Supporting Information - Linking Device Dynamics to Neural Network Performance in Ionically Gated Synaptic Transistors

Nithil Harris Manimaran[1], Huayuan Han[2], Cory Merkel[3], and Ke Xu[1,2,4*]

[1]*Microsystems Engineering, Rochester Institute of Technology, Rochester, New York 14623, United States*

[2]*School of Physics and Astronomy, Rochester Institute of Technology, Rochester, New York 14623, United States*

[3]*Department of Electrical and Computer Engineering, Rochester Institute of Technology, Rochester, New York 14623, United States*

[4]*School of Chemistry and Material Science, Rochester Institute of Technology, Rochester, New York 14623, United States*

**Corresponding author – Email:* ke.xu@rit.edu

Supporting Information

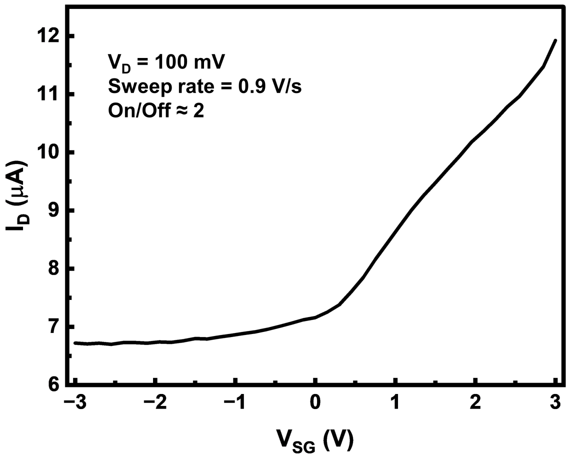

Figure S1: I_D-V_G transfer characteristics of the MoS_2 EDLT with a low on/off ratio.

Fig. S1 shows the I_D-V_G transfer characteristics of the MoS_2 EDLT with a relatively low on/off ratio (~2), in contrast to the high on/off device discussed in Fig. 1b of the main text. The limited conductance range in this device leads to earlier saturation during pulsed operation, which constrains the number of distinguishable conductance states and contributed to increased weight-update nonlinearity.

