
Neural Correlates of Serial Dependence: Synaptic Short-term Plasticity Orchestrates Repulsion and Attraction

Anonymous Author(s)

Affiliation

Address

email

1 A Model Details

2 **Single-layer CANN and two-layer CANN.** The neural dynamics in a single-layer CANN are de-
3 scribed as:

$$\tau \frac{\partial h(\theta, t)}{\partial t} = -h(\theta, t) + \rho \int J(\theta, \theta') u(\theta', t) x(\theta', t) r(\theta', t) d\theta' + I_{\text{ext}}(\theta, t) + \mu_b \xi_b(\theta, t) \quad (1)$$

4 The calculation of $u(\theta, t)$, $x(\theta, t)$, $r(\theta, t)$ were identical to the two-layer CANN in Sec. 2. The
5 time constants for short-term depression (τ_d) and short-term facilitation (τ_f) were set for the STD-
6 dominated CANN and the STF-dominated CANN models, respectively.

7 We considered heterogeneity in intra-layer and inter-layer neuronal connections across different
8 participants by including noise: $\tilde{J}_{KM}(\theta_K, \theta'_M) = J_{KM}(\theta_K, \theta'_M)(1 + \mu_J \xi_J)$, with $K, M \in \{L, H\}$.
9 ξ_J denotes white Gaussian noise of zero mean and unit variance and μ_J the noise strength. We
10 included noise in the neuronal connections of the single-layer CANN similarly. Model parameters
11 are shown in Tab. 1.

12 **Population vector method.** We decoded orientation value from neural activity across CANN layers
13 using the population vector method, expressed as:

$$\theta^d = \frac{\int \theta \langle r(\theta, t) \rangle d\theta}{\int \langle r(\theta, t) \rangle d\theta} \quad (2)$$

14 where $\langle r(\theta, t) \rangle$ is the averaged firing rate of neurons at θ during the cueing period.

15 B Experimental Paradigm and Statistical Methods

16 **Post-cueing paradigm.** Each trial consists of a first stimulus (S_1) lasting 200 ms, an inter-stimulus
17 interval (ISI) lasting 1000 ms, a second stimulus (S_2) lasting 200 ms, a delay period of 3400 ms,
18 a retrieval cue ($S_i, i = 1, 2$) lasting 500 ms, and an inter-trial interval (ITI) of 1000 ms. In the
19 single-layer CANN, S_2 was recalled during the cue phase to examine the influence of S_1 on S_2
20 within a trial. Simulations were performed for a single trial. In the two-layer CANN, the effects
21 within and between trials were examined by recalling S_1 or S_2 during the cue phase in a randomized
22 order across 100 consecutive trials. Simulations were conducted with 20 participants, each with
23 various neuronal synaptic connection strengths as in Appendix A. For each participant, 100 trials
24 were conducted.

25 **Stimulus orientation settings.** Orientations were selected from a uniform distribution over $[-90^\circ,$
26 $90^\circ)$ with a step size of 1° . ΔS refers to the angular difference between the previous stimulus and
27 the current stimulus. To ensure an equal probability of its occurrence, ΔS_{within} and $\Delta S_{\text{between}}$ were

