Subdiffusive Dynamics of Bump Attractors: Mechanisms and Functional Roles

Yang Qi[1], **Michael Breakspear**[2], **Pulin Gong**[1, 3]

[1]School of Physics, University of Sydney, Sydney, NSW, Australia.

[2]Queensland Institute of Medical Research, Brisbane, QLD, Australia

[3]Sydney Medical School, University of Sydney, Sydney, NSW, Australia.

Keywords: Subdiffusion, long-range temporal correlation, working memory

Abstract

Bump attractors are localized activity patterns that can self-sustain after stimulus presentation, and they are regarded as the neural substrate for a host of perceptual and cognitive processes. One of the characteristic features of bump attractors is that they are neutrally stable, so that noisy inputs cause them to drift away from their initial locations, severely impairing the accuracy of bump location-dependent neural coding. Previous modeling studies of such noise-induced drifting activity of bump attractors

have focused on normal diffusive dynamics, often with an assumption that noisy inputs are uncorrelated. Here we show that long-range temporal correlations and spatial correlations in neural inputs generated by multiple interacting bumps cause them to drift in an anomalous subdiffusive way. This mechanism for generating subdiffusive dynamics of bump attractors is further analyzed based on a generalized Langevin equation. We demonstrate that subdiffusive dynamics can significantly improve the coding accuracy of bump attractors, since the variance of the bump displacement increases sub-linearly over time, and is much smaller than that of normal diffusion. Furthermore, we re-analyze existing psychophysical data concerning the spread of recalled cue position in spatial working memory tasks, and show that its variance increases sub-linearly with time, consistent with subdiffusive dynamics of bump attractors. Based on the probability density function of bump position, we also show that the subdiffusive dynamics result in a long-tailed decay of firing rate, greatly extending the duration of persistent activity.

1 Introduction

Localized, self-sustained activity patterns often arise in neural network models with local excitation and lateral inhibition when responding to transient stimuli. Such persistent activity patterns can be described by bell-shaped tuning curves with respect to the positions of the stimuli, so they are often referred to as bump states or bump attractors. Bump attractors have been widely studied in both firing-rate and spiking neural network models (Amari, 1977; Ben-Yishai et al., 1995; Pouget et al., 1998; Compte

et al., 2000; Wu et al., 2008). These self-sustained bumps can be centered at any continuously varying locations; therefore they can be used to encode continuous analogue quantities. Due to this property, bump attractors have been proposed to be the neural mechanism underlying a host of perceptual and cognitive processes, including stabilization of eye gaze (Seung et al., 2000), orientation tuning in visual systems (Ben-Yishai et al., 1995), direction selection in head direction cells (Zhang, 1996), path integration in hippocampus (McNaughton et al., 2006), and spatial working memory (Wang, 2001; Compte et al., 2000). Notably, a recent experimental study has found physiological evidence supporting the hypothesis that spatial working memory is represented by bump attractors in the prefrontal cortex (Wimmer et al., 2014).

Bump attractors are neutrally stable, so that they drift away from their initial locations in the presence of stochastic effects. Noise-induced drifting behavior of bump attractors exhibits normal diffusive dynamics when noise is white; such normal diffusive dynamics of bump attractors have been studied in several different frameworks (Compte et al., 2000; Miller, 2006; Wu et al., 2008; Polk et al., 2012; Kilpatrick et al., 2013). Since the location of a bump can be used to encode information, diffusion of the bump results in errors in such bump location-dependent coding; for instance, in the context of spatial working memory, diffusion of a bump degrades the memory stored at its initial location, resulting in inaccuracies in the behavioral report of remembered locations (White et al., 1994; Ploner et al., 1998; Wimmer et al., 2014).

To reduce such diffusion-induced coding errors of bump attractors, it is important to make the diffusion coefficient of the corresponding diffusive process as small as possible. Recently, Kilpatrick et al. (2013) show that the diffusion coefficient can be

reduced by incorporating spatial heterogeneity to the network connectivity. The diffusion of bump attractors in this setting is described by a random walk in a periodic potential well, resulting in an effective diffusion coefficient smaller than a free random walk. In another study, Polk et al. (2012) show that noisy inputs with correlations distributed globally across populations are able to reduce the diffusion coefficient. Nevertheless, in these modeling studies, bump attractors still move in a normal diffusive way, since noisy input fluctuations considered in these models are assumed to be white noise without temporal correlations. However, neural fluctuations exhibit temporal and spatial correlations; particularly, long-range temporal correlations have been found in recordings of nearly all levels of neural systems, ranging from spike trains of individual neurons (Bhattacharya et al., 2005; Teich et al., 1990), local field potentials (Destexhe et al., 1999), to EEG and MEG recordings (Linkenkaer-Hansen et al., 2001; Freeman et al., 2003; Gong et al., 2003), and fMRI data (Leopold et al., 2003). Spatial correlations of neural fluctuations have been found to be distance-dependent, *i.e.*, nearby neurons tend to be correlated much more than distant neurons (Smith and Sommer, 2013; Ecker et al., 2014). Despite the ubiquity of temporal and spatial correlations in neural activities, their effects on the dynamics of bump attractors remain unknown.

In this study, we investigate the effect of long-range temporal correlations in combination with spatial correlations in neural input fluctuations generated by multiple, interacting bumps in a simple spiking neural circuit, and find that such correlations causes the bumps to undergo anomalous subdiffusion. This mechanism is further validated in a biologically more plausible, conductance-based spiking neural circuit model. For a bump undergoing subdiffusion, the variance of its displacement relative to its initial

position, hence the variance of its errors, increases sub-linearly with time, instead of increasing linearly as in normal diffusion. Such a sub-linear increase is characterized by a diffusion exponent less than one. Due to this property, the subdiffusive dynamics of bump attractors provide a more effective way for improving the coding accuracy than previous methods by reducing the diffusion coefficient of normal diffusion.

To provide further theoretical insights into the origin of the anomalous subdiffusion of bump attractors, we regard individual bumps as point particles and apply a generalized Langevin equation (GLE) with a long-range temporally correlated force to describe the motion of the bumps. This approach is particularly suited for describing anomalous diffusion involving multiparticle interactions (Sokolov, 2012), analogous to our case of multiple bumps. By solving the GLE analytically, we find that the mean-squared displacement of bump position increases sub-linearly with time, consistent with numerical simulations of the spiking neural circuit models. Based on the corresponding probability density function of bump position derived from the GLE and that obtained from the spiking neural network, we find two additional characteristic features of bump attractors with subdiffusive dynamics, including a slow decay in the firing rate of persistent activity and a slow decay in the rate of correct recall of the stimulus cue based on a simple threshold read-out criterion.

2 A simple spatially extended spiking neural circuit model

In this study we consider a simple, spatially extended spiking neural circuit model, in which the essential dynamics of individual neurons is described by three states: the

spiking, the resting, and the refractory states (Gong and Robinson, 2012). The numerical and the analytical tractability of this type of model facilitate an understanding of the mechanisms of spatio-temporal dynamics in neural systems. In a similar spirit, even two-state network models have been widely used to illustrate principles of emergent neural computation by attractors (Hopfield, 1982), up and down states in studies of repeating patterns of cortical activity (Roxin et al., 2008), statistical properties of populations of neurons (Buice and Cowan, 2007), and avalanches and noisy limit cycles or quasi-cycles in neural systems (Wallace et al., 2011; Benayoun et al., 2010).

