

A APPENDIX

A.1 METHODS DETAILS

DC-RNN weights were initialized using He initialization, a uniform distribution $\mathcal{U}(-\sqrt{\frac{1}{hidden_size}}, \sqrt{\frac{1}{hidden_size}})$, which is the default in PyTorch (Paszke et al., 2017). The DC-RNN was implemented with PyTorch (Paszke et al., 2017). The learning rate and batch size were selected after initial parameter sweep with learning rate of 1e-2, 1e-3, and batch size of 16, 32.

A.2 THE USE OF LARGE LANGUAGE MODELS (LLMs)

The Large Language Model ChatGPT (<https://chatgpt.com/>) was used only to revise grammar in writing.

A.3 SUPPLEMENTARY MATERIALS

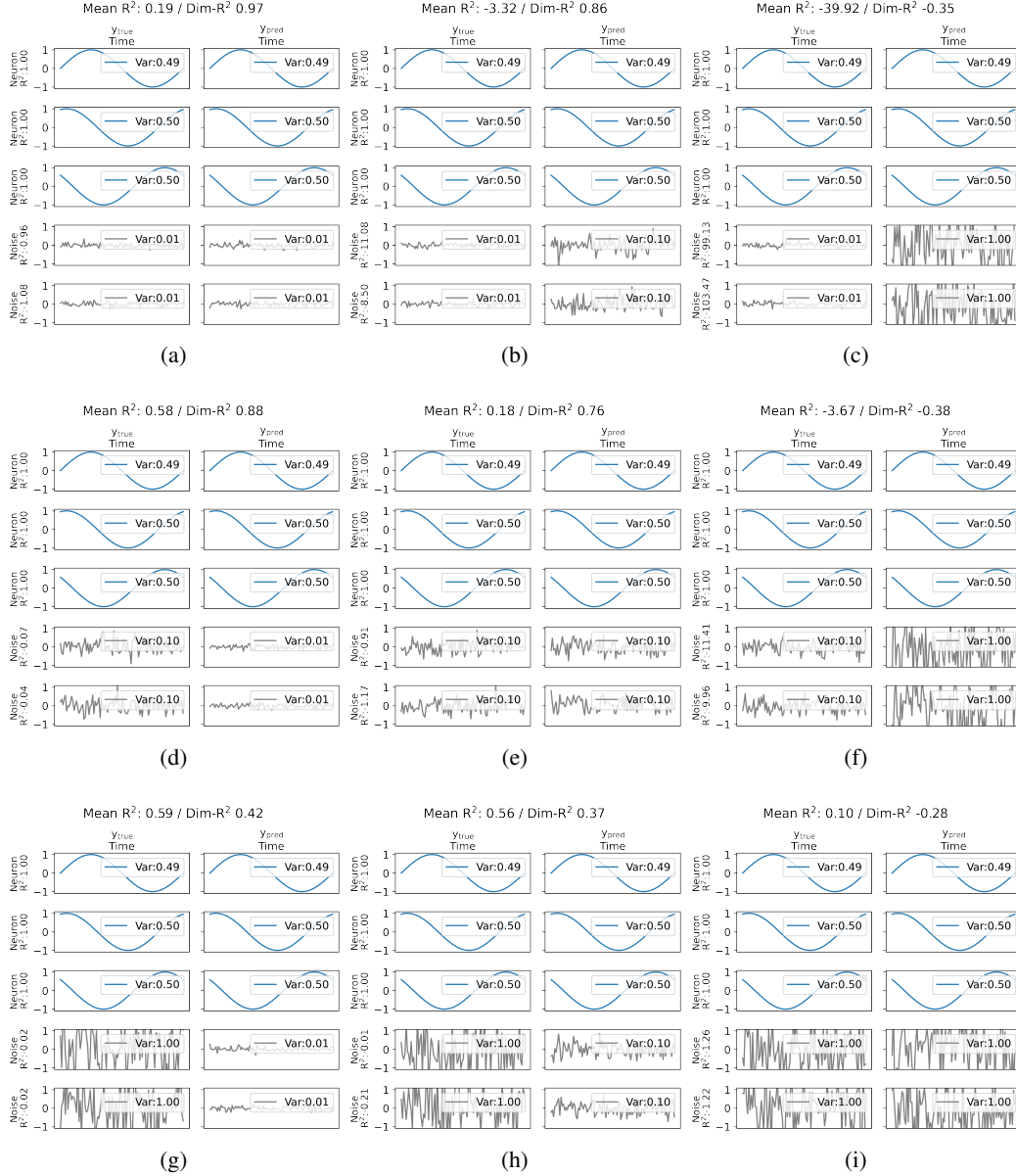