Table 1: Model Parameters

Parameter	Meaning	Value
General Network Parameters		
N	Number of neurons	100
τ	Time constant of synaptic current	0.01 s
μ_J	Neuronal interaction noise strength	0.01
Parameters for STD-dominated single-layer CANN (Fig. 2A-C)		
J^0	Maximum synaptic connection strength	0.13
a	Range of neuronal interaction	0.5
k	Global inhibition strength	0.0018
μ_b	Background noise strength	0.5
τ_d	Time constant of x	3 s
τ_f	Time constant of u	0.3 s
\dot{U}_0	Increment of u produced by a spike	0.5
Parameters for STF-dominated single-layer CANN (Fig. 2D-F)		
J^0	Maximum synaptic connection strength	0.09
a	Range of neuronal interaction	0.15
k	Global inhibition strength	0.0095
μ_b	Background noise strength	0.5
τ_d	Time constant of x	0.3 s
τ_f	Time constant of u	5 s
\dot{U}_0	Increment of u produced by a spike	0.2
Parameters for two-layer network (Fig. 3)		
J_{LL}^0	Max synaptic connection strength (low-to-low)	0.13
a_L	Range of neuronal interaction (low-level)	0.5
k_L	Global inhibition strength (low-level)	0.0018
μ_L	Background noise strength (low-level)	0.5
τ_d^L	Time constant of x (low-level)	3 s
τ_f^L	Time constant of u (low-level)	0.3 s
\dot{U}_0^L	Increment of u produced by a spike (low-level)	0.5
J_{HL}^0	Max synaptic connection strength (low-to-high)	0.02
a_{HL}	Range of neuronal interaction (low-to-high)	0.15
J_{HH}^0	Max synaptic strength (high-to-high)	0.09
a_H	Range of neuronal interaction (high-level)	0.15
k_H	Global inhibition strength (high-level)	0.0095
μ_H	Background noise strength (high-level)	0.5
τ_d^H	Time constant of x (high-level)	0.3 s
τ_f^H	Time constant of u (high-level)	5 s
\dot{U}_0^H	Increment of u produced by a spike (high-level)	0.2
External Input		
α_{sti}	Strength of external stimulus	20
a_{sti}	Spatial scale of external stimulus	0.3
μ_{sti}	Noise strength of external stimulus	0.5
α_{cue}	Strength of external cue	2.5
a_{cue}	Spatial scale of external cue	0.4
μ_{cue}	Noise strength of external cue	1

28 randomly generated with equal probability over $[-90^\circ, 90^\circ]$. Take the first trial in the two-layer
29 CANN as an example, $\theta_{\text{cue}}^{s,1}$ was selected from a uniform distribution over $[-90^\circ, 90^\circ]$ randomly, then
30 $\theta_{\text{uncue}}^{s,1}$ and $\theta_{\text{cue}}^{s,2}$ were calculated:

$$\theta_{\text{uncue}}^{s,1} = \theta_{\text{cue}}^{s,1} + \Delta S_{\text{within}} \quad (3)$$

$$\theta_{\text{cue}}^{s,2} = \theta_{\text{cue}}^{s,1} - \Delta S_{\text{between}} \quad (4)$$

Subsequent stimuli were generated in the same way. For the n th trial, $\theta_{\text{uncue}}^{s,n}$ of the current trial and $\theta_{\text{cue}}^{s,n+1}$ of the subsequent trial were calculated according to the following formulas:

$$\theta_{\text{uncue}}^{s,n} = \theta_{\text{cue}}^{s,n} + \Delta S_{\text{within}} \quad (5)$$

$$\theta_{\text{cue}}^{s,n+1} = \theta_{\text{cue}}^{s,n} - \Delta S_{\text{between}} \quad (6)$$

θ_{cue}^n and θ_{uncue}^n were assigned to S_1 or S_2 randomly, meanwhile the cue index in S_{cue}^n was determined.

Behavioral readout. In a single-layer CANN, external signals are input into the network, and the output is read from the same network using the population vector method. In a two-layer CANN, external signals are input into the lower layer, and the output is read from the higher layer.

Adjustment error. For each trial, we measured the adjustment error as the angular distance between the cued stimulus angle and the decoded response angle.

$$\text{Error} = \theta_{\text{cue}}^d - \theta_{\text{cue}}^s \quad (7)$$

Serial bias analysis. Serial bias was calculated as the averaged adjustment error using a 30° sliding window, as a function of within-trial and between-trial angular differences. In Fig. 3D, E:

$$\Delta S_{\text{within}} = \theta_{\text{uncue}}^{s,n} - \theta_{\text{cue}}^{s,n} \quad (8)$$

$$\Delta S_{\text{between}} = \theta_{\text{cue}}^{s,n} - \theta_{\text{cue}}^{s,n+1} \quad (9)$$

In Fig. 3F, we further analyzed the average error as a function of the memory representation angular difference:

$$\Delta R_{\text{between}} = \theta_{\text{cue}}^{d,n} - \theta_{\text{cue}}^{s,n+1} \quad (10)$$

DoG curve. To measure the amplitude of serial dependence, we fit the error plot with the first derivative of a Gaussian curve (DoG) $G(x) = \sqrt{e}(x/\sigma)A_{\text{DoG}} \exp(-x^2/2\sigma^2)$, where A_{DoG} is the amplitude of the curve peak, σ is the curve width.