We consider a two-dimensional lattice of $n \times n$ neurons, each labeled by integer coordinates (i, j). The state of a neuron located at (i, j) at time t is denoted by $u_{ij}(t)$, which can be one of the firing, resting, or refractory states. The neurons are coupled to each other with local excitation and lateral inhibition, a commonly used coupling structure in neural modeling. The coupling strength $W_{ij,kl}$ between two neurons located at (i, j) and (k, l) is constructed from a distance dependent 'Mexican hat' function:

$$W'_{ij,kl} = C_E \exp\left(-\frac{d^2_{ij,kl}}{d^2_E}\right) - C_I \exp\left(-\frac{d^2_{ij,kl}}{d^2_I}\right), \tag{1}$$

where $C_E = 0.4$, $C_I = 0.1$, $d_E = 3.74$, $d_I = 6.55$ and $d_{ij,kl} = |(i, j) - (k, l)|$ is the Euclidean distance between the neurons located at (i, j) and (k, l). The interactions between neurons are excitatory with $W'_{ij,kl} \geq 0$ when $d_{ij,kl} \leq d_0$, and inhibitory with $W'_{ij,kl} < 0$ when $d_{ij,kl} > d_0$, where $d_0 = 5.1$. Since the interactions between neurons in the brain are finite-ranged, the connections between neurons are constrained within $d_{ij,kl} \leq d_m$ by setting $W'_{ij,kl} = 0$ for $d_{ij,kl} > d_m$, where we set $d_m = 15$. To enable the

excitatory and the inhibitory coupling strengths to be changed without affecting their spatial ranges, we scale the excitatory and the inhibitory parts separately. The final coupling strength $W_{ij,kl}$ is: $W_{ij,kl} = W_E W'_{ij,kl}$ for $d_{ij,kl} \leq d_0$ and $W_{ij,kl} = W_I W'_{ij,kl}$ for $d_{ij,kl} > d_0$, where $W_E = 8.82$ and $W_I = 69.8$ are the excitatory and the inhibitory weights respectively. There is no coupling to a neuron itself, which means $W_{ij,ij} = 0$. Note that the results presented in this study are not sensitive to changes in these parameter values. Periodic boundary conditions are applied to the two-dimensional lattice.

The state of each neuron changes at discrete time steps according to the following rules. (1) A neuron in the firing state sends outputs to other connected neurons and then transitions to the refractory state at the next time step. (2) A neuron in the resting state transitions to the firing state at the next time step if the total input it receives from other neurons is greater than a threshold value; otherwise it remains at rest. (3) A neuron in the refractory state returns to the resting state with a probability p, otherwise remaining unchanged. The randomness of the transition is consistent with the fact that noise is common in neural systems; for instance, neural firing activity depends on ionic channels whose activity is stochastic in nature (Faisal et al., 2008). Assigning the firing, the resting, and the refractory states with values of 1, 0 and -1 respectively, the full dynamics of the neural network can be described as,

$$u_{ij}(t+1) = \begin{cases} H[I_{ij}(t) - 1], & \text{if } u_{ij}(t) = 0, \\ -1, & \text{if } u_{ij}(t) = 1, \\ 0 \text{ with probability } p, \text{ otherwise } -1, & \text{if } u_{ij}(t) = -1, \end{cases} \tag{2}$$

where $H(x)$ is the Heaviside step function with $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ otherwise. The total synaptic input $I_{ij}(t)$ received by a neuron located at (i, j) is,

$$I_{ij}(t) = \sum_{k,l} W_{ij,kl}\delta[u_{kl}(t) - 1], \tag{3}$$

where $W_{ij,kl}$ is the coupling strength and $\delta(x)$ is the delta function with $\delta(0) = 1$ and $\delta(x) = 0$ for all $x \neq 0$.

3 Subdiffusive dynamics of localized bumps

Spatially localized persistent activities, also known as 'bump states', can emerge from the network in response to transient external stimuli. Depending on the value of the parameter p, either bumps ($p > 0.5$) or crescent-shaped propagating waves ($p < 0.5$) can emerge from the network (Gong and Robinson, 2012). In this study, we focus on the dynamics of bumps, so we set $p = 0.7$ for all subsequent studies in this paper (similar results can be obtained for any $0.5 < p < 1$). A bump state can be induced by injecting spikes to a small patch (3-by-3) of neurons at arbitrary locations, which persists over time after the stimulus is removed. We evoke $N = 8$ such bumps in the network; a snapshot of the network activity is shown in Fig. 1A. Due to the random nature of the model, *i.e.*, the random transition from the refractory state to the resting state, the bumps diffuse away from their initial positions along the network.

To investigate the diffusive dynamics of the bumps, we calculate the position of each bump and then trace its trajectory over time. The bump position is estimated using the population vector, $r(t) = \frac{\sum_{ij} S_{ij}(t) r_{ij}}{\sum_{ij} S_{ij}(t)}$, where $r_{ij} = (i, j)$ is a location on the network,

$S_{ij}(t)$ is the spike count of an ij-th neuron within the bump over a time window of $[t, t+1]$ (Takeda and Funahashi, 2004; Georgopoulos et al., 1986; Compte et al., 2000). Figure 1B shows the trajectory of a bump undergoing random motion. This seemingly random motion of the bump can be characterized by its mean-squared displacement (MSD),

$$\sigma^2(t) = \langle |r(t) - r(0)|^2 \rangle, \tag{4}$$

where $r(t)$ is the bump position at time t and $r(0)$ is the initial position. The firing rate of the bump is 0.29 spikes per time step in our model, whereas it is around 30 Hz for persistent activity in working memory (Wimmer et al., 2014); the unit time step of our model therefore corresponds to $t_0 = 0.01$ s, which allows us to present our results in the unit of second. In this study, we focus on the time scale less than 10 s, over which working memory typically operates (Funahashi et al., 1989; Wimmer et al., 2014). We compare the MSD for two cases, *i.e.*, one case in which there is only one bump moving freely on the network ($N = 1$) and another case where the network is saturated with $N = 8$ bumps (Fig. 1C). By fitting the slopes of the MSD on a log-log scale we find distinctive results: when $N = 1$, the MSD increases linearly with time $\sigma^2(t) \sim t^\beta$ with a diffusion exponent of $\beta \approx 1.03$, indicating the bump moves in a normal diffusive way; however when $N = 8$, the MSD of each bump scales sub-linearly over time with a diffusion exponent of $\beta \approx 0.75 < 1$, indicating the bump has anomalous subdiffusive dynamics (Metzler and Klafter, 2000). As described in Eq. 4, the MSD quantifies the mean-squared difference between the current bump position $r(t)$ and the initial bump position $r(0)$. Because bump positions are often used to encode continuous variables such as spatial locations in working memory, the MSD is also equal to the variance of

the coding error (Compte et al., 2000). The subdiffusive dynamics thus indicate the variance of the coding error increases with time in a sub-linear way.

We now show that subdiffusive dynamics of a bump attractor provide an effective way for improving its coding accuracy. Consider a bump undergoing normal diffusion with its MSD, or coding variance, described by $\sigma^2(t) = Dt/t_0$, where D is the diffusion coefficient and $t_0 = 0.01$ s is the time step in our model. Existing theoretical studies achieve a smaller variance by reducing the diffusion coefficient (Kilpatrick et al., 2013; Polk et al., 2012), that is, $\sigma^2(t) = \eta Dt/t_0$, with $0 < \eta < 1$. In contrast, we do so by reducing the diffusion exponent, that is, $\sigma^2(t) = D(t/t_0)^\beta$, with $0 < \beta < 1$. The ratio between the two coding variances is,

$$h(\beta, \eta) = \eta(t/t_0)^{1-\beta}. \tag{5}$$

For subdiffusion to be more effective than normal diffusion at improving coding accuracy of bump attractors, it requires $h(\beta, \eta) > 1$. In particular, we are interested in whether this can be satisfied for the neurophysiological time scale for working memory, which is typically below 10 s. Figure 1D shows the curves corresponding to $h(\beta, \eta) = 1$ when $t = 1$ s and $t = 10$ s respectively, where the region above each of the curves corresponds to $h(\beta, \eta) > 1$. This condition is almost always satisfied if both β and η are reduced by the same amount, as indicated by the dotted line in Fig. 1D above the curves. To further illustrate this, consider the MSD in our model with $\beta = 0.75$, which is less than normal diffusion for all $t > 1$ s even if η is reduced to a value as low as 0.32. This result therefore shows that subdiffusive dynamics is very effective for improving

the coding accuracy of bump attractors.

