Supplementary Materials for

A spatiotemporal mechanism of visual attention: Superdiffusive motion and theta oscillations of neural population activity patterns

Guozhang Chen and Pulin Gong*

*Corresponding author. Email: pulin.gong@sydney.edu.au

Published 22 April 2022, Sci. Adv. 8, eabl4995 (2022)
DOI: 10.1126/sciadv.abl4995

The PDF file includes:

Figs. S1 to S11
Legends for movies S1 to S7

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S7
Snapshots of spiking patterns in State I and III. (A) Black dots denote spontaneous spikes in the past 10 ms period in State I, showing a localized activity pattern with a regular, smooth propagation path. Circle represents one standard deviation of the fitted 2D Gaussian firing rate profile. The line with color gradient shows the trajectory of the pattern in the previous 200 ms. (B) Same as in (A) but for State III. For State III (the asynchronous state), no structured patterns are formed.
Fig. S2.

Dynamics of the attention activity pattern account for neural temporal fluctuations illustrated on a different trial. (A) Snapshots of firing rates during a 10 ms period at different time points. Red circles denote the 1 s.d. of Gaussian profile of two input objects. One object is in the center and another one is in one of the corners. Due to the periodic boundary conditions of the two-dimensional circuit model, the latter appears to occupy one quarter of each corner. (B) Spiking raster showing the spike times of a sub-population of 100 excitatory neurons where the center object is added to the circuit. This indicates transitions between episodes of vigorous and faint spiking. (C) Same as in (B) but for the corner area where the corner object is added to the circuit. (D) The time course of MUA in the circuit center. (E) Time course of MUA in the corner of the circuit. (F) Snapshots of gamma-band LFP (30-80 Hz) amplitude at time points in (A). Amplitude is extracted by using the Hilbert transform. (G) Normalized time-frequency diagram...
of LFPs recorded in the center position. White solid and dashed lines denote 30-80 Hz bandpass LFPs (gamma band) and 3-10 Hz bandpass LFPs (theta band), respectively. Black line represents the raw LFP. The amplitudes of raw, gamma- and theta-band LFP are normalized. (H) Same as in (G) but for the corner position.
Fig. S3.

**Pattern dynamics give rise to theta-band neural oscillations.** (A) Averaged power spectrum of LFPs when two inputs are added to the neural circuit. An approximately straight line on this log-log plot indicates that the power spectrum follows \( \text{Power} \propto 1/f^\beta \), with the exponent \( \beta = 2.5 \). (B-D) Averaged power spectra of the trajectories of spontaneous activity patterns, MUAs, and LFPs, respectively. Approximately straight lines on these log-log plots indicate that the power spectrum follows \( \text{Power} \propto 1/f^\beta \), with the exponent \( \beta = 1.5, 2.2, 2.1 \), respectively. 1/f activity is disentangled from the oscillatory components for considering distinct peaks that exceed the estimated background 1/f spectrum. The shaded areas represent the s.e.m. calculated across 500 trials.
Fig. S4.

Simulation of reaction time. (A) Temporal averaged MUA with a window of 10 ms of the center area as a function of time. One input object (Eq. 13) is first added to the center of the circuit, which is followed by another object (Eq. 13) to a corner of the circuit in addition 500 ms later. As pointed out in (9), such a sequential presentation of the two objects resets attention to one of two objects (i.e. the object in the corner). A target, which is a 50% increase of the contrast level of the center object, is then presented after a randomized period within up to 1000 ms (940 ms in this figure) following the onset of the second object. Reaction time is the period from target onset to the response. The decrease of MUA activity after adding the second object is consistent with experimental results (9); in our model, this decrease is caused by lateral inhibitory interactions between the neurons in the center and those in the corner. Note that in our circuit model, the MUA activity increases after the target onset; this is also consistent with experimental studies (9). (B) Reaction time as a function of target onset time. (C) Distribution of
responses at different phases of MUA. We obtain the phases by applying the Hilbert transform to the MUA time series. (D) Peak frequency of reaction time as the function of the sampling rate (median). As in Fig. 5E, the contrast level is increased from 0 to 2.
Global, uniform object added to the circuit. We add global, uniform stimulus as external current, $I_{(stim)}^E$, to the circuit. $I_{(stim)}^E(t) = \mu H(t - t_{onset})$. The first and the second rows show the properties of the attention activity pattern when $\mu = 0.001$, and 0.01, respectively. (A) Averaged power spectrum of LFP, averaged over 500 trials. The shaded area represents the standard error of mean (s.e.m). (B) Same as in (A) but for the MUA activity. (C) Averaged power spectrum of the trajectories of the attention activity pattern on the X coordinate (Y coordinate is similar), averaged over 500 trials. The pattern center is estimated by fitting a 2D Gaussian function (Eq. 17). (D) Mean square displacement (MSD) of trajectories of the attention activity pattern demonstrates superdiffusion. Dashed red line shows $MSD(\tau) \propto \tau^{1.2}$. Shaded area represents the standard error of mean (s.e.m.) across 500 trials. (E) Distribution of the increments over 100 s. Red line denotes a fitted symmetric Lévy $\alpha$-stable distribution. Inset: Distribution of the displacement of the attention activity patterns shows a power-law tail. The red dashed reference line has an exponent of -2.2. P. Probability; D., Density; Dis., Displacement.
**Fig. S6.**

**Objects with higher contrast levels attract more attention.** We increase the contrast level from 0 to 2 for one object. The dwelling time increases linearly as the contrast level increases. The dwelling time is quantified by the duration that the attention activity pattern spends in one object location before switching to the other. The error bar represents the standard deviation over 100 trials.
Fig. S7.