Partial no-report paradigm. To test whether serial dependence occurs without a prior reported stimulus, the cue was absent in 40% of trials. Trials following those without a report were used to analyze serial biases.

C Numerical Simulation of u and x

The fitting of x curve in an STD-dominated single-layer CANN. In an STD-dominated single-layer CANN, x is fitted by the following equation:

$$x(\theta, t) = 1 - A_x^1(t) \exp[(\theta - \theta_1^s)^2/2a^2] - A_x^2(t) \exp[(\theta - \theta_2^s)^2/2a^2] \quad (11)$$

For example, in Fig. 2B, we set $\theta_1^s = -30^\circ$, $\theta_2^s = 0^\circ$ as the initial guess value. The amplitudes $A_x^1(t)$ and $A_x^2(t)$ decayed from the early delay period (Fig. S1A) to the late delay period (Fig. S1B). We calculated these parameters from the time point after S_2 disappeared ($t_0 = 1.5$ s) to the time point before the cue period ($t_e = 4.8$ s), shown as dots in Fig. S1C, and found that they decayed exponentially with a time constant τ_d (curve in Fig. S1C), described as:

$$A_x^i(t) = A_x^i(t_0) \exp[(t - t_0)/\tau_d], i = 1, 2 \quad (12)$$

where $A_x^1(t_0) = 0.46$, $A_x^2(t_0) = 0.68$.

The above analysis is applicable to other combinations of θ_1^s and θ_2^s .

Table 2: The fitting parameters of x curve

t (s)	θ_1^s ($^\circ$)	θ_2^s ($^\circ$)	A_x^1	A_x^2	a
2.0	-32.72	6.13	0.39	0.57	0.31
4.5	-32.72	6.13	0.17	0.25	0.31

The fitting of u curve in a STF-dominated single-layer CANN. In an STF-dominated single-layer CANN, the u curve was fitted by the following equation:

$$u(\theta, t) = A_u^1(t) \exp[(\theta - \theta_1^s)^2/2a^2] + A_u^2(t) \exp[(\theta - \theta_2^s)^2/2a^2] \quad (13)$$

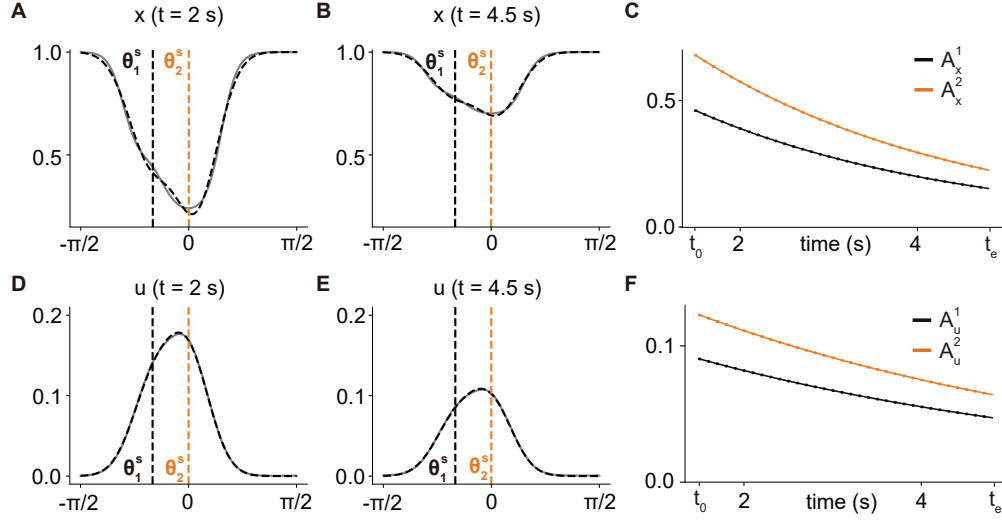


Figure S1: The fitting of x and u curves in a single-layer CANN. (A-C) x in an STD-dominated single-layer CANN. (D-F) u in an STF-dominated single-layer CANN. (A, D) Simulation (gray curve) and fitting results (dotted curve) in the early delay period ($t = 2.0$ s). (B, E) Simulation (gray curve) and fitting results (dotted curve) the late delay period ($t = 4.5$ s). (C, F) Exponential decay during the delay period. Dots represent the simulation results, and curves represent the exponential fitting results.