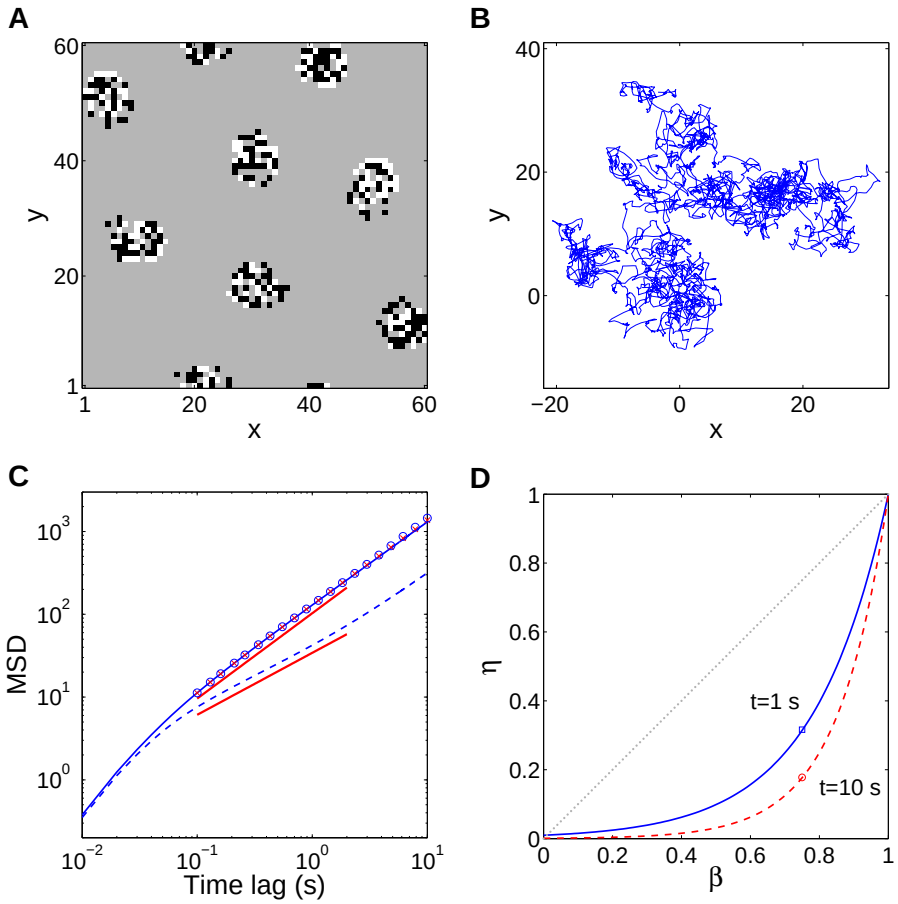

Figure 1: **The MSD of bump positions.** (A) A snapshot of the network activity showing a total number of $N = 8$ localized bumps undergoing random motion. Each pixel in white, gray, or black represents a neuron in the spiking, the resting, or the refractory state, respectively. (B) The trajectory of a bump over time resembling that of a random walk. (C) Mean-squared displacement (MSD) of bump position. The solid curve is the MSD for a single bump ($N = 1$) with normal diffusive dynamics and the dashed curve is the MSD for multiple bumps ($N = 8$) with subdiffusive dynamics. The short lines show the fitted slopes of the MSDs on a log-log scale, with $\beta \approx 1.03$ and $\beta \approx 0.75$ for the respective fittings. The crosses and the circles indicate the MSD of a single bump when temporally shuffled (crosses) or spatially shuffled (circles) external inputs are applied (see Section 4). (D) Improvements in the coding accuracy of bump attractors by reducing diffusion exponent and diffusion coefficient. The curves represent when both methods achieve the same improvements in coding accuracy at $t = 1$ s (solid) and $t = 10$ s (dashed) respectively. The square and the circle indicates the diffusion exponent $\beta = 0.75$ as obtained in our model.

4 The origin of subdiffusive dynamics from correlated inputs

Given that bump attractors with subdiffusive dynamics can significantly improve coding accuracy, it is important to understand the mechanism underlying this subdiffusive dynamics. We begin with investigating the temporal properties of the inputs received by the individual bumps shown in Fig. 1A. Since the influence of other bumps on a given bump depends on their distances and these bumps wander persistently in space, we introduce a co-moving frame for each bump to construct input sequences it receives. A position $\bar{r}$ in such a co-moving frame is given by $\bar{r} = r_{ij} - [r(t)]$, where $r_{ij} = (i, j)$ is a point in the original coordinates of the lattice and $r(t)$ is the bump position as calculated based on the population vector. The coordinates of the co-moving frame are rounded to the nearest integer as the inputs are only defined at discrete lattice points. The input to a position $\bar{r}$ and at a time t is therefore defined in a similar manner as Eq. 3:

$$I(\bar{r}, t) = \sum_{\bar{r}' \in \mathbb{Z}^2} W(\bar{r}, \bar{r}')\delta[u(\bar{r}', t) - 1], \tag{6}$$

where $W(\bar{r}, \bar{r}')$ is the synaptic coupling between two positions located at $\bar{r}$ and $\bar{r}'$, and $\delta(x)$ is the delta function. For convenience, we will drop the bar for the rest of this section and simply write $I(r, t)$ to mean inputs in the co-moving frame.

In order to reveal the effect of inputs from surrounding bumps, we consider separately the internal input and the external input to each bump: the internal input $I_{\text{int}}(r, t)$ to a location r refers to the total input from the spikes of the same bump, and the external

input $I_{\text{ext}}(r,t)$ refers to the total input from spikes belonging to the surrounding bumps. For the case with only one bump ($N = 1$), all inputs are purely internal. As shown in Fig. 2A, the time series of the internal input appears to be unstructured fluctuations similar to white noise. In contrast, the time series of the external input shows some self-similar temporal structures (Fig. 2B). To quantify the temporal property of these inputs, we calculate their autocorrelation functions (ACFs), as defined below. First, we calculate the unnormalized ACF,

$$C'(\tau; r) = \frac{1}{T}\sum_{t}[I(r,t) - \langle I\rangle][I(r,t+\tau) - \langle I\rangle], \tag{7}$$

where $\langle I\rangle = \frac{1}{T}\sum_t I(r,t)$ is the input mean over the full length T of the input time series. We then normalize the above expression to obtain the ACF:

$$C(\tau; r) = \langle I(r,t), I(r,t+\tau)\rangle = \frac{C'(\tau; r)}{C'(0; r)}, \tag{8}$$

where $C'(0; r)$ is the temporal variance of $I(r,t)$.

Figure 2C shows that the internal and the external inputs display very different temporal properties; the internal inputs I_{int} have an ACF that rapidly decays to zero, consistent with white noise whose ACF is a delta function. The ACF of the external input, however, exhibits a long tail that can be described using a power-law function, $I_{\text{ext}} \sim t^{-\alpha}$. The example here is for the input at $|r| = 5$, whose ACF has a power-law exponent of $\alpha \approx 0.76$. This power-law function of the ACF indicates that the external inputs to each bump have long-range temporal correlations. This is in contrast to previous theoretical studies of bump attractors, which only consider temporally uncorrelated

noise whose ACF is a delta function. The distinction between the internal and the external inputs are further confirmed by their power spectrum as shown in Fig. 2D. We also find that when there is only one bump in the network ($N = 1$), the internal inputs received by the bump have identical ACFs and power spectra as the internal inputs of the subdiffusive case ($N = 8$). In other words, bumps undergoing normal diffusion and subdiffusion both receive temporally uncorrelated internal inputs, but only the latter receive long-range temporally correlated external inputs, suggesting that such correlated inputs could be responsible for generating subdiffusive dynamics.