Figure 11: Example schematic waveforms of y_{true} and y_{pred} with their corresponding mean R^2 and Dim- R^2 scores. Each example shows sine waves with varying phases across 5 channels, where 2 channels have been replaced with Gaussian noise of specified variance. Rows share the same y_{true} noise channel variance; columns share the same y_{pred} noise channel variance. Legends show variances.

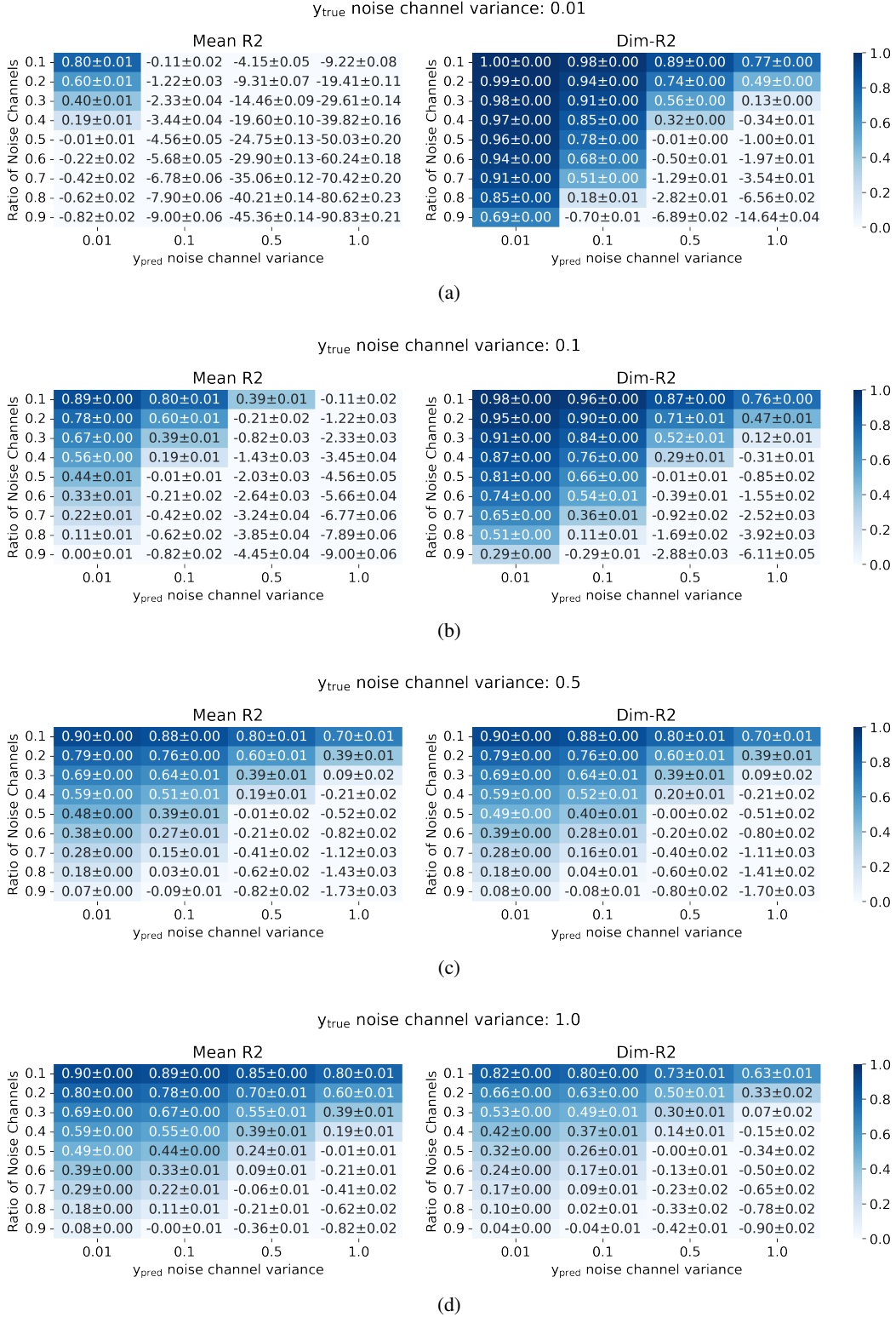


Figure 12: Dim-R2 highlights the presence of channels with high predictive accuracy in the presence of noisy channels. Scores were measured on simulated sinusoidal data (Fig. 8) across hyperparameter sweeps. Each entry shows the mean $\pm$ standard deviation across 100 random repetitions.

