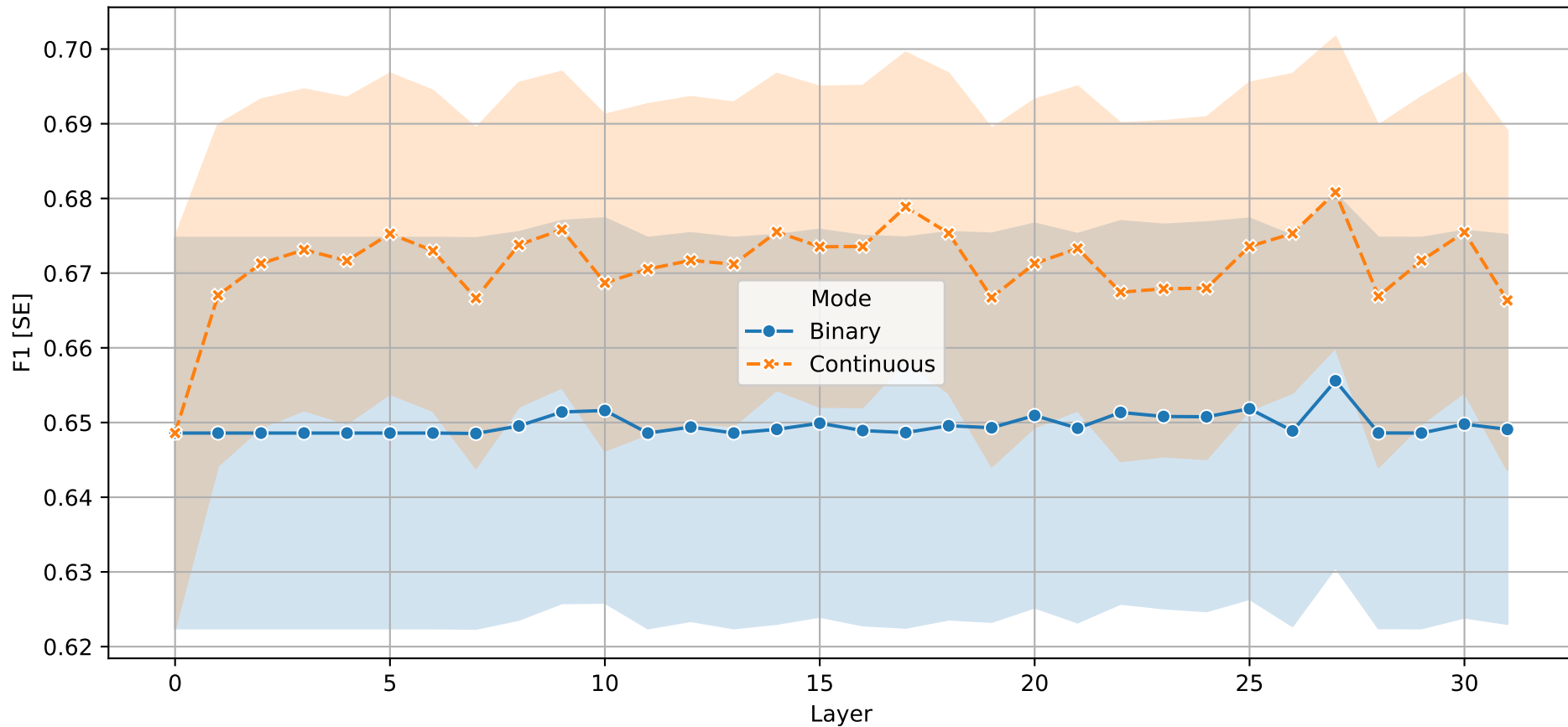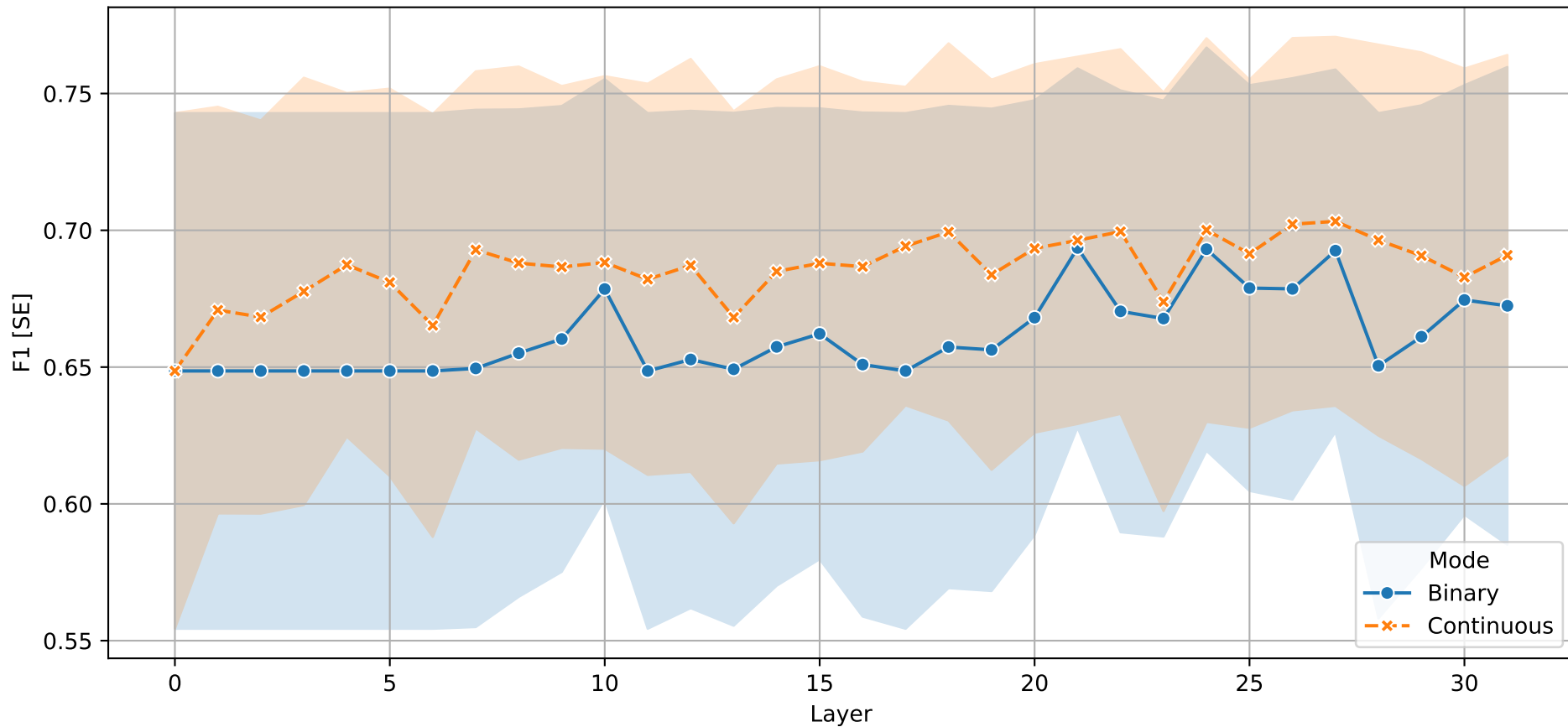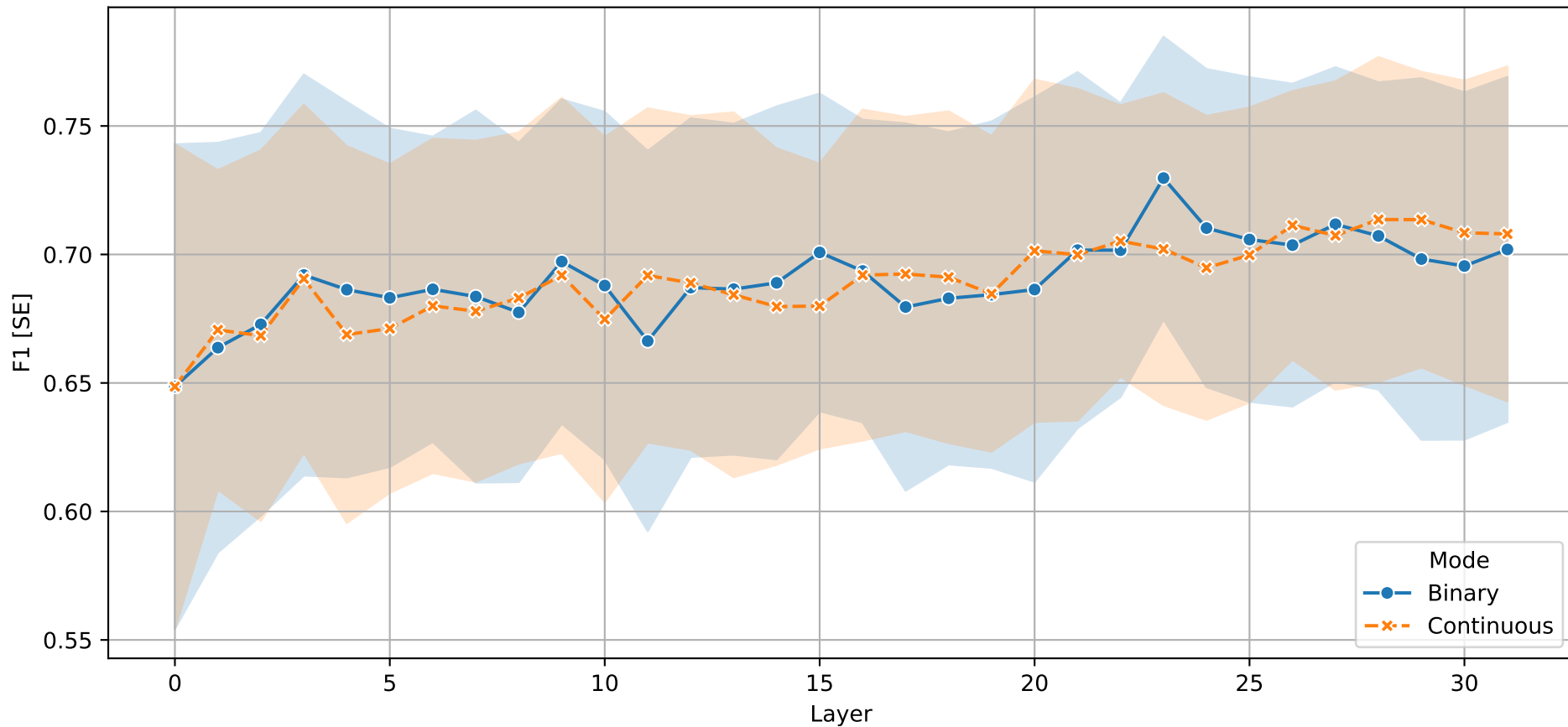F1 per Layer – Single Neuron Probing for music
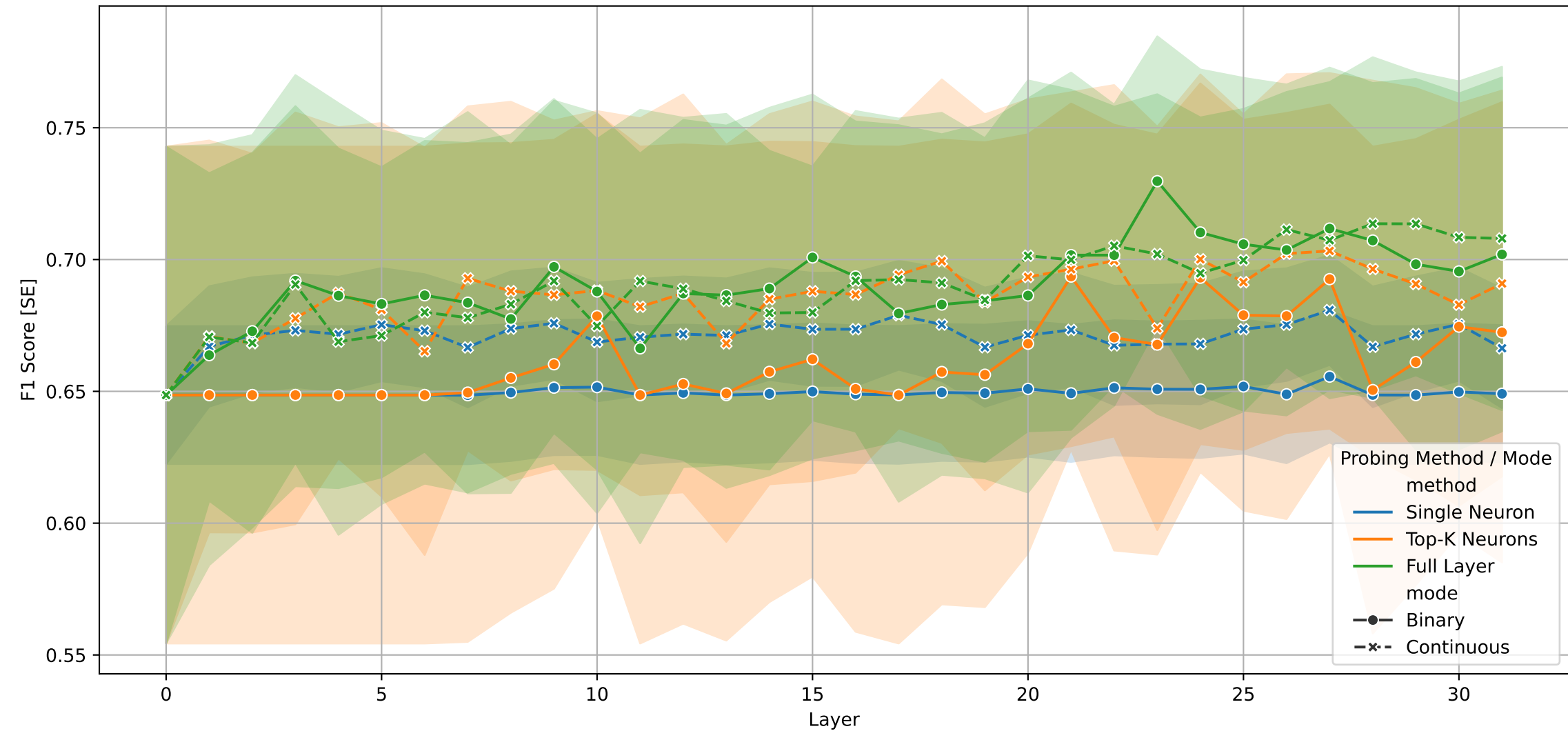
