and limited annotation of transformations, LSBD can be achieved.

G.2 Quessard Arrow

In the main text we mentioned that we did not reproduce good results with Quessard et al. (2020)'s method on the Arrow and Square dataset. We highlight a particular case for the Arrow dataset, where the method clearly learns the rotations of the arrow but fails to learn color. Figure 3 shows reconstructed Arrow images. Since color isn't learned well, this example doesn't get a good $\mathcal{D}_{\mathrm{LSBD}}$ score, even though rotation is properly linearly disentangled.

(a) Input   (b) Reconstructions
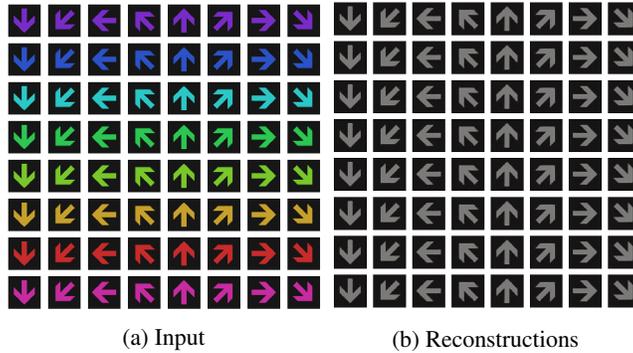
Figure 3: Results from Quessard et al. (2020)'s method on the Arrow dataset

References

Community, B. O. (2020). *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam.

Hall, B. C. (2015). *Lie Groups, Lie Algebras, and Representations*, volume 222 of *Graduate Texts in Mathematics*. Springer International Publishing, Cham.

Higgins, I., Amos, D., Pfau, D., Racaniere, S., Matthey, L., Rezende, D., and Lerchner, A. (2018). Towards a definition of disentangled representations. *arXiv preprint arXiv:1812.02230*.

Locatello, F., Bauer, S., Lucic, M., Gelly, S., Schölkopf, B., and Bachem, O. (2019). Challenging common assumptions in the unsupervised learning of disentangled representations. In *International Conference on Machine Learning*.

187 Nene, S. A., Nayar, S. K., Murase, H., et al. (1996). Columbia object image library (coil-20).

188 Quessard, R., Barrett, T. D., and Clements, W. R. (2020). Learning Group Structure and Disentangled
189 Representations of Dynamical Environments. *Advances in Neural Information Processing Systems*,
190 33.

191 Soatto, S. (2011). Steps Towards a Theory of Visual Information: Active Perception, Signal-
192 to-Symbol Conversion and the Interplay Between Sensing and Control. *arXiv preprint*
193 *arXiv:1110.2053*.

194 TensorFlowDatasets, . (2021). TensorFlow Datasets, a collection of ready-to-use datasets. `https:`
195 `//www.tensorflow.org/datasets`.

196 Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J. (2014). 3D ShapeNets: A
197 Deep Representation for Volumetric Shapes.

Table 5: Scores for the Square dataset.

MODEL	BETA ↑	FACTOR ↑	SAP ↑	DCI ↑	MIG ↑	MOD ↑	$\mathcal{D}_{\text{LSBD}}$ ↓
VAE	$.945_{\pm.061}$	$.835_{\pm.140}$	$.019_{\pm.004}$	$.009_{\pm.005}$	$.013_{\pm.004}$	$.579_{\pm.202}$	$.634_{\pm.440}$
β-VAE	$.980_{\pm.033}$	$.913_{\pm.095}$	$.021_{\pm.006}$	$.017_{\pm.011}$	$.021_{\pm.014}$	$.642_{\pm.147}$	$.732_{\pm.488}$
CC-VAE	$.508_{\pm.023}$	$.000_{\pm.000}$	$.003_{\pm.002}$	$.007_{\pm.002}$	$.014_{\pm.004}$	$.222_{\pm.110}$	$1.905_{\pm.023}$
FACTOR	$.974_{\pm.048}$	$.910_{\pm.104}$	$.020_{\pm.003}$	$.019_{\pm.017}$	$.017_{\pm.010}$	$.712_{\pm.183}$	$.667_{\pm.428}$
DIP-I	$.972_{\pm.042}$	$.861_{\pm.097}$	$.020_{\pm.005}$	$.010_{\pm.002}$	$.011_{\pm.002}$	$.618_{\pm.117}$	$1.109_{\pm.312}$
DIP-II	$.930_{\pm.119}$	$.848_{\pm.137}$	$.018_{\pm.004}$	$.010_{\pm.004}$	$.015_{\pm.007}$	$.607_{\pm.207}$	$.907_{\pm.559}$
QUESSARD	$.504_{\pm.021}$	$.000_{\pm.000}$	$.004_{\pm.003}$	$.007_{\pm.004}$	$.018_{\pm.008}$	$.354_{\pm.213}$	$1.686_{\pm.294}$
LSBD-VAE /0	$.970_{\pm.079}$	$.913_{\pm.121}$	$.018_{\pm.003}$	$.052_{\pm.052}$	$.018_{\pm.004}$	$.884_{\pm.183}$	$.749_{\pm.554}$
LSBD-VAE /256	$1.000_{\pm.000}$	$1.000_{\pm.001}$	$.021_{\pm.004}$	$.267_{\pm.152}$	$.027_{\pm.007}$	$.986_{\pm.023}$	$.104_{\pm.147}$
LSBD-VAE /512	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.021_{\pm.006}$	$.393_{\pm.022}$	$.025_{\pm.005}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /768	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.019_{\pm.004}$	$.387_{\pm.014}$	$.025_{\pm.004}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1024	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.005}$	$.398_{\pm.020}$	$.024_{\pm.003}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1280	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.023_{\pm.003}$	$.389_{\pm.016}$	$.023_{\pm.003}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1536	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.004}$	$.398_{\pm.013}$	$.027_{\pm.002}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1792	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.020_{\pm.004}$	$.397_{\pm.016}$	$.027_{\pm.005}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /2048	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.021_{\pm.006}$	$.380_{\pm.027}$	$.027_{\pm.005}$	$.999_{\pm.000}$	$.000_{\pm.000}$