What about the spatial correlation property of the external input, aside from its long-range temporal correlations? Figure 2E shows the spatial profile of the external input at a fixed time instant, which is locally peaked at the origin and decreases smoothly outwards. We calculate spatial correlation as a function of displacement ρ, defined in the same way as Eq. 8, *i.e.*, $C(\rho; t) = \langle I(r, t), I(r + \rho, t)\rangle$, where $I(r, t)$ is given in Eq. 6 and the brackets denote averaging over space with $r \in [-\frac{L}{2}, \frac{L}{2}]^2$ and $L = 20$, the typical distance between two bumps when $N = 8$. Figure 2F shows the cross-section of the spatial correlation $C(\rho, t)$ of the external inputs along the x-axis, decreasing monotonically as distance increases. Unlike the long-range temporal correlation, the spatial correlation is short-ranged, reaching the half maximum at $x \approx 4$ and zero at $x \approx 10$.

To test if both the temporal and the spatial correlations are required for generating subdiffusion, we apply modified versions of the external inputs to a single bump with $N = 1$. In one case, a random shuffling is applied to the temporal sequences of the external inputs, resulting in a delta-shaped temporal ACF (solid orange line, Fig. 2C)

without changing the spatial correlations. In another case, we randomly shuffle the spatial locations of the external inputs, thereby destroying the spatial correlations (dashed line, Fig. 2F) while preserving the temporal correlations. In both cases, the bump remains normally diffusive (crosses and circles, Fig. 1C), indicating the combination of the temporal and the spatial correlations, as originated from multiple bumps and their interactions, is required for generating the subdiffusive dynamics of bump attractors.

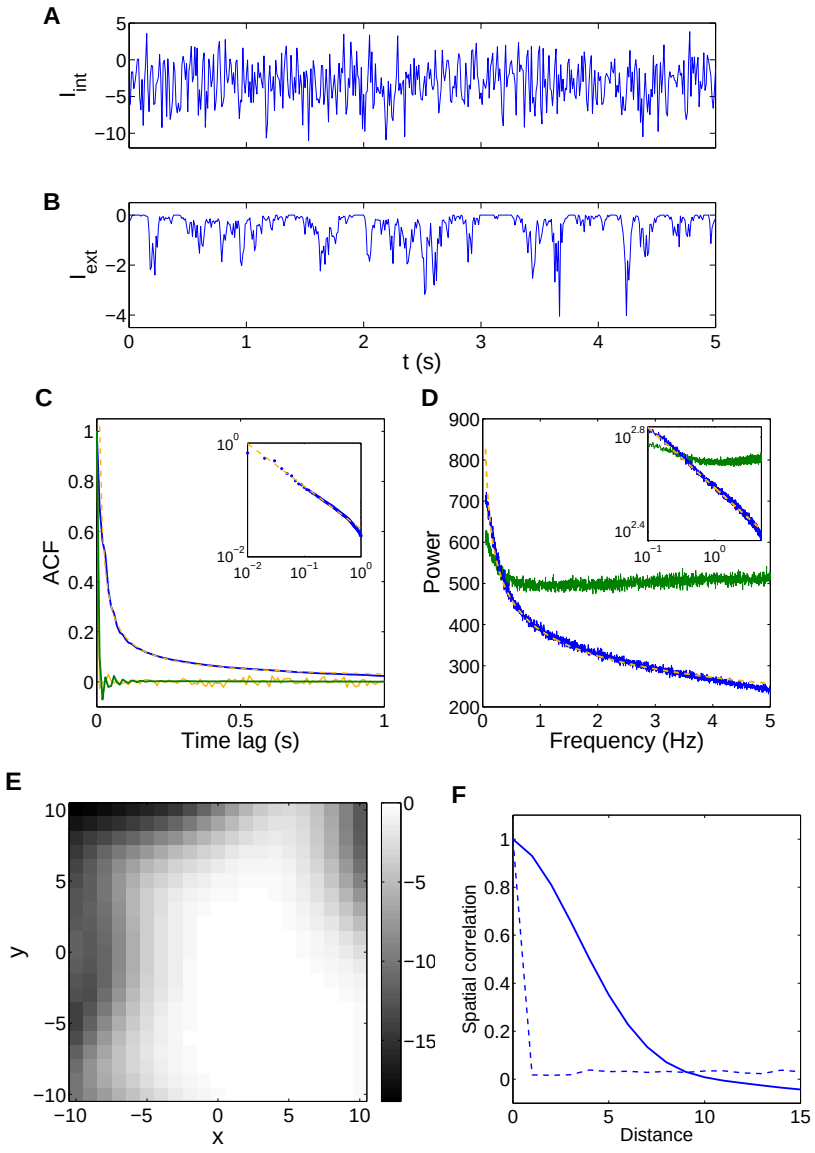

Figure 2: **Correlation properties of the inputs.** (A,B) Time series of internal and external inputs. The former shows completely random fluctuations similar to white noise, whereas the latter shows apparent self-similar temporal structures. (C) Autocorrelation functions of internal and external inputs. The ACF of the internal input decays rapidly towards zero (solid green line), resembling a delta function for white noise. In contrast, the external input has a long-tailed ACF (solid blue line), which follows a power-law, $C(\tau) \sim \tau^{-\alpha}$, with an exponent of $\alpha \approx 0.76$ (dashed orange line). The ACF of temporally shuffled external input also resembles a delta function (solid orange line). *Inset*: log-log plot of the ACFs. (D) Power spectra of internal and external inputs. The power spectrum of the internal input (solid green line) is flat across the frequency range, characteristic of white noise, whereas the spectrum of the external input (solid blue line) decays in the form of a power-law $f^{-\gamma}$, corresponding to noise with long-range temporal correlations. The dashed orange line is a power-law fit with exponent $\gamma = 0.25$. *Inset*: log-log plot of the power spectra. (E) Snapshot of the spatial profile of external inputs. (F) Spatial correlation of internal inputs resembles a delta function (dashed), whereas the external inputs decreases monotonically as distance increases (solid).

5 Subdiffusion in a conductance-based spiking neural circuit

Based on the simple spiking neural network model, we have shown that noisy inputs with long-range temporal correlations and spatial correlations underlie the subdiffusion of bump attractors. To further confirm this mechanism, we investigate the bump attractor dynamics in a conductance-based, spiking neural circuit model (see Appendix). In the conductance-based model, we use a transient stimulus to evoke a localized bump; as in our simple model, the bump persists over time after the stimulus is removed. We compare the diffusive dynamics of the bumps in two cases, one with uncorrelated noisy inputs and another with correlated inputs. Various numerical methods have been proposed for generating noise with long-range correlations (Hamzehpour and Sahimi, 2006; Makse et al., 1996; Garcia-Ojalvo and Sancho, 1999), but they could not reliably generate noise with the desired correlations. For our purpose, however, we mainly require a proof of principle for subdiffusive bump dynamics based on correlated inputs. Therefore we apply directly the external inputs from the three-state model (Fig. 2B) to a single bump in the conductance-based model as an external additive current (see Appendix). When the bump receives uncorrelated Gaussian white noise, the MSD of the bump grows linearly with time, corresponding to normal diffusion as found in previous studies of bump attractors. However, when we add external inputs with long-range temporal correlations and spatial correlations, the MSD scales sub-linearly in the form of a power-law function, $\sigma^2(t) \sim t^\beta$, with $\beta \approx 0.78 < 1$ (Fig. 3). Similar to the simple model, random shuffling of either the temporal sequences or the spatial positions of the

external inputs destroys the subdiffusion. These results therefore confirm the importance of correlated inputs for generating subdiffusive dynamics in bump attractors.

