
Disentanglement Beyond Static vs. Dynamic: A Benchmark and Evaluation Framework for Multi-Factor Sequential Representations

Tal Barami* Nimrod Berman* Ilan Naiman* Amos H. Hason Rotem Ezra Omri Azencot
Faculty of Computer and Information Science
Ben-Gurion University of the Negev
{baramit, bermann, naimani, hasona, rotomez}@post.bgu.ac.il
azencot@bgu.ac.il

dMelodies-WAV Results

Table 1: Swap intervention on the dMelodies-WAV dataset for Sparse-AE.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 2: Swap intervention on the dMelodies-WAV dataset for VAE.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

*Equal contribution

Table 3: **Swap intervention** on the **dMelodies-WAV** dataset for β -VAE.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 4: **Swap intervention** on the **dMelodies-WAV** dataset for **SKD**.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 5: **Swap intervention** on the **dMelodies-WAV** dataset for **SSM-SKD**.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.31 \pm 2e-03$	$0.08 \pm 2e-03$	$0.33 \pm 1e-02$	$0.04 \pm 2e-03$	$0.51 \pm 1e-02$	$0.50 \pm 9e-03$
Tonic	$0.31 \pm 2e-03$	$0.08 \pm 4e-03$	$0.33 \pm 9e-03$	$0.05 \pm 5e-03$	$0.51 \pm 2e-02$	$0.50 \pm 1e-02$
Scale	$0.30 \pm 2e-03$	$0.08 \pm 3e-03$	$0.34 \pm 2e-02$	$0.06 \pm 6e-03$	$0.56 \pm 1e-02$	$0.51 \pm 1e-02$
Rhythm bar1	$0.30 \pm 5e-03$	$0.08 \pm 3e-03$	$0.34 \pm 1e-02$	$0.06 \pm 4e-03$	$0.51 \pm 1e-02$	$0.50 \pm 9e-03$
Arp chord1	$0.31 \pm 2e-03$	$0.08 \pm 6e-03$	$0.34 \pm 9e-03$	$0.05 \pm 3e-03$	$0.52 \pm 1e-02$	$0.51 \pm 9e-03$
Arp chord2	$0.30 \pm 4e-03$	$0.08 \pm 5e-03$	$0.33 \pm 1e-02$	$0.06 \pm 7e-03$	$0.52 \pm 9e-03$	$0.57 \pm 8e-03$

Table 6: **Sample intervention** on the **dMelodies-WAV** dataset for **Sparse-AE**.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 7: **Sample intervention** on the **dMelodies-WAV** dataset for **VAE**.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 8: **Sample intervention** on the **dMelodies-WAV** dataset for β -VAE.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.24 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 9: **Sample intervention** on the **dMelodies-WAV** dataset for SKD.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Tonic	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Scale	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Rhythm bar1	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord1	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$
Arp chord2	$0.29 \pm 0e+00$	$0.08 \pm 0e+00$	$0.32 \pm 0e+00$	$0.03 \pm 0e+00$	$0.49 \pm 0e+00$	$0.49 \pm 0e+00$

Table 10: **Sample intervention** on the **dMelodies-WAV** dataset for SSM-SKD.

None	Instrument	Tonic	Scale	Rhythm bar1	Arp chord1	Arp chord2
Instrument	$0.29 \pm 2e-03$	$0.09 \pm 4e-03$	$0.33 \pm 1e-02$	$0.04 \pm 3e-03$	$0.50 \pm 9e-03$	$0.50 \pm 8e-03$
Tonic	$0.29 \pm 2e-03$	$0.08 \pm 2e-03$	$0.33 \pm 1e-02$	$0.05 \pm 5e-03$	$0.52 \pm 1e-02$	$0.51 \pm 1e-02$
Scale	$0.29 \pm 1e-03$	$0.07 \pm 5e-03$	$0.33 \pm 2e-02$	$0.06 \pm 4e-03$	$0.58 \pm 3e-02$	$0.52 \pm 2e-03$
Rhythm bar1	$0.29 \pm 2e-03$	$0.09 \pm 2e-03$	$0.33 \pm 2e-02$	$0.07 \pm 5e-03$	$0.52 \pm 1e-02$	$0.52 \pm 1e-02$
Arp chord1	$0.29 \pm 2e-03$	$0.08 \pm 5e-03$	$0.33 \pm 2e-02$	$0.05 \pm 4e-03$	$0.54 \pm 1e-02$	$0.51 \pm 9e-03$
Arp chord2	$0.29 \pm 1e-03$	$0.07 \pm 6e-03$	$0.33 \pm 1e-02$	$0.06 \pm 2e-03$	$0.52 \pm 1e-02$	$0.57 \pm 2e-02$

Table 11: **DCI** on the **dMelodies-WAV** dataset for **Sparse-AE**.