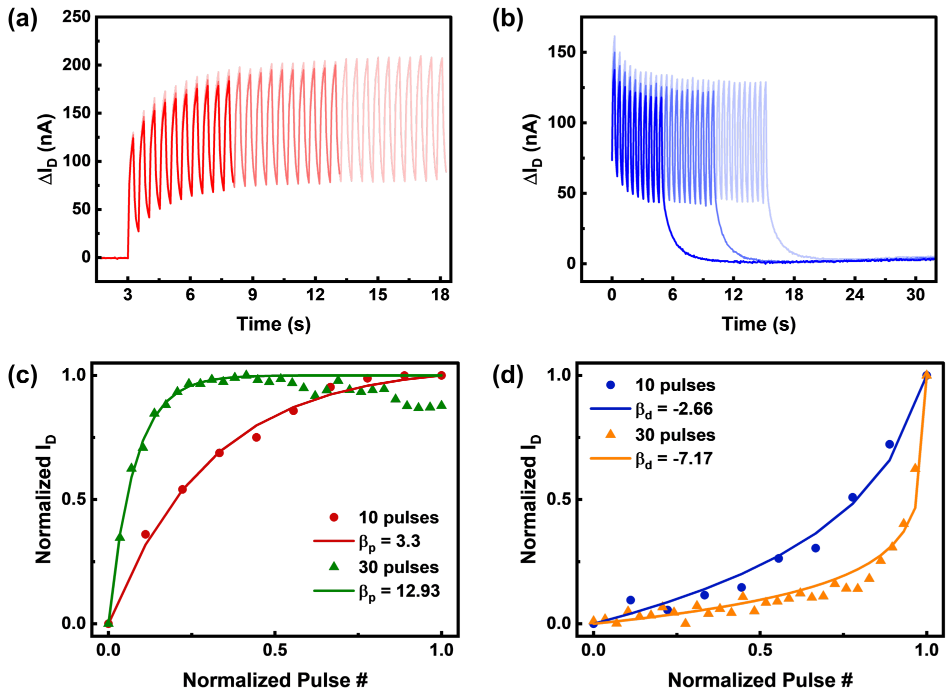

Figure S2: (a, b) Experimentally measured transient drain current response to 10, 20, and 30 gate voltage pulses, applied to the high on/off device, from which the normalized synaptic weights are extracted for Fig. 2a. (c, d) Extracted asymmetric nonlinearity (ANL) factor β for potentiation and depression as a function of state number N for the low on/off device, depicting increased nonlinearity and loss of distinguishable conductance states beyond ~10 pulses due to early saturation.