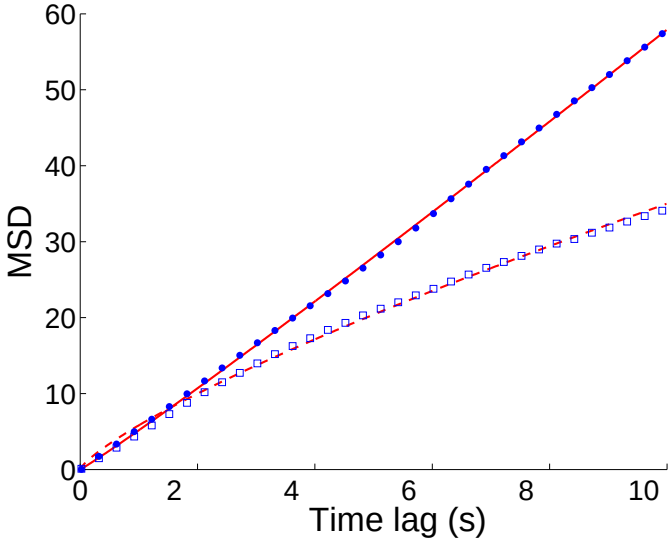

Figure 3: **Mean-squared displacement (MSD) of a bump attractor in a conductance-based spiking neural network**. White noise and spatio-temporally correlated noise show distinct effects on the diffusive motion of a bump. The former shows an MSD growing linearly with time (dots), consistent with normal diffusive dynamics, whereas the latter shows an MSD that increases sub-linearly over time, indicating subdiffusive dynamics (squares). The solid and the dashed lines show the corresponding fittings to power-law function of the form $\sigma^2(t) \sim t^\beta$, with the exponents found to be $\beta \approx 1.05$ and $\beta \approx 0.78$, respectively.

6 The Langevin theory of subdiffusive dynamics

Having shown that noisy inputs with long-range temporal correlations and spatial correlations are essential for subdiffusive bump dynamics in spiking neural circuits, we now analyze a simplified model in order to gain further theoretical insights into the origin of subdiffusive dynamics of bump attractors. We consider each bump as a point particle, a simplification that has been used previously (Kilpatrick et al., 2013; Renart et al., 2003), and regard the collective influence of the inputs as a random force act-

ing on the particle; such a collective force is called a Langevin force in the Langevin description of particle motion. Due to spatial isotropy, it is sufficient to consider one spatial component of the force. Suppose the external input fluctuations are small, the Langevin force $F(t)$ exerting on a bump can be approximated by their linear combination, $F(t) = \sum_r f(r)\delta I(r,t)$, where $f(r)$ is some undetermined coefficient depending on spatial location. The ACF of the force is a bilinear form of the input cross-correlations,

$$C(\tau) = \langle F(t), F(t+\tau)\rangle = \sum_{r,r'} f(r)f(r')\langle I(r,t), I(r',t+\tau)\rangle. \tag{9}$$

As verified numerically, the ACFs ($r = r'$) are the dominant terms on the right-hand side of Eq. 9, whereas the cross-correlation terms ($r \neq r'$) follow power-law functions with smaller amplitudes. Since the external inputs have long-range temporal correlations, the Langevin force should also inherit this temporal structure. Moreover, the spatial correlation ensures that the sum of the cross-correlations (including the ACFs) is enhanced, instead of being canceled out.

Based on this argument, the bump can be abstracted as a particle driven by a long-range temporally correlated force, and its motion can thus be formulated with a generalized Langevin equation (GLE), (Porrà et al., 1996; Wang and Tokuyama, 1999; Viñales and Despósito, 2007):

$$\ddot{X}(t) + \int_0^t \kappa(t-t')\dot{X}(t')dt' = F(t), \tag{10}$$

where $X(t)$ is the bump position and $\kappa(t)$ is a memory kernel. The quantity $F(t)$ is a

Langevin force with zero mean and an ACF,

$$C(\tau) = \langle F(t), F(t+\tau) \rangle \sim \tau^{-\alpha}, \tag{11}$$

where τ is the time lag and $0 < \alpha < 1$ is the same exponent as for the noisy external inputs received by the bump, *i.e.*, $\alpha \approx 0.76$ as in Eq. 8. Under the condition of thermal equilibrium, the GLE satisfies the fluctuation-dissipation theorem that relates the force ACF to the memory kernel (Henery, 1971):

$$C(\tau) = k_B T \kappa(\tau), \tag{12}$$

where k_B is Boltzmann's constant and T is the absolute temperature.

A standard way to formally solve the GLE is via the method of Laplace transform (Viñales and Despósito, 2007). The Laplace transform of a function $x(t)$ is defined as:

$$\mathcal{L}x(t) = \hat{x}(s) = \int_0^\infty x(t) e^{-st} dt. \tag{13}$$

By applying Laplace transform to the GLE (Eq. 10), along with the relation $\mathcal{L}\dot{X} = s\hat{X} - X(0)$ and the convolution theorem $\mathcal{L}(\kappa * x) = \hat{\kappa}\hat{x}$, we get,

$$\hat{X}(t) = \frac{1}{s} X(0) + \hat{G}(s)\dot{X}(0) + \hat{G}(s)\hat{F}(s), \tag{14}$$

where,

$$\hat{G}(s) = \frac{1}{s^2 + \hat{\kappa}(s)s}. \tag{15}$$

Laplace inversion of Eq. 14 then gives the formal solution of the GLE of Eq. 10:

$$X(t) = \langle X(t) \rangle + \int_0^t G(t - t')F(t')dt', \tag{16}$$

where $\langle X(t) \rangle = X(0) + \dot{X}(0)G(t)$. The variance in position can be calculated directly from Eq. 16 (Porrà et al., 1996):

$$\sigma_{xx}^2 = \langle [X(t) - \langle X(t) \rangle]^2 \rangle \tag{17}$$

$$= 2 \int_0^t G(t')dt' \int_0^{t'} G(t'')C(t' - t'')dt''. \tag{18}$$

Note that the variance has the same expression as the MSD defined in Eq. 4. Based on the fluctuation-dissipation theorem (Eq. 12), the variance is simplified into,

$$\sigma_{xx}^2 = k_B T \left[2 \int_0^t G(t')dt' - G(t)^2 \right] \tag{19}$$

In the case $F(t)$ is a Gaussian white noise whose ACF is a delta function, $C(\tau) = \delta(\tau)$, the GLE is reduced to an ordinary Langevin equation that describes normal diffusion, with a variance equal to $\sigma_{xx}^2 = 2k_B T t$. However, anomalous diffusion arises when the random force has a long-range temporal correlation, *i.e.*, $C(\tau) \sim \tau^{-\alpha}$. From Eq. 12 we also obtain the memory kernel $\kappa(\tau) \sim \tau^{-\alpha}$. We can then derive the asymptotic behavior of the MSD as $t \to \infty$ by considering the corresponding asymptote in the Laplace space as $s \to 0$ (Wong, 1975). The memory kernel in the Laplace space is:

$$\hat{\kappa} \sim s^{\alpha - 1}. \tag{20}$$

Therefore $\hat{G}(s)$ of Eq. 15 becomes:

$$\hat{G}(s) \sim \frac{1}{s^2 + s^\alpha}. \tag{21}$$

Since $0 < \alpha < 1$, the square term vanishes as $s \to 0$. Therefore,

$$\hat{G}(s) \sim s^{-\alpha}. \tag{22}$$

Laplace inversion of the above expression gives:

$$G(t) \sim t^{\alpha-1}. \tag{23}$$

Substituting $G(t)$ to Eq. 19, we obtain the final asymptotic expression for the variance of the bump position (in one dimension):

$$\sigma_{xx}^2 \sim k_B T \left[2 \int_0^t t'^{\alpha-1} dt' - (t^{\alpha-1})^2 \right] \sim t^\alpha. \tag{24}$$