None	DCI-M \uparrow	DCI-C \uparrow	DCI-E \uparrow
Instrument	$0.05 \pm 6e-03$	$0.70 \pm 2e-03$	$0.73 \pm 1e-02$
Tonic	$0.01 \pm 8e-03$	$0.68 \pm 7e-04$	$0.08 \pm 5e-03$
Scale	$0.01 \pm 3e-03$	$0.68 \pm 2e-03$	$0.33 \pm 2e-02$
Rhythm bar1	$0.01 \pm 6e-03$	$0.68 \pm 5e-04$	$0.09 \pm 2e-02$
Arp chord1	$0.01 \pm 4e-03$	$0.68 \pm 2e-03$	$0.51 \pm 3e-02$
Arp chord2	$0.01 \pm 5e-03$	$0.68 \pm 2e-03$	$0.51 \pm 3e-02$

Table 12: **DCI** on the **dMelodies-WAV** dataset for **VAE**.

None	DCI-M \uparrow	DCI-C \uparrow	DCI-E \uparrow
Instrument	$0.00 \pm 1e-03$	$0.69 \pm 2e-03$	$0.25 \pm 2e-02$
Tonic	$0.00 \pm 3e-03$	$0.69 \pm 5e-04$	$0.08 \pm 5e-03$
Scale	$0.00 \pm 2e-03$	$0.69 \pm 3e-03$	$0.31 \pm 4e-02$
Rhythm bar1	$0.06 \pm 2e-02$	$0.68 \pm 4e-04$	$0.10 \pm 2e-02$
Arp chord1	$0.00 \pm 3e-03$	$0.69 \pm 9e-03$	$0.50 \pm 3e-02$
Arp chord2	$0.00 \pm 4e-03$	$0.69 \pm 6e-03$	$0.49 \pm 1e-02$

Table 13: **DCI** on the **dMelodies-WAV** dataset for β -**VAE**.

None	DCI-M \uparrow	DCI-C \uparrow	DCI-E \uparrow
Instrument	$0.64 \pm 8e-03$	$0.87 \pm 4e-03$	$0.78 \pm 9e-03$
Tonic	$0.06 \pm 4e-03$	$0.73 \pm 1e-03$	$0.09 \pm 8e-03$
Scale	$0.06 \pm 6e-03$	$0.74 \pm 2e-03$	$0.33 \pm 2e-02$
Rhythm bar1	$0.06 \pm 2e-02$	$0.71 \pm 6e-04$	$0.25 \pm 2e-02$
Arp chord1	$0.06 \pm 5e-03$	$0.74 \pm 6e-03$	$0.52 \pm 2e-02$
Arp chord2	$0.05 \pm 2e-03$	$0.74 \pm 4e-03$	$0.49 \pm 1e-02$

Table 14: **DCI** on the **dMelodies-WAV** dataset for **SKD**.

None	DCI-M \uparrow	DCI-C \uparrow	DCI-E \uparrow
Instrument	$0.06 \pm 2e-02$	$0.68 \pm 1e-02$	$0.99 \pm 0e+00$
Tonic	$0.07 \pm 1e-02$	$0.56 \pm 5e-03$	$0.23 \pm 1e-02$
Scale	$0.09 \pm 2e-02$	$0.56 \pm 7e-03$	$0.36 \pm 3e-03$
Rhythm bar1	$0.12 \pm 6e-03$	$0.76 \pm 3e-02$	$0.94 \pm 1e-03$
Arp chord1	$0.08 \pm 5e-03$	$0.56 \pm 8e-03$	$0.68 \pm 2e-03$
Arp chord2	$0.08 \pm 1e-02$	$0.57 \pm 1e-02$	$0.67 \pm 1e-03$

Table 15: **DCI** on the **dMelodies-WAV** dataset for **SSM-SKD**.

None	DCI-M \uparrow	DCI-C \uparrow	DCI-E \uparrow
Instrument	$0.37 \pm 1\text{e-}02$	$0.80 \pm 3\text{e-}03$	$1.00 \pm 0\text{e+}00$
Tonic	$0.06 \pm 1\text{e-}02$	$0.63 \pm 2\text{e-}03$	$0.29 \pm 9\text{e-}03$
Scale	$0.06 \pm 2\text{e-}03$	$0.63 \pm 2\text{e-}03$	$0.38 \pm 6\text{e-}03$
Rhythm bar1	$0.11 \pm 4\text{e-}03$	$0.66 \pm 1\text{e-}02$	$0.99 \pm 3\text{e-}03$
Arp chord1	$0.12 \pm 1\text{e-}02$	$0.64 \pm 3\text{e-}03$	$0.70 \pm 3\text{e-}03$
Arp chord2	$0.06 \pm 2\text{e-}03$	$0.66 \pm 4\text{e-}03$	$0.66 \pm 1\text{e-}03$