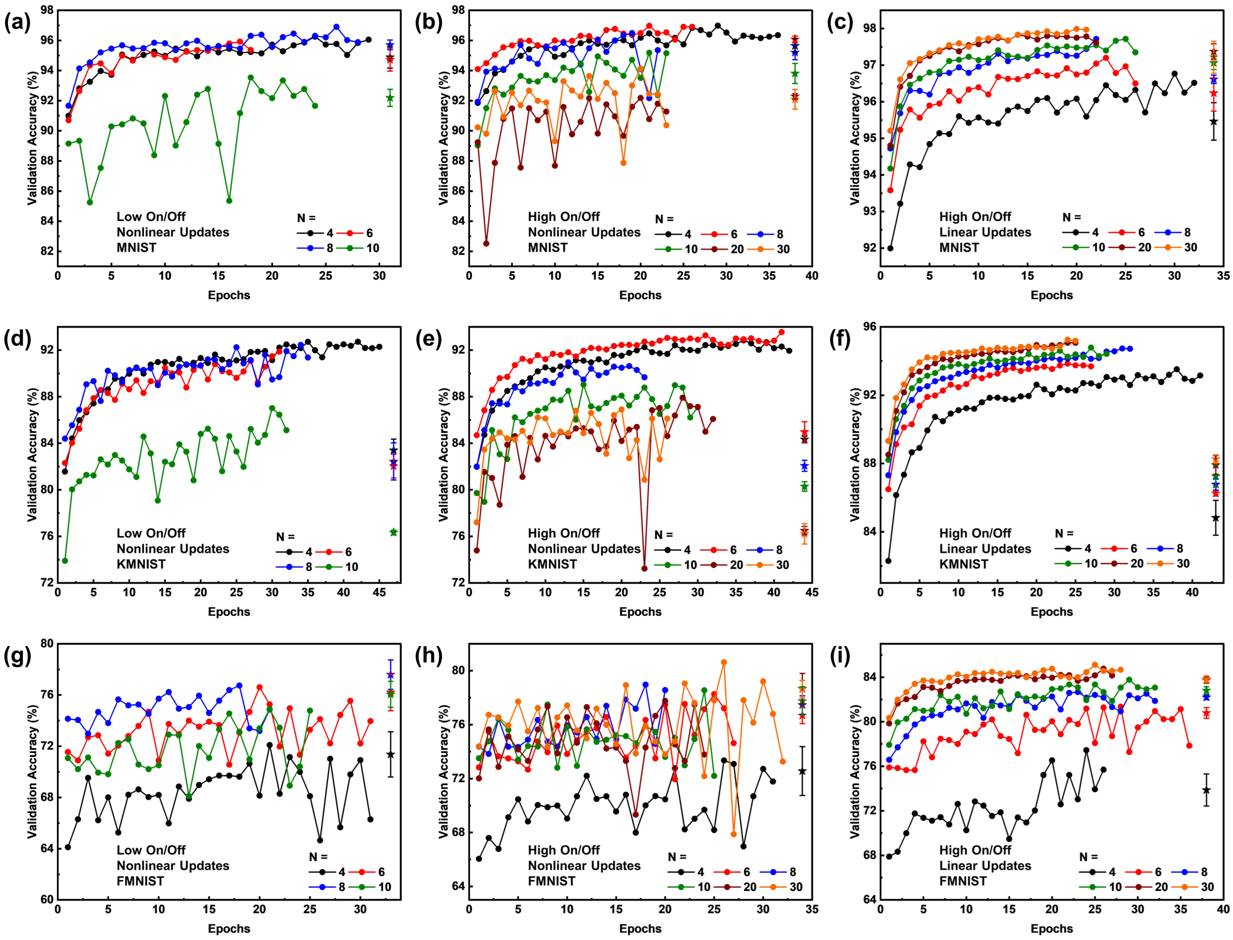

Figure S3: Validation accuracy versus training epochs for different number of conductance states, with each row representing the MNIST, KMNIST, and FMNIST tasks while each column represents the low on/off device with nonlinear weight updates and the high on/off device with both nonlinear and linear weight updates, respectively. This validation accuracies are used to determine convergence behavior and the final testing accuracy via early stopping, as described in the Methods section.