The MSD of the bump position in two spatial dimensions is simply $\sigma^2 = \sigma_{xx}^2 + \sigma_{yy}^2 \sim t^\alpha$. Matching this expression of MSD derived from the GLE to the definition of MSD in Eq. 4, we find that the power-law exponent of the ACF, α, is equal to the diffusion exponent, β. This implies the bump dynamics is subdiffusive as $\alpha = \beta < 1$, qualitatively in agreement the subdiffusive bump dynamics obtained from the simulation. Quantitatively, this is also consistent with the simulation result as the exponent $\alpha \approx 0.76$ of the ACF of the inputs (Fig. 2C) is approximately equal to the diffusion exponent

with $\beta \approx 0.75$ (Fig. 1C). These results therefore indicate that the GLE with long-range temporally correlated force is able to account for the subdiffusive bump dynamics.

To further show how the bump dynamics can be captured by the GLE, we consider the probability density function (PDF) of bump position. The Fokker-Planck equation (FPE) associated with the GLE with two spatial dimensions $r = (x, y) \in \mathbb{R}^2$ is (Wang and Tokuyama, 1999):

$$\frac{\partial P(r,t)}{\partial t} = -g(t)\dot{r}(0) \cdot \nabla P(r,t) + k_B T G(t)[1 - g(t)]\nabla^2 P(r,t), \tag{25}$$

where $P(r,t)$ is the PDF for the bump position, $G(t)$ is given by Eq. 23 and $g(t)$ is the Laplace inversion of $\hat{g}(s) = \frac{1}{s+\hat{\kappa}(s)}$. Using the similar argument as before, the asymptotic form of $g(t)$ as $t \to \infty$ is found to be $g(t) \sim (\alpha - 1)t^{\alpha-2}$. For a pulse initial condition $P(r,0) = \delta(r)$, the solution for Eq. 25 is:

$$P(r,t) = \frac{1}{4\pi D t^{\alpha}} \exp\left(-\frac{|r - \langle r \rangle|^2}{4Dt^{\alpha}}\right), \tag{26}$$

where D is a generalized diffusion coefficient proportional to $k_B T/[\alpha\Gamma(|1-\alpha|)\Gamma(2-\alpha)]$, α is the exponent of the ACF of the force (as in Eq. 11), and $\langle r \rangle = 0$ is the mean position. This PDF of Eq. 26 gives the probability density of the bump position r at time t.

We compare this theoretical PDF to that obtained from numerical simulations of the three-state neural circuit model. We track the trajectory $r(t)$ of a bump over $m = 1000$ trials, and calculate the histogram $H(|r|,t)$ of the bump's radial distance $|r(t)|$ with a bin size of one. The radial profile of the PDF is therefore calculated as $P(|r|,t) =$

$H(|r|,t)/(m2\pi|r|)$. The time evolution of the peak of the PDF, $P(0,t)$, is shown in Fig. 4A, which according to Eq. 26 should decay over time in the form of a power law $\frac{1}{4\pi D t^{\alpha}}$. As shown by the solid line, the theoretical curve according to this expression fits well to the numerical data, with parameters $D \approx 9.66$ s$^{-\alpha}$ and $\alpha \approx 0.74$. Note this exponent α is consistent with the exponent used to describe the ACF of the input (Fig. 2C) and the exponent of the bump MSD (Fig. 1C) from the simulation. As in Eq. 26, the corresponding radial profiles of the PDF follow Gaussian distributions; Fig. 4B shows two examples at $t = 1$ s and $t = 2.5$ s respectively, both with agreement between the numerical (dotted) and the theoretical (solid) curves. These results further demonstrate that GLE with long-range temporally correlated force is able to describe the subdiffusive dynamics of bump attractors.

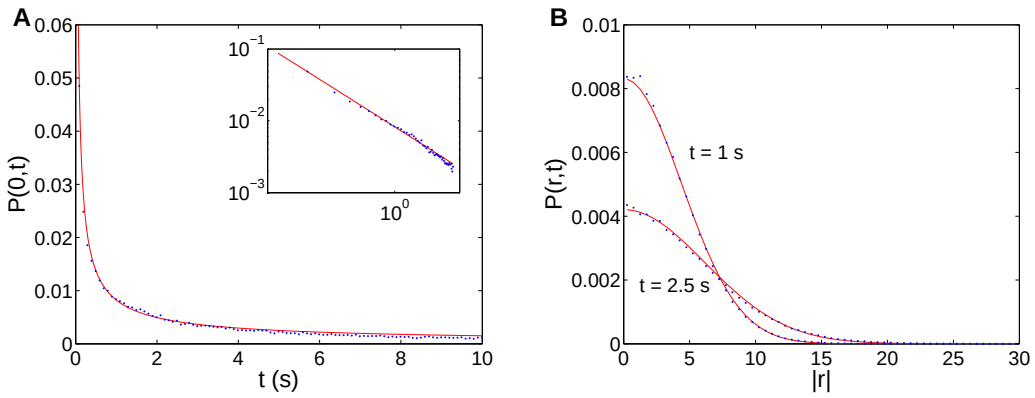

Figure 4: **Probability density function of bump position.** (A) The amplitude of the PDF $P(0,t)$ decays as a power-law function of time $\frac{1}{4\pi D t^{\alpha}}$. The dotted plot shows the numerically calculated PDF at $|r| = 0$; the solid line is the fitted curve with $D \approx 9.66$ s$^{-\alpha}$ and $\alpha \approx 0.74$. The exponent α is consistent with those of the input ACF and the MSD. *Inset*: log-log plot of the PDF. (B) Radial profile of the PDF for the bump position follows a Gaussian distribution. The dotted lines are the simulated data at $t = 1$ s and $t = 2.5$ s respectively; the solid lines are the theoretical prediction using the same values of D and α as in (A).

7 Power-law decay of firing rate and working memory performance

As introduced above, one of the characteristic features of the subdiffusive dynamics of bump attractors is that the variance of coding errors increases sub-linearly with time. In this section, we present two other features of bump attractors with subdiffusive dynamics; one is the power-law decay of firing rate of persistent activity, and another is the power-law decay of working memory performance. These two properties are derived from the PDF in Eq. 26, and are also calculated numerically from the three-state neural network model.