Table 6: Scores for the Arrow dataset.

MODEL	BETA ↑	FACTOR ↑	SAP ↑	DCI ↑	MIG ↑	MOD ↑	$\mathcal{D}_{\text{LSBD}}$ ↓
VAE	$1.000_{\pm.000}$	$.646_{\pm.032}$	$.017_{\pm.004}$	$.009_{\pm.003}$	$.013_{\pm.004}$	$.961_{\pm.012}$	$1.316_{\pm.193}$
β-VAE	$.999_{\pm.002}$	$.588_{\pm.045}$	$.018_{\pm.004}$	$.008_{\pm.002}$	$.015_{\pm.005}$	$.898_{\pm.032}$	$1.178_{\pm.065}$
CC-VAE	$.982_{\pm.056}$	$.707_{\pm.102}$	$.019_{\pm.004}$	$.011_{\pm.005}$	$.016_{\pm.004}$	$.980_{\pm.038}$	$1.013_{\pm.096}$
FACTOR	$1.000_{\pm.000}$	$.659_{\pm.028}$	$.017_{\pm.003}$	$.008_{\pm.003}$	$.014_{\pm.002}$	$.935_{\pm.037}$	$1.526_{\pm.125}$
DIP-I	$1.000_{\pm.000}$	$.624_{\pm.042}$	$.020_{\pm.004}$	$.008_{\pm.002}$	$.012_{\pm.003}$	$.967_{\pm.027}$	$1.521_{\pm.113}$
DIP-II	$1.000_{\pm.000}$	$.644_{\pm.064}$	$.020_{\pm.004}$	$.009_{\pm.003}$	$.013_{\pm.004}$	$.973_{\pm.011}$	$1.616_{\pm.102}$
QUESSARD	$1.000_{\pm.000}$	$.596_{\pm.032}$	$.016_{\pm.006}$	$.008_{\pm.004}$	$.017_{\pm.008}$	$.999_{\pm.000}$	$1.183_{\pm.412}$
LSBD-VAE /0	$1.000_{\pm.001}$	$.664_{\pm.105}$	$.016_{\pm.002}$	$.009_{\pm.004}$	$.019_{\pm.005}$	$.897_{\pm.108}$	$1.627_{\pm.104}$
LSBD-VAE /256	$1.000_{\pm.000}$	$.662_{\pm.046}$	$.017_{\pm.005}$	$.009_{\pm.004}$	$.020_{\pm.005}$	$.963_{\pm.010}$	$1.475_{\pm.121}$
LSBD-VAE /512	$1.000_{\pm.000}$	$.956_{\pm.119}$	$.021_{\pm.006}$	$.297_{\pm.157}$	$.023_{\pm.003}$	$.967_{\pm.092}$	$.245_{\pm.474}$
LSBD-VAE /768	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.006}$	$.390_{\pm.022}$	$.026_{\pm.003}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1024	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.003}$	$.396_{\pm.026}$	$.026_{\pm.006}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1280	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.019_{\pm.005}$	$.401_{\pm.018}$	$.026_{\pm.004}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1536	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.019_{\pm.005}$	$.397_{\pm.017}$	$.026_{\pm.007}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /1792	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.020_{\pm.004}$	$.399_{\pm.018}$	$.026_{\pm.004}$	$.999_{\pm.000}$	$.000_{\pm.000}$
LSBD-VAE /2048	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.020_{\pm.006}$	$.444_{\pm.186}$	$.027_{\pm.004}$	$.999_{\pm.000}$	$.000_{\pm.000}$

Table 7: Scores for the Airplane dataset.

