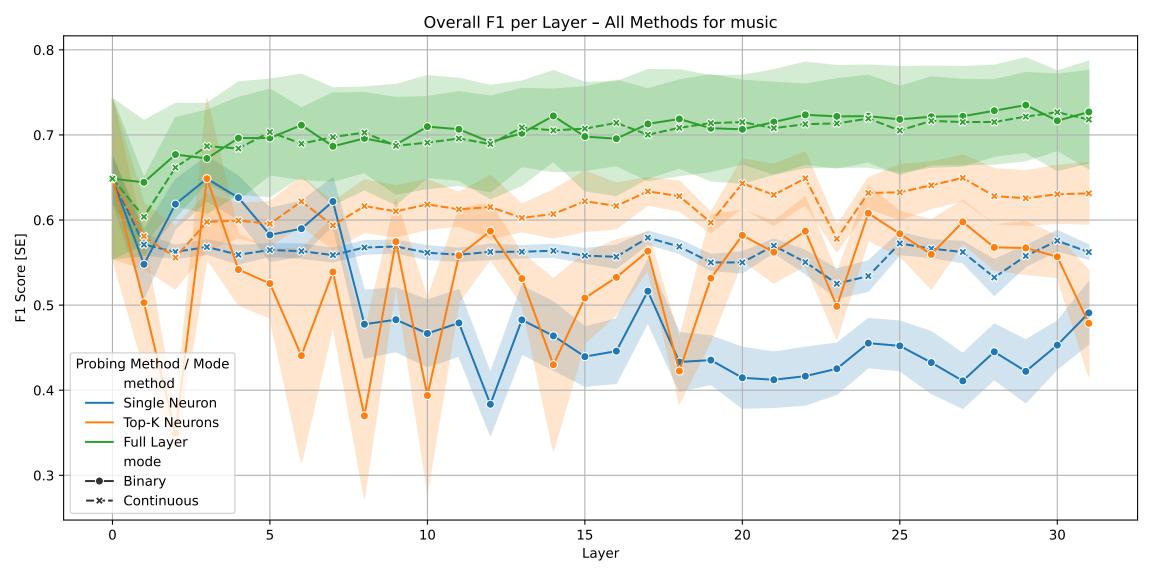


F1 per Layer – Top-K Neurons Probing for music Mode Binary 0.7 Continuous 0.6 F1 [SE] 0.4 0.3 20 30 10 15 25 Layer

F1 per Layer - Full Layer Probing for music 0.80 Mode Binary Continuous 0.75 0.70 F1 [SE] 0.65 0.60 0.55 20 25 30 10 15 Layer



F1 Score Summary by Probing Method Method Metric Binary f1 best layer

f1 max

f1 mean

f1_std

Full Layer

Full Laver

Top-K Neurons

Top-K Neurons

- 1		_		I I
Ī	Full Layer	f1_mean	0.7044	0.6996
	Full Layer	f1_std	0.1121	0.0957
ſ	Single Neuron	f1_best_layer	0.0	0.0
Ī	Single Neuron	f1_max	0.8526	0.8526
Ī	Single Neuron	f1_mean	0.4882	0.563
	Single Neuron	f1_std	0.2251	0.0757
	Top-K Neurons	f1_best_layer	0.0	27.0
ſ	Top-K Neurons	f1_max	0.8526	0.8526
-	Single Neuron Top-K Neurons	f1_std f1_best_layer	0.2251 0.0	0.0757 27.0

29.0

0.8597

0.5297

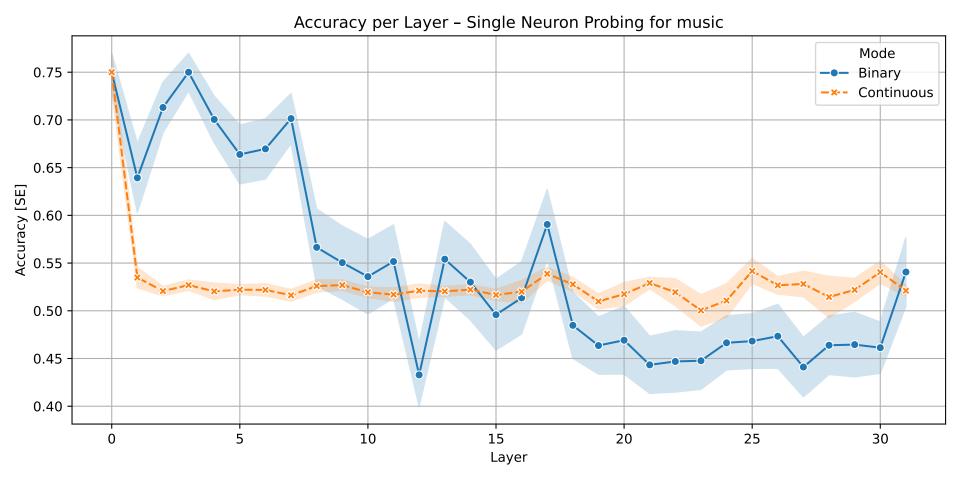
0.1275

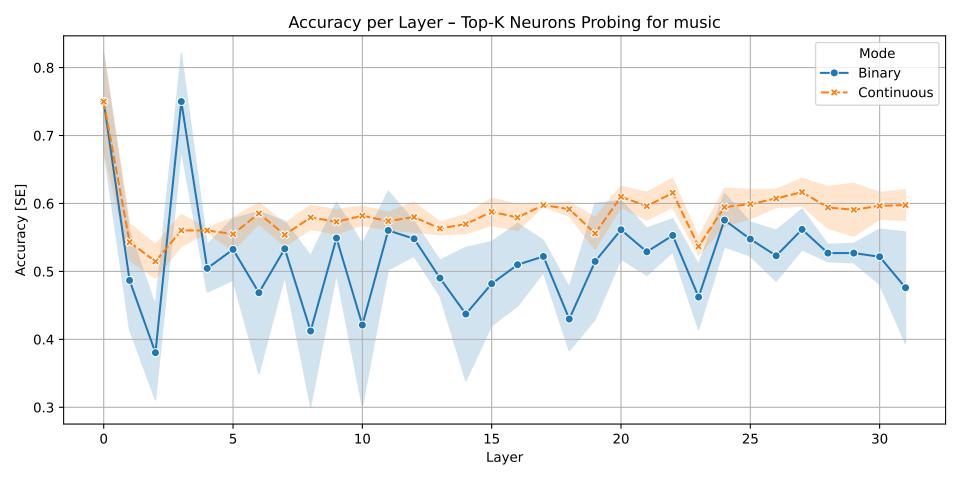
Continuous

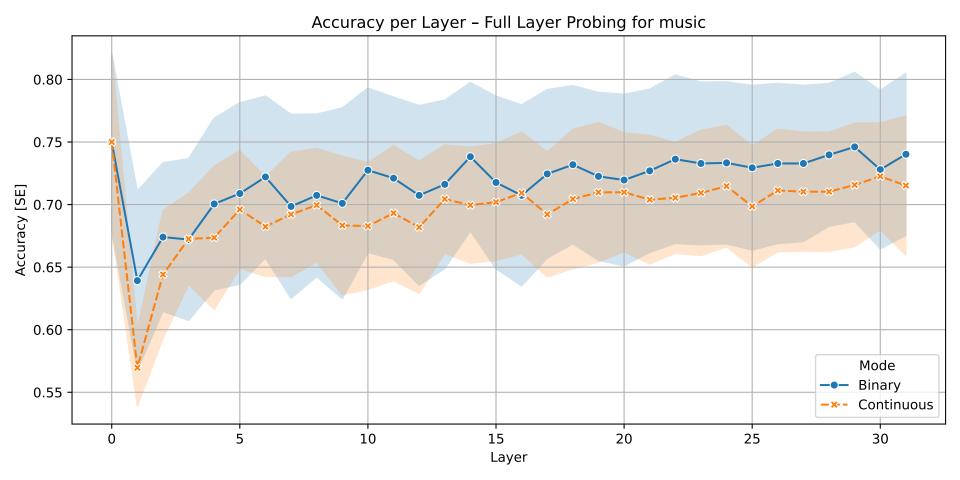
30.0 0.8526

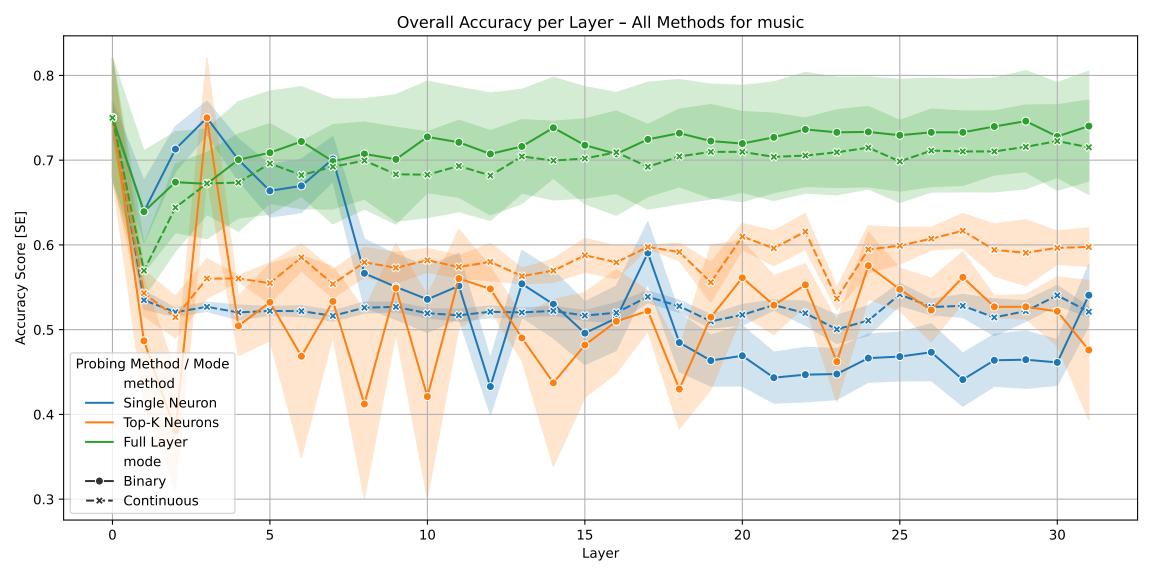
0.617

0.061









	Accuracy Score Summi	ary by ribbing Method
Method	Metric	Binary

accuracy best layer

accuracy_std

Full Layer

Top-K Neurons

,			I .
Full Layer	accuracy_max	0.9	0.9
Full Layer	accuracy_mean	0.7183	0.6959
Full Layer	accuracy_std	0.1179	0.0909
Single Neuron	accuracy_best_layer	0.0	0.0
Single Neuron	accuracy_max	0.9	0.9
Single Neuron	accuracy_mean	0.5451	0.5297
Single Neuron	accuracy_std	0.2276	0.0768
Top-K Neurons	accuracy_best_layer	0.0	0.0
Top-K Neurons	accuracy_max	0.9	0.9
Top-K Neurons	accuracy mean	0.5202	0.5847

Accuracy Score Summary by Probing Method

0.0

0.131

Continuous

0.0

0.0556