In Fig. 2E, we fixed $\theta_1^s = -30^\circ$ and $\theta_2^s = 0^\circ$. The amplitude parameters $A_u^1(t)$ and $A_u^2(t)$ decayed from the early delay period (Fig. S1D) to the late delay period (Fig. S1E). We calculated these parameters from the time point after S_2 disappeared ($t_0 = 1.5$ s) to the time point before the cue period ($t_e = 4.8$ s), shown as dots in Fig. S1F, and found that they decayed exponentially with the time constant τ_f (curve in Fig. S1F), described as:

$$A_u^i(t) = A_u^i(t_0) \exp[(t - t_0)/\tau_f], i = 1, 2 \quad (14)$$

where $A_u^1(t_0) = 0.09$, $A_u^2(t_0) = 0.12$.

The above analysis is applicable to other combinations of θ_1^s and θ_2^s .

Table 3: The fitting parameters of u curve

t (s)	θ_1^s ($^\circ$)	θ_2^s ($^\circ$)	A_u^1	A_u^2	a
2.0	-30.00	0.00	0.08	0.11	0.31
4.5	-30.00	0.00	0.05	0.07	0.31

The fitting of x_L and u_H curve in a two-layer CANN. In a two-layer CANN, the x_L curve in the $(n + 1)th$ trial was fitted by the following equation:

$$x_L(\theta_L, t) = 1 - A_x^1(t) \exp[(\theta_L - \theta_1^{s, n+1})^2 / 2a_x^2] - A_x^2(t) \exp[(\theta_L - \theta_2^{s, n+1})^2 / 2a_x^2] - A_x^{\text{cue}}(t) \exp[(\theta_L - \theta_{\text{cue}}^{s, n})^2 / 2a_x^2] \quad (15)$$

In Fig. 4B, we set $\theta_{\text{cue}}^n = 20^\circ$, $\theta_1^{n+1} = -30^\circ$, and $\theta_2^{n+1} = 0^\circ$ as the initial guess values. The amplitude parameters $A_x^{\text{cue}}(t)$, $A_x^1(t)$ and $A_x^2(t)$ decayed from the early delay period (Fig. S2A) to the late delay period (Fig. S2B). We calculated the fitted parameters from the time point after S_2^{n+1} disappeared ($t_0 = 7.8$ s) to the time point before the cue period ($t_e = 11.1$ s), shown as dots in Fig. S2C, and found that they decayed exponentially with a time constant of τ_d (curve in Fig. S2C), which can also be described as:

$$A_x^i(t) = A_x^i(t_0) \exp[(t - t_0)/\tau_d], i = 1, 2, \text{cue} \quad (16)$$

Table 4: The fitting parameters of x_L curve

t (s)	$\theta_{\text{cue}}^{d,n}$ ($^\circ$)	$\theta_1^{s,n+1}$ ($^\circ$)	$\theta_2^{s,n+1}$ ($^\circ$)	A_x^{cue}	A_x^1	A_x^2	a_x
8.3	39.13	-35.41	3.26	0.19	0.37	0.57	0.33
10.8	39.13	-35.41	3.26	0.08	0.16	0.25	0.33

where $A_x^{\text{cue}}(t_0) = 0.23$, $A_x^1(t_0) = 0.44$, $A_x^2(t_0) = 0.68$.