Attention sampling of multiple objects. (A–E) From the first to the third row, neural dynamics and attention pattern dynamics of network with three, four, and five objects are shown. (A) Averaged power spectrum of the trajectories of the attention activity pattern on the X coordinate (Y coordinate is similar), averaged over 500 trials. The pattern center is estimated by fitting a 2D Gaussian function (Eq. 17). (B) Averaged power spectrum of MUA, averaged over 500 trials. The shaded area represents the standard error of mean (s.e.m.). (C) Mean square displacement (MSD) of trajectories of the attention activity pattern demonstrates superdiffusion. Dashed red line shows $MSD(\tau) \propto \tau^{1.2}$. Shaded area represents the standard error of mean (s.e.m.) across 500 trials. (D) Distribution of the increments over 100 s. Red line denotes a fitted symmetric Lévy $\alpha$-stable distribution. Inset: Distribution of the displacement of the attention activity patterns shows a power-law tail. The red dashed reference line has an exponent of -2.2. P. Probability; D., Density; Dis., Displacement. (E) Colormap represents the logarithmic probability of patterns visiting one cell of the grid. The dashed and solid circles represent the stimulus locations within
the 1 or 2 s.d. of the stimulus object’s Gaussian profile, respectively. (F) Sampling rate as a function of object number. (G) The proportion of irrelevant sampling as a function of object number.
Scan paths emerging from our circuit model exhibits similar properties as human scan paths. (A) Distribution of fixation duration of the attention activity pattern in our model (top) and human scan paths (MIT1003, data bottom). The fixation duration is defined as the duration that the attention activity pattern spends in one object location before switching to another one. Inset: Distribution of sampling rates. (B) Averaged power spectrum of human scan paths on the X coordinate (Y coordinate is similar), averaged over 15 subjects viewing 1003 images. The power spectrum follows $Power \propto 1/\omega^\beta$, with the exponent $\beta = 2.15$. The shaded area represents the s.e.m.. (C) Distribution of the displacement of human eye movement in 4.167 ms (the temporal resolution of the MIT1003 eye-tracking data) shows a heavy tail. The bin size is around 19 pixels. Red dashed reference line has an exponent of -2.2. (D) Cross correlations between the human attention maps and the simulated attention maps as a function of I-E ratio, $\zeta$. The error bar represents the 10 s.e.m calculated over 500 trials.
Fig. S9.
The focus of attention in the classical WTA follows the order from the largest to the smallest saliency values without variability when sampling natural scenes. (A) Black line denotes the scan path generated by the WTA model with the saliency map (the background colormap) produced by the method of Judd et al (45). Numbers represent the temporal order of the scan paths. Two rows illustrate the scan paths on two different trials, showing the lack of attention sampling variability across trials. (B-D) Same as in (A) but for saliency map produced by the AWS (46), Itti et al (15), and GBVS (47) methods, respectively.
Fig. S10.

**Gamma bursts only emerge in State II.** The spiking neural circuits with $\zeta/\zeta_C = 0.7$ and $\Delta g_k = 0.003 \text{ nS}$ (A), $\zeta/\zeta_C = 1.3$ and $\Delta g_k = 0.003 \text{ nS}$ (B), $\zeta/\zeta_C = 1$ and $\Delta g_k = 0.001 \text{ nS}$ (C), and $\zeta/\zeta_C = 1$ and $\Delta g_k = 0.008 \text{ nS}$ (D) do not exhibit clear gamma bursts shown in Fig. 3C with the default parameters ($\zeta/\zeta_C = 1$ and $\Delta g_k = 0.003 \text{ nS}$) of State II.
Fig. S11.
Proportion of irrelevant sampling as a function of the I-E ratio $\zeta$, showing the exploration degree can be modulated by changing the I-E ratio in State II. The task-irrelevant location is defined either by 1 or 2 s.d. of object as in Fig. 5C.
Movie S1
Spatiotemporal dynamics of spiking activity patterns in the neural circuit model with the I-E ratio $\zeta = 3.03$ without inputting objects. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile.

Movie S2
Spatiotemporal dynamics of spiking activity patterns in the neural circuit model with the I-E ratio $\zeta = 4.38$ without inputting objects. Blue dots represent the spikes.

Movie S3
Spatiotemporal dynamics of spiking activity patterns in the neural circuit model with the I-E ratio $\zeta = 3.31$ without inputting objects. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile.

Movie S4
Spatiotemporal dynamics of the attention activity pattern sampling two objects. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile. The black circles denote one standard deviation of the inputs.

Movie S5
Attention activity pattern sampling a natural scene. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile.

Movie S6
Spatiotemporal dynamics of spiking activity patterns in the neural circuit model with the I-E ratio $\zeta = 4.38$ when two inputs are added to the circuit. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile. The black circles denote one standard deviation of the inputs.

Movie S7
Spatiotemporal dynamics of spiking activity in the neural circuit model with the I-E ratio $\zeta = 3.03$ when two inputs are added to the circuit. Blue dots represent the spikes and red circle represents one standard deviation of the fitted 2D Gaussian firing rate profile. The black circles denote one standard deviation of the inputs.