F1 per Layer – Top-K Neurons Probing for music

F1 per Layer – Full Layer Probing for music

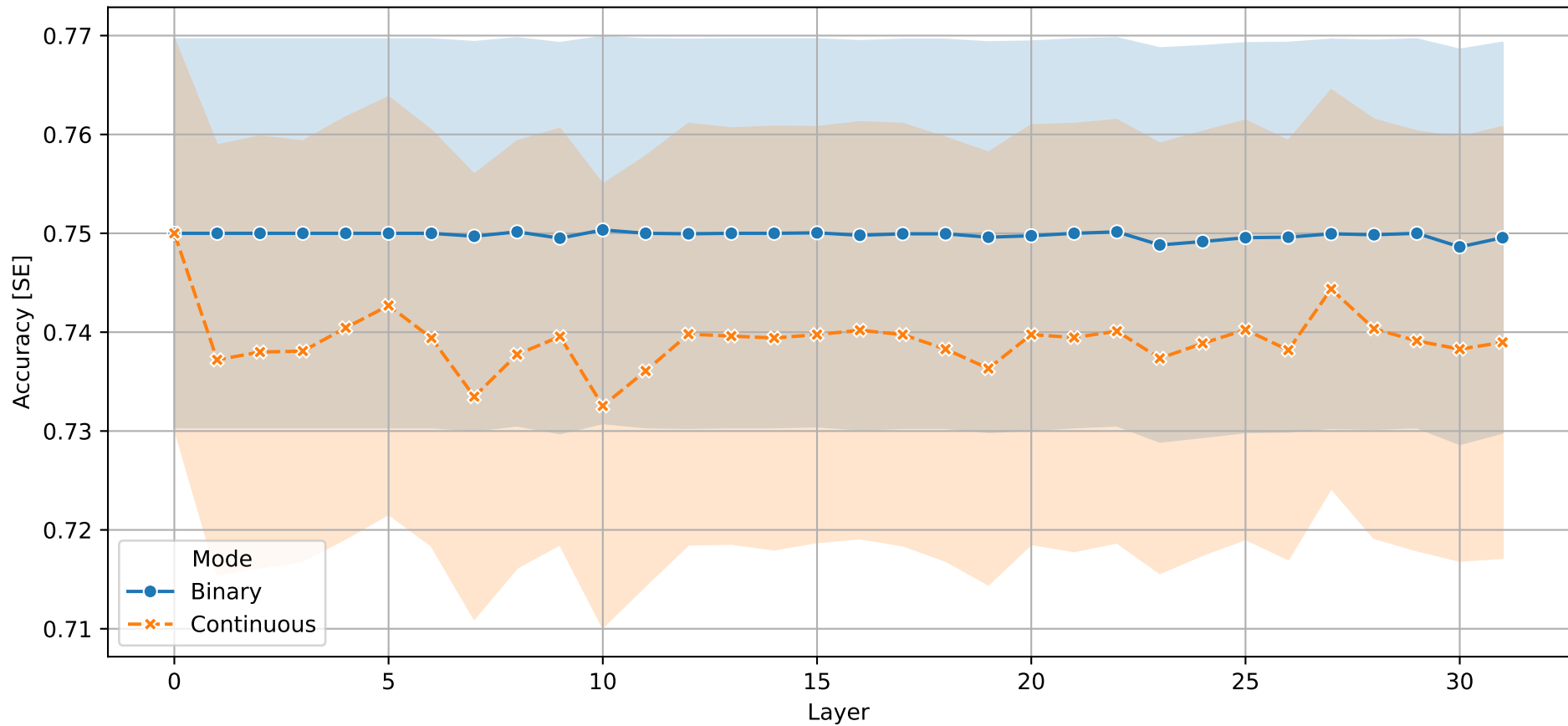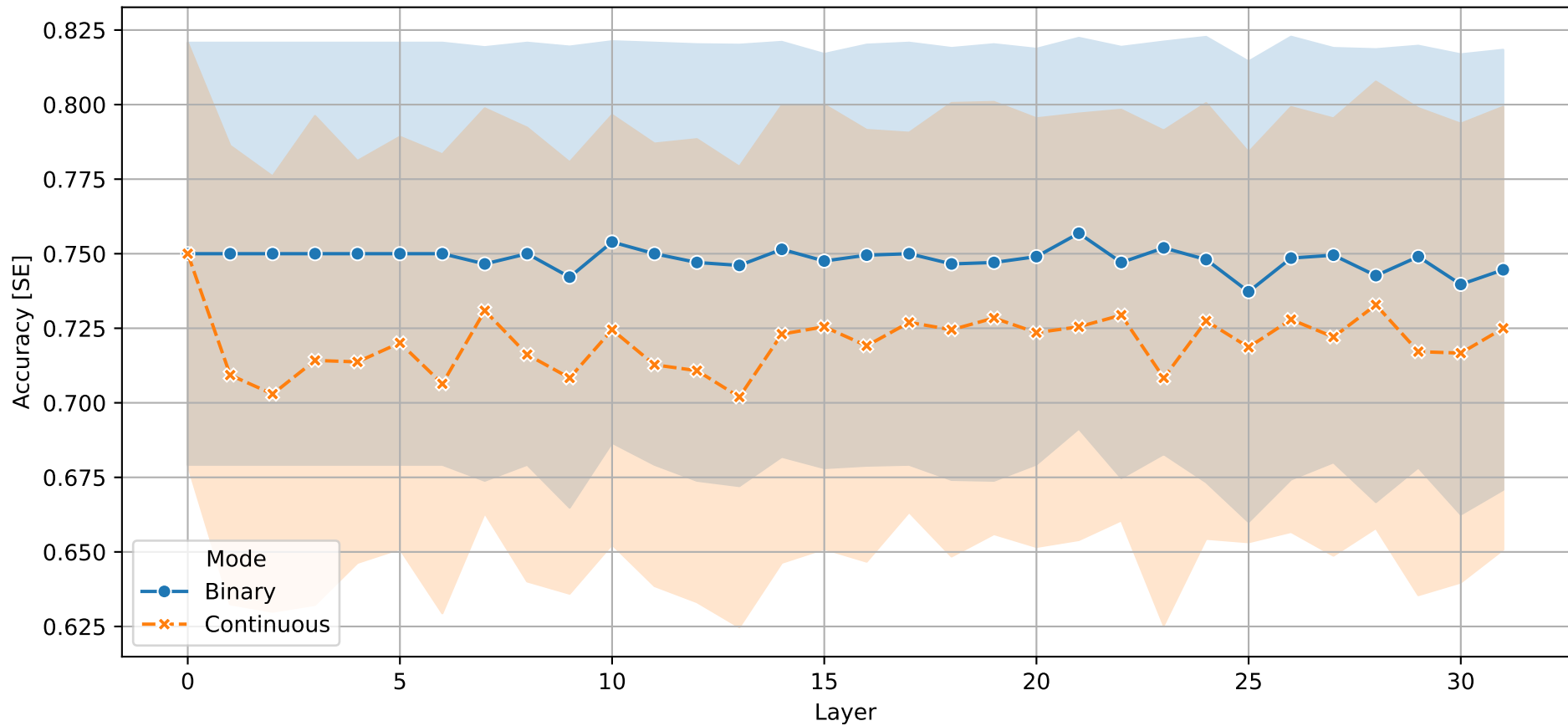Overall F1 per Layer – All Methods for music