The above analysis is applicable to other combinations of θ_{cue}^n , θ_1^{n+1} and θ_2^{n+1} .

u_H curve in the $(n+1)$ th trial can be fit by the following equation:

$$u_H(\theta_H, t) = A_u^1(t) \exp \left[(\theta_H - \theta_1^{s,n+1})^2 / 2a_u^2 \right] + A_u^2(t) \exp \left[(\theta_H - \theta_2^{s,n+1})^2 / 2a_u^2 \right] + A_u^{\text{cue}}(t) \exp \left[(\theta_H - \theta_{\text{cue},H}^{d,n})^2 / 2a_u^2 \right] \quad (17)$$

In Fig. 3C and Fig. 4B, we fixed $\theta_{\text{cue}}^n = 20^\circ$, $\theta_1^{n+1} = -30^\circ$, and $\theta_2^{n+1} = 0^\circ$. $A_u^{\text{cue}}(t)$, $A_u^1(t)$ and $A_u^2(t)$ decayed from the early delay period (Fig. S2D) to the late delay period (Fig. S2E). We calculated these parameters from the time point after S_2^{n+1} disappeared ($t_0 = 7.8$ s) to the time point before the cue period ($t_e = 11.1$ s), shown as dots in Fig. S2F, and found that they decayed exponentially with the time constant τ_f (curve in Fig. S2F), which can also be described as:

$$A_u^i(t) = A_u^i(t_0) \exp[(t - t_0)/\tau_f], i = 1, 2, \text{cue} \quad (18)$$

where $A_u^{\text{cue}}(t_0) = 0.22$, $A_u^1(t_0) = 0.14$, $A_u^2(t_0) = 0.08$.

The above analysis is applicable to other combinations of θ_{cue}^n , θ_1^{n+1} and θ_2^{n+1} .

Table 5: The fitting parameters of u_H curve

t (s)	$\theta_{\text{cue}}^{d,n}$ ($^\circ$)	$\theta_1^{s,n+1}$ ($^\circ$)	$\theta_2^{s,n+1}$ ($^\circ$)	A_u^{cue}	A_u^1	A_u^2	a_u
8.3	20.00	-30.00	0.00	0.20	0.12	0.07	0.31
10.8	20.00	-30.00	0.00	0.31	0.08	0.04	0.31

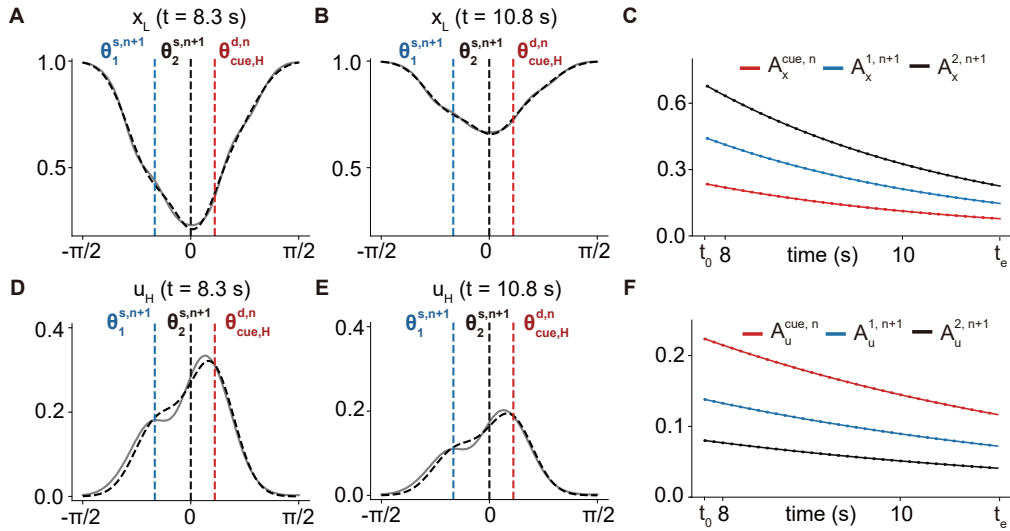


Figure S2: The fitting of x_L and u_H curves in a two-layer CANN. (A-C) x_L in an STD-dominated low layer. (D-F) u_H in an STF-dominated high layer.

Computer resources. We used Python 3.11.9 and Brainpy 2.6.0.post20240918 for simulations. All experiments were conducted on a consumer-grade desktop computer (AMD Ryzen 9 9950X, 32GB DDR5 RAM) with hour-scale runtimes.