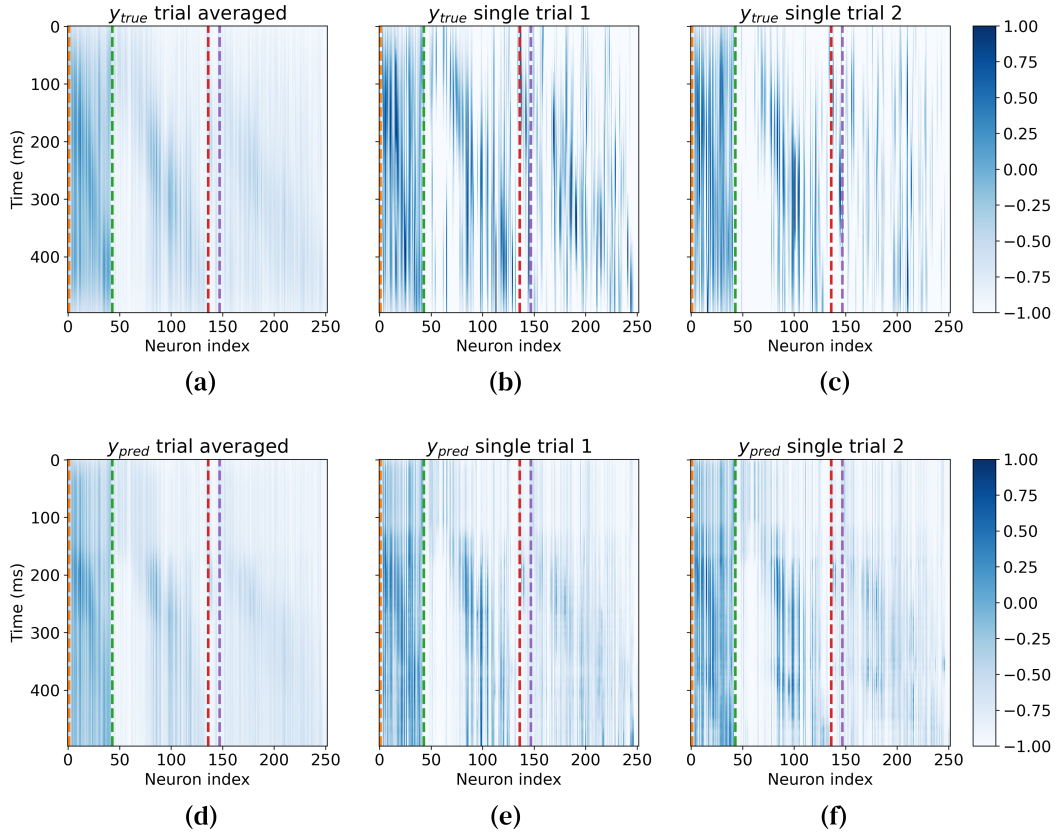


Figure 13: Examples of y_{true} and y_{pred} from DC-RNN trained to reproduce neural activity. This single session contains 78 trials, with 42, 93, 11, and 106 neurons from DCN, M1, Striatum, and Thalamus, respectively. This session corresponds to Session index 21 with a 50ms Gaussian filter in Fig. 10. Dashed lines separate neurons by brain region (from left): DCN (orange), M1 (green), Striatum (red), Thalamus (purple). (a) & (d) Trial averaged activity, (b) & (e) Single trial example 1, (c) & (f) Single trial example 2. (a)-(c) y_{true} , (d)-(f) y_{pred} .

	Trial averaged (a) & (d)	Single trial 1 (b) & (e)	Single trial 2 (c) & (f)
Dim-R2	0.92	0.44	0.57
Mean R2	0.57	-0.28	-0.41

Table 1: Dim-R2 and Mean R2 measured on y_{true} and y_{pred} of Fig. 13. Dim-R2 yields higher scores than mean R2 when the y_{true} and y_{pred} are similar, even in the presence of noise.