MODEL	BETA ↑	FACTOR ↑	SAP ↑	DCI ↑	MIG ↑	MOD ↑	$\mathcal{D}_{\text{LSBD}}$ ↓
VAE	$1.000_{\pm.001}$	$.947_{\pm.054}$	$.023_{\pm.005}$	$.013_{\pm.005}$	$.020_{\pm.017}$	$.801_{\pm.045}$	$1.342_{\pm.084}$
β-VAE	$1.000_{\pm.001}$	$.997_{\pm.005}$	$.018_{\pm.005}$	$.036_{\pm.012}$	$.028_{\pm.012}$	$.816_{\pm.104}$	$1.481_{\pm.129}$
CC-VAE	$.858_{\pm.194}$	$.646_{\pm.353}$	$.010_{\pm.006}$	$.021_{\pm.011}$	$.018_{\pm.009}$	$.969_{\pm.034}$	$1.481_{\pm.174}$
FACTOR	$1.000_{\pm.000}$	$.984_{\pm.015}$	$.020_{\pm.003}$	$.021_{\pm.008}$	$.026_{\pm.013}$	$.810_{\pm.040}$	$1.382_{\pm.171}$
DIP-I	$1.000_{\pm.000}$	$.994_{\pm.008}$	$.022_{\pm.004}$	$.029_{\pm.012}$	$.026_{\pm.012}$	$.842_{\pm.073}$	$1.289_{\pm.150}$
DIP-II	$.998_{\pm.005}$	$.972_{\pm.031}$	$.021_{\pm.004}$	$.022_{\pm.013}$	$.030_{\pm.019}$	$.780_{\pm.054}$	$1.367_{\pm.129}$
QUESSARD	$.999_{\pm.003}$	$.987_{\pm.026}$	$.018_{\pm.007}$	$.016_{\pm.009}$	$.018_{\pm.005}$	$.795_{\pm.107}$	$.558_{\pm.239}$
LSBD-VAE /0	$.536_{\pm.065}$	$.000_{\pm.000}$	$.002_{\pm.001}$	$.007_{\pm.004}$	$.005_{\pm.003}$	$.956_{\pm.046}$	$1.165_{\pm.180}$
LSBD-VAE /256	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.006}$	$.144_{\pm.011}$	$.023_{\pm.004}$	$.870_{\pm.039}$	$.153_{\pm.021}$
LSBD-VAE /512	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.023_{\pm.008}$	$.151_{\pm.015}$	$.020_{\pm.004}$	$.846_{\pm.032}$	$.168_{\pm.022}$
LSBD-VAE /768	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.004}$	$.140_{\pm.014}$	$.022_{\pm.005}$	$.832_{\pm.034}$	$.180_{\pm.030}$
LSBD-VAE /1024	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.020_{\pm.005}$	$.160_{\pm.015}$	$.022_{\pm.005}$	$.859_{\pm.032}$	$.165_{\pm.021}$
LSBD-VAE /1280	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.024_{\pm.004}$	$.153_{\pm.013}$	$.022_{\pm.003}$	$.876_{\pm.016}$	$.151_{\pm.015}$
LSBD-VAE /1536	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.021_{\pm.005}$	$.160_{\pm.016}$	$.022_{\pm.004}$	$.896_{\pm.025}$	$.140_{\pm.018}$
LSBD-VAE /1792	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.022_{\pm.005}$	$.163_{\pm.022}$	$.023_{\pm.003}$	$.904_{\pm.016}$	$.138_{\pm.010}$
LSBD-VAE /2048	$1.000_{\pm.000}$	$1.000_{\pm.000}$	$.016_{\pm.008}$	$.161_{\pm.024}$	$.021_{\pm.006}$	$.913_{\pm.018}$	$.132_{\pm.009}$

Table 8: Scores for the Modelnet40 Airplanes dataset.