F1 Score Summary by Probing Method

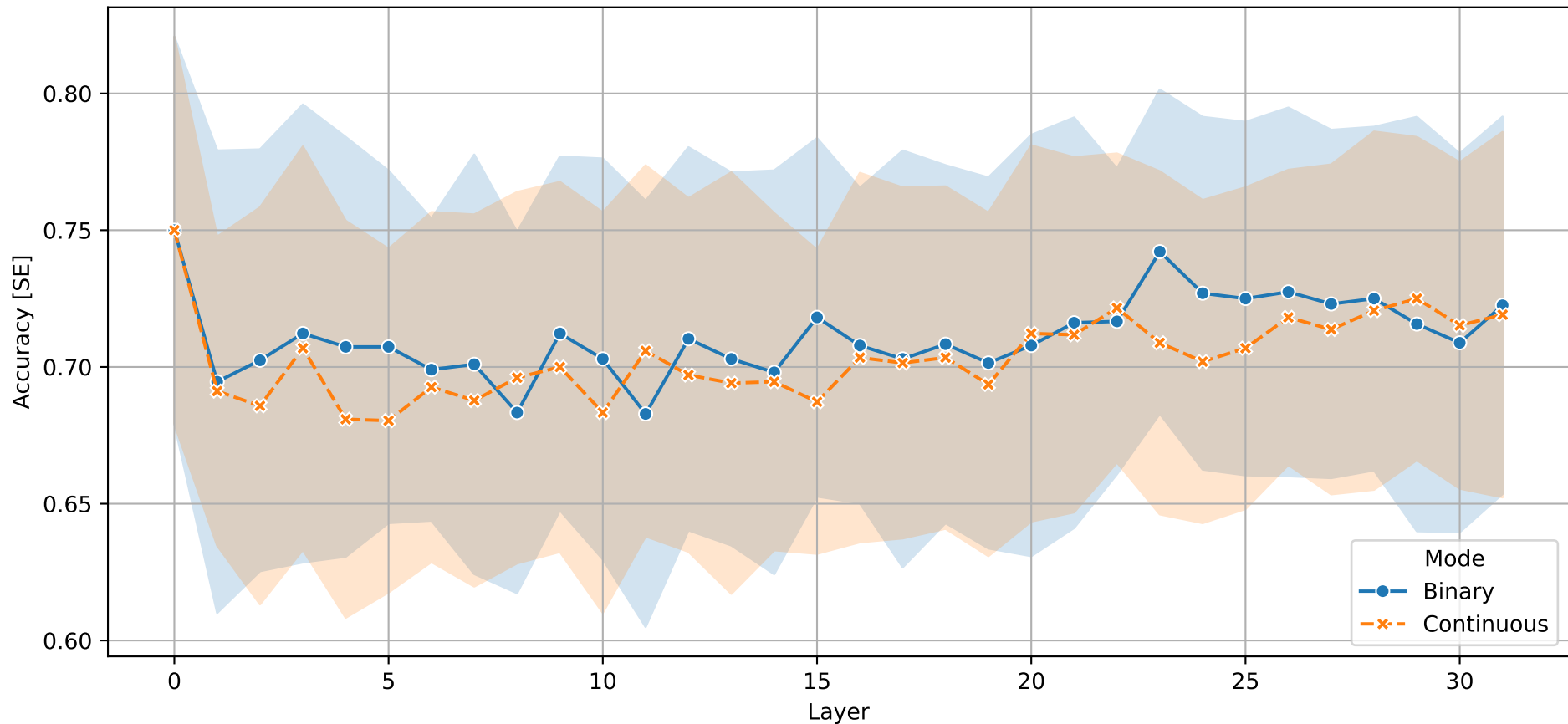| Method	Metric	Binary	Continuous
Full Layer	f1_best_layer	23.0	28.0
Full Layer	f1_max	0.8631	0.8534
Full Layer	f1_mean	0.6907	0.6899
Full Layer	f1_std	0.1191	0.1135
Single Neuron	f1_best_layer	27.0	27.0
Single Neuron	f1_max	0.8526	0.8526
Single Neuron	f1_mean	0.6496	0.6713
Single Neuron	f1_std	0.1621	0.1377
Top-K Neurons	f1_best_layer	21.0	27.0
Top-K Neurons	f1_max	0.8526	0.8608
Top-K Neurons	f1_mean	0.6615	0.6859
Top-K Neurons	f1_std	0.1511	0.1246

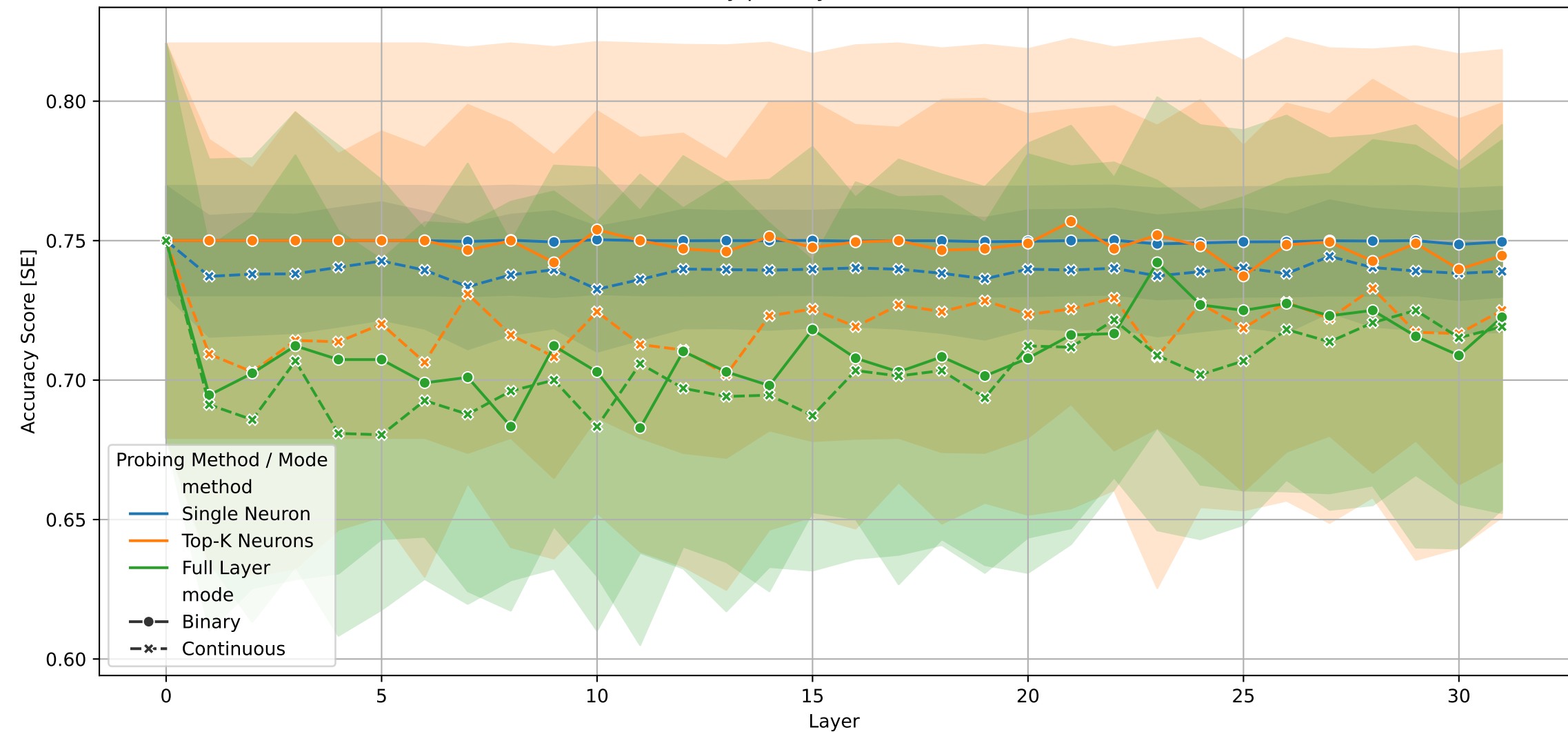Accuracy per Layer – Single Neuron Probing for music

Accuracy per Layer – Top-K Neurons Probing for music

Accuracy per Layer – Full Layer Probing for music

Overall Accuracy per Layer – All Methods for music

Accuracy Score Summary by Probing Method

Method	Metric	Binary	Continuous
Full Layer	accuracy_best_layer	0.0	0.0
Full Layer	accuracy_max	0.9	0.9
Full Layer	accuracy_mean	0.7113	0.7035
Full Layer	accuracy_std	0.1221	0.1138
Single Neuron	accuracy_best_layer	10.0	0.0
Single Neuron	accuracy_max	0.9	0.9
Single Neuron	accuracy_mean	0.7498	0.7392
Single Neuron	accuracy_std	0.1228	0.1333
Top-K Neurons	accuracy_best_layer	21.0	0.0
Top-K Neurons	accuracy_max	0.9	0.902
Top-K Neurons	accuracy_mean	0.7482	0.7201
Top-K Neurons	accuracy_std	0.1249	0.1285