First, we derive the firing rate $f(r,t)$ from the PDF (Eq. 26) through a spatial convolution: $f(r,t) = \int_{\mathbb{R}^2} \phi(r')P(r-r',t)dr'$, where $\phi(r)$ is the firing rate profile of a bump. At $t = 0$ the PDF is a delta function $P(r,0) = \delta(r)$, and the firing rate is simply $f(r,0) = \phi(r)$. For $t > 1$ s such that the width of the PDF $P(r,t)$ is larger than the diameter of the bump $\phi(r)$, the firing rate converges to:

$$f(r,t) \sim \bar{\phi} P(r,t), \tag{27}$$

where $\bar{\phi} = \int_{\mathbb{R}^2} \phi(r)dr$. The above equation states that the firing rate is proportional to the PDF (Eq. 26) for large t, therefore it decays asymptotically to a power-law function, *i.e.*, $f(r,t) \sim t^{-\alpha}$ with $0 < \alpha < 1$, where α is the same as the exponent of the force ACF in Eq. 11. We compare this theoretical result with the trial-averaged firing rates obtained from the simulation using the formula: $f_{ij}(t) = \frac{1}{m}\sum_{k=1}^{m} \delta[u_{ij}^k(t) - 1]$, where

$f_{ij}(t)$ is the firing rate of an ij-th neuron, $m = 1000$ is the total number of trials, and k is the trial index. The firing rate calculated here excludes the spike counts due to surrounding bumps. Similarly, the bump profile $\phi(r)$ can be calculated numerically as $\phi(r) = \frac{1}{m}\sum_{k=1}^{m}\delta[u^k(\bar{r},t)-1]$, where $\bar{r}$ denotes the co-moving frame. Figure 5A shows the simulation result of the firing rate (solid line) for two neurons located at $|r| = 0$ and $|r| = 5$ respectively, in agreement with the corresponding theoretical asymptotes (dashed line) as predicted by Eq. 27. This form of long-tailed firing rate decay with a power-law exponent $\alpha < 1$ is significantly slower than the decay based on normal diffusion with $\alpha = 1$. Bump attractors with subdiffusive dynamics therefore provides persistent activity with significantly longer duration than those with normal diffusion.

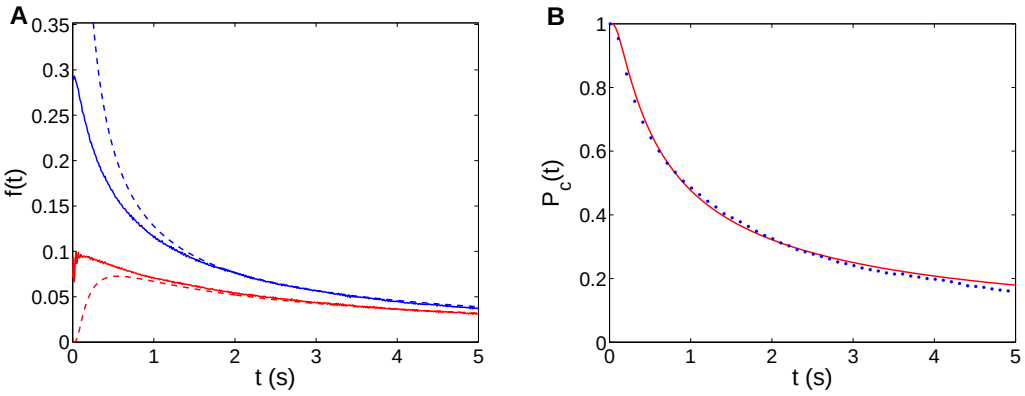

Figure 5: **Power-law decay of firing rate and performance.** (A) The firing rates of neurons decay in the form of power-law as predicted by the PDF. The solid lines show the firing rates of neurons at fixed locations with $|r| = 0$ (blue) and $|r| = 5$ (red) from the initial stimulus. The dashed lines show the theoretical asymptotes for the firing rates as predicted by the PDF, with excellent convergence for $t > 2$ s. (B) Performance of working memory is measured using correct rate of recall with a simple threshold criterion. The simulation results (dotted line) is in agreement with the theoretical asymptote (solid line) as predicted by the PDF.

The second prediction is the power-law decay of working memory performance over time. A common way for quantifying working memory performance in behavioral studies, especially those involving verbal recall of words, is to measure the rate of

correct recall after a delay period t (Wixted and Ebbesen, 1991; Donkin and Nosofsky, 2012). In the framework of using bump attractor to store working memory, the initial bump position $X(0)$ encodes the memorized item and the bump location $X(t)$ after a delay period t corresponds to the recalled item. Therefore, a recall is considered as correct if the final bump position is within a radius of the initial bump position, i.e., $|r(t) - r(0)| < r_0$. Here we set $r_0 = 5$, roughly the same as the diameter of the bump, but note that choices with different threshold values would not significantly affect our results. The probability for a correct recall can therefore be calculated from the PDF $P(r, t)$ of Eq. 26 as,

$$P_c(t) = \int_{|r|<r_0} P(r,t)dr = 1 - \exp\left(-\frac{r_0^2}{4Dt^\alpha}\right). \tag{28}$$

For $t \gg (\frac{r_0^2}{4D})^{1/\alpha} \approx 0.6$ s, we can expand the exponential function to the first order, and find that the correct recall rate converges rapidly to an asymptote $P_c(t) \sim \frac{r_0^2}{4Dt^\alpha}$. In other words, the rate of correct recall approximately follows a power-law decay for large t with $\alpha < 1$, where α is equal to the exponent of the force ACF. The numerical result of the correct rate of recall as a function of time is shown in Fig. 5B (dotted line), which is in agreement to the theoretical prediction (solid line).

8 Existing experimental evidence for subdiffusive bump dynamics

One of the consequences of subdiffusive dynamics of bump attractors is that the variance of coding errors increases sub-linearly with time. To test this, we re-examine existing psychophysical data about spatial working memory. White et al. (1994) studied the saccadic eye movements of monkeys during a delayed spatial working memory task, in which the monkeys were trained to track the location of an initial stimulus after a delay period. The monkeys' eye movements showed random error from the target location that on average increased with time. The qualitative increase of the error with time has been considered as evidence for normal diffusion of bump attractors (Compte et al., 2000; Kilpatrick et al., 2013). We re-analyze their data of the error (White et al., 1994) and demonstrate quantitatively that the variance grows sub-linearly with time, which can be accounted for by anomalous subdiffusion, instead of normal diffusion. There are two points we need to highlight for interpreting the data: firstly, the original paper presents the standard deviation $\sigma(t)$ of the targeting error, instead of the variance $\sigma^2(t)$; secondly, the total variance $\sigma^2(t) = \sigma_1^2(t) + \sigma_0^2$ has two components, $\sigma_1^2(t)$ the variance of the genuine coding error that grows with time, and $\sigma_0^2 = \sigma^2(0)$ a constant intrinsic variance of the eye movement. It is the former $\sigma_1^2(t)$ we are interested in, as shown in the scatter plot of Fig. 6A [redrawn from Fig.5B of White et al. (1994)]. There are two series of data, both clearly showing that the variance increases sub-linearly over time. The data also show a clear trend of a straight line in the corresponding log-log plot, as in Fig. 6B and Fig. 6C, suggesting the variance increases with time as a power-

law function, *i.e.*, $\sigma_1^2(t) = At^\beta$. The exponents are found to be $\beta = 0.72$, with a 95% confidence interval of [0.52,0.92], and $\beta = 0.37$, with a 95% confidence interval of [0.13,0.60], respectively; both exponents are significantly less than one, indicating that the eye movement during delayed spatial working memory tasks involves subdiffusive dynamics.

We further show that the power-law model is a better fit than a linear model of the form $\sigma_1^2(t) = At$. We first compare the residual standard deviations of each fit: the power-law fit yields 0.75 and 0.57 whereas the linear fit yields 1.07 and 1.27, for the two data series respectively. The power-law function produces less error in the fit than the linear function for both data series. Since the power-law model has one additional parameter than the linear model, it is necessary to conduct a suitable model comparison that takes this into account. For this purpose, we calculate the Bayes factor (Kass and Raftery, 1995), $B = \frac{P(D|M_1)}{P(D|M_0)}$, where $P(D|M)$ is the marginal likelihood for data D under model M (see Appendix for details). The Bayes factors are found to be $B = 40$ and $B = 4.6 \times 10^5$ for the two data series respectively. In both cases, the Bayes factors are significantly greater than one, suggesting that the power-law function is a better model than the linear function (Kass and Raftery, 1995). Our re-examination of these existing data therefore shows that the variance of the recalled position of the stimulus cue increases sub-linearly with time, largely consistent with subdiffusive dynamics of a bump attractor.