MODEL	BETA ↑	FACTOR ↑	SAP ↑	DCI ↑	MIG ↑	MOD ↑	$\mathcal{D}_{\text{LSBD}}$ ↓
VAE	$.995_{\pm.004}$	$.838_{\pm.030}$	$.013_{\pm.002}$	$.013_{\pm.002}$	$.009_{\pm.002}$	$.415_{\pm.058}$	$.393_{\pm.110}$
β-VAE	$.995_{\pm.005}$	$.857_{\pm.045}$	$.012_{\pm.003}$	$.015_{\pm.003}$	$.009_{\pm.002}$	$.447_{\pm.067}$	$.285_{\pm.045}$
CC-VAE	$.997_{\pm.003}$	$.818_{\pm.093}$	$.011_{\pm.003}$	$.017_{\pm.004}$	$.011_{\pm.003}$	$.567_{\pm.063}$	$.281_{\pm.191}$
FACTOR	$.996_{\pm.004}$	$.856_{\pm.052}$	$.012_{\pm.002}$	$.014_{\pm.003}$	$.010_{\pm.003}$	$.444_{\pm.077}$	$.388_{\pm.096}$
DIP-I	$.988_{\pm.009}$	$.783_{\pm.070}$	$.012_{\pm.002}$	$.013_{\pm.002}$	$.008_{\pm.001}$	$.343_{\pm.082}$	$.416_{\pm.142}$
DIP-II	$.994_{\pm.006}$	$.832_{\pm.042}$	$.013_{\pm.003}$	$.014_{\pm.003}$	$.011_{\pm.002}$	$.433_{\pm.080}$	$.379_{\pm.130}$
QUESSARD	$.907_{\pm.192}$	$.727_{\pm.384}$	$.010_{\pm.005}$	$.015_{\pm.007}$	$.009_{\pm.004}$	$.563_{\pm.108}$	$.134_{\pm.294}$
LSBD-VAE /0	$.990_{\pm.009}$	$.863_{\pm.038}$	$.011_{\pm.003}$	$.015_{\pm.003}$	$.014_{\pm.003}$	$.538_{\pm.103}$	$.731_{\pm.068}$
LSBD-VAE /256	$1.000_{\pm.000}$	$.990_{\pm.004}$	$.012_{\pm.005}$	$.052_{\pm.009}$	$.020_{\pm.006}$	$.947_{\pm.007}$	$.041_{\pm.007}$

Table 9: Scores for COIL 100 dataset.

MODEL	BETA ↑	FACTOR ↑	SAP ↑	DCI ↑	MIG ↑	MOD ↑	$\mathcal{D}_{\mathrm{LSBD}}$ ↓
VAE	$1.000_{\pm.000}$	$.674_{\pm.049}$	$.014_{\pm.003}$	$.016_{\pm.003}$	$.011_{\pm.002}$	$.986_{\pm.001}$	$.463_{\pm.030}$
β-VAE	$1.000_{\pm.001}$	$.740_{\pm.024}$	$.015_{\pm.004}$	$.014_{\pm.004}$	$.013_{\pm.003}$	$.982_{\pm.001}$	$.579_{\pm.095}$
CC-VAE	$.999_{\pm.003}$	$.723_{\pm.026}$	$.013_{\pm.005}$	$.014_{\pm.003}$	$.013_{\pm.004}$	$.985_{\pm.001}$	$.406_{\pm.057}$
FACTOR	$1.000_{\pm.001}$	$.684_{\pm.041}$	$.014_{\pm.002}$	$.012_{\pm.002}$	$.013_{\pm.004}$	$.984_{\pm.001}$	$.490_{\pm.024}$
DIP-I	$.999_{\pm.002}$	$.631_{\pm.025}$	$.013_{\pm.004}$	$.012_{\pm.002}$	$.010_{\pm.002}$	$.986_{\pm.001}$	$.525_{\pm.109}$
DIP-II	$1.000_{\pm.001}$	$.643_{\pm.043}$	$.013_{\pm.003}$	$.014_{\pm.002}$	$.011_{\pm.002}$	$.985_{\pm.001}$	$.568_{\pm.079}$
QUESSARD	$1.000_{\pm.000}$	$.780_{\pm.044}$	$.014_{\pm.004}$	$.014_{\pm.002}$	$.011_{\pm.003}$	$.973_{\pm.004}$	$.396_{\pm.055}$
LSBD-VAE /0	$1.000_{\pm.001}$	$.739_{\pm.047}$	$.014_{\pm.003}$	$.014_{\pm.001}$	$.011_{\pm.001}$	$.982_{\pm.004}$	$.515_{\pm.099}$
LSBD-VAE /256	$1.000_{\pm.000}$	$.655_{\pm.028}$	$.015_{\pm.004}$	$.029_{\pm.003}$	$.013_{\pm.003}$	$.802_{\pm.056}$	$.112_{\pm.026}$