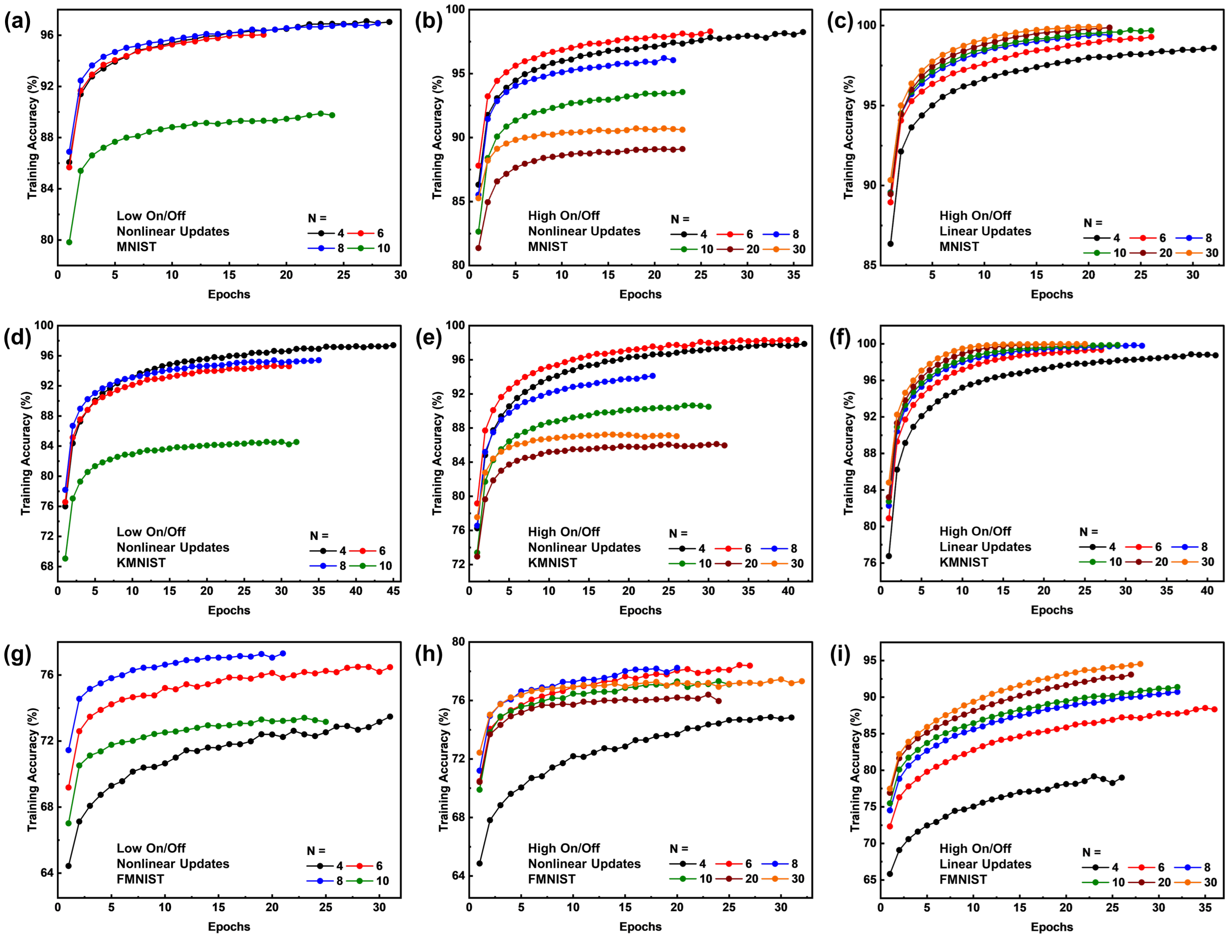

Figure S4: Training accuracy versus training epochs for different number of conductance states, with each row representing the MNIST, KMNIST, and FMNIST tasks while each column represents the low on/off device with nonlinear weight updates and the high on/off device with both nonlinear and linear weight updates, respectively.

Table 1: ANL factor β extracted using Eq. 1 for all ANN simulations, highlighting the increase in nonlinearity with N for fixed-amplitude pulsing and near-zero β achieved using linear pulse engineering

N	**Low on/off Nonlinear**		**High on/off Nonlinear**		**High on/off Linear**	
	β_p	β_d	β_p	β_d	β_p	β_d
4	1.48	0.82	1.33	0.48	0.2	0.19
6	2.02	1.55	1.5	1.28	-0.03	-0.076
8	2.3	2.23	2.02	1.62	-0.08	-0.05
10	3.3	2.66	2.19	1.64	0.04	0.18
20	7.39	5.46	3.24	3.09	0.22	-0.04
30	12.93	7.17	3.82	3.9	0.04	-0.1

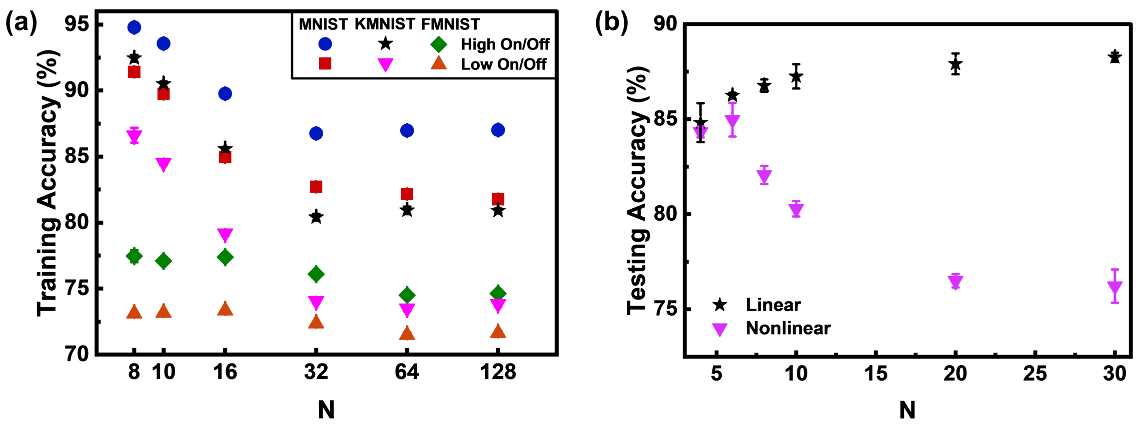

Figure S5: (a) Training accuracy as a function of states for the hypothetical case shown in Fig. 4b, depicting a decrease followed by saturation for MNIST and KMNIST, while FMNIST remains relatively unchanged. (b) ANN classification accuracy for the KMNIST dataset demonstrating improved performance with linear weight updates (~7% at N = 10 and ~12% at N = 30).