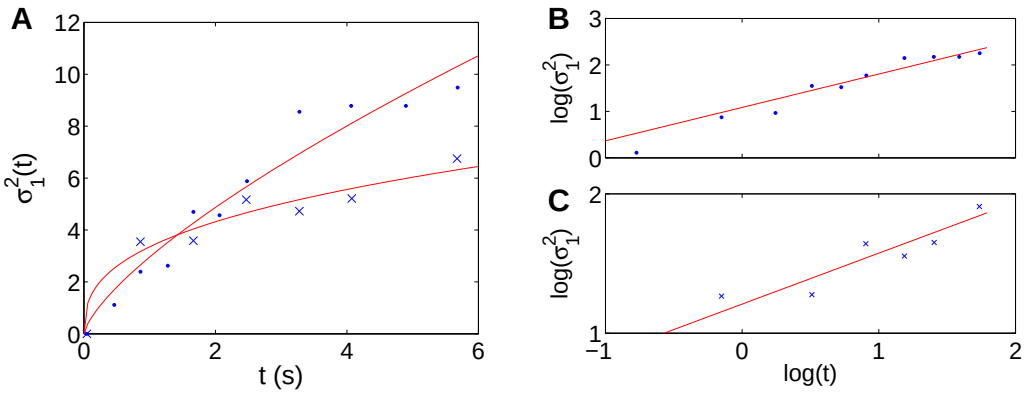

Figure 6: **Targeting error of saccadic eye movement during delayed spatial working memory tasks.** (A) The variances of the eye position (degree) of two subjects (dots and crosses) show sub-linear increase over time during the delay period. The data are fitted to power-laws of the form At^{β} with exponents found to be $\beta = 0.72$, 95% CI [0.52,0.92], and $\beta = 0.37$, 95% CI [0.13,0.60], respectively. The data points are replotted from Fig. 5B of White et al. (1994). Note that the original paper reported the standard deviation instead of the variance. (B,C) Each of the data series of variance as in (A) plotted on a log-log scale.

9 Discussion and conclusion

In this study, we have investigated the anomalous subdiffusive dynamics of bump attractors in a simple spiking neural network model and we have found such dynamics are caused by fluctuating inputs with long-range temporal correlations and spatial correlations generated by multiple, interacting bumps. The importance of correlated inputs for anomalous diffusion of bump attractor is further verified in a biologically more plausible, conductance-based spiking network. To demonstrate that long-range temporally correlated inputs are essential for the subdiffusive dynamics, we have simplified a bump as a particle and then applied a generalized Langevin equation to investigate its motion. Based on our analysis of the GLE, we have obtained several analytical results, including the diffusion exponent β that is determined by the exponent α of the power-law function

describing the autocorrelation function of the input, the probability density function of bump position, the long-tailed firing rate decay for persistent activity, and the decay of correct rate of recall. These analytical results are in good agreement with those obtained by simulating the spiking neural circuit models, suggesting that the GLE is a suitable theoretical approach for describing the dynamics of bump attractors with correlated inputs. The subdiffusive dynamics provide an efficient way for improving the coding accuracy of bump attractors by reducing the diffusion exponent, instead of the diffusion coefficient of normal diffusion (Kilpatrick et al., 2013; Polk et al., 2012). Our study therefore suggests that neural fluctuations with long-range temporal correlations and spatial correlations, which have been widely observed in the brain (Teich et al., 1990; Destexhe et al., 1999; Linkenkaer-Hansen et al., 2001; Leopold et al., 2003; Werner, 2010), can have significant functional benefits for neural information processing.

However, we would like to note that the subdiffusive dynamics of bump attractors in our model are limited to the case with multiple, interacting bumps and the particular combination of spatio-temporal correlations originated from them. Despite this limitation, bump attractors with subdiffusive dynamics show several characteristic features that are fundamentally different from those with normal diffusion, which can be verified experimentally. Firstly, the subdiffusion of a bump attractor results in a sub-linear increase in the variance of bump positions with time, instead of linear increase as in normal diffusion. Since drifts of bump positions is predictive of memory-dependent inaccuracies in the behavioral report of the remembered locations (Compte et al., 2000), the sub-linear increase in the variance of bump position means that there should be a corresponding sub-linear increase in the variance of recall error. Secondly, the trial-averaged

firing rate of persistent activity generated by bump attractors with sub-diffusion follows a long-tailed decay in the form of a power-law function. With sufficient experimental data in the future, this prediction may be tested by fitting the firing rate decay to a power-law function to show that it is statistically better than fitting to exponential functions as in (Miri et al., 2011). Lastly, working memory performance as measured by the rate of correct recall also decays as a power-law function with an exponent smaller than one. This is supported by existing behavioral studies of working memory performance, in which power-law decay has been identified in a diverse range of measures on forgetting, including the proportion of correct recall of word lists, facial recognition, and savings in relearning lists of nonsense syllables (Donkin and Nosofsky, 2012; Wixted and Ebbesen, 1991).

The temporal increases in the spread of recall error in psychophysical studies of working memory have been considered qualitatively as an indication for the underlying diffusive dynamics of bump states. Based on our re-analysis of the psychophysical data from White et al. (1994), we have found that the variance of recall error increases sub-linearly with time, and can be fitted with a power law function with an exponent smaller than one, consistent with the prediction from the subdiffusion of bump attractors; however, in our model there are multiple bumps, a scenario not fully compatible with working memory experiments presented by White et al. (1994), which only involve a single memorized item. Nonetheless, existing theoretical studies simply overlook the quantitative details when citing these data, without acknowledging that normal diffusion is inadequate for explaining them. Compte et al. (2000) acknowledge that the variance of recall errors reported by White et al. (1994) cannot be fitted to a straight

line over the full duration of the data, but unable to provide an explanation. Here not only we show that the variance grows sub-linearly in the form of a power-law, but also be able to explain it through the subdiffusion of bump states. It is worth noting that different quantifications for errors of behavioral recall have been used by experimentalists, including the standard deviation (White et al., 1994) and the inter-quartile range of 'gain' (Ploner et al., 1998). We argue that in order to show the differences between errors of normal diffusion-based coding and subdiffusion-based coding, it is the best for future studies to use variance for quantifying behavioral inaccuracy; this is because normal diffusion and subdiffusion can be distinguished by their respective linear and sub-linear increases in the variance. Recently, Wimmer et al. (2014) have found specific physiological evidence for diffusive bump attractor dynamics in the pre-frontal cortex in explaining behavioral inaccuracies. Similar research maybe conducted in the future to show that the bump dynamics is subdiffusive rather than normal diffusive.

References

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27(2):77–87.

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848.

Benayoun, M., Cowan, J. D., van Drongelen, W., and Wallace, E. (2010). Avalanches in a stochastic model of spiking neurons. *PLoS Computational Biology*, 6(7):e1000846.

Bhattacharya, J., Edwards, J., Mamelak, A. N., and Schuman, E. M. (2005). Long-range temporal correlations in the spontaneous spiking of neurons in the hippocampal-amygdala complex of humans. *Neuroscience*, 131(2):547 – 555.

Buice, M. A. and Cowan, J. D. (2007). Field-theoretic approach to fluctuation effects in neural networks. *Physical Review E*, 75:051919.

Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10(9):910–923.

Destexhe, A., Contreras, D., and Steriade, M. (1999). Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. *Journal of Neuroscience*, 19(11):4595–4608.

Donkin, C. and Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short- and long-term recognition. *Psychological Science*, 23(6):625–634.

Ecker, A. S., Berens, P., Cotton, R. J., Subramaniyan, M., Denfield, G. H., Cadwell, C. R., Smirnakis, S. M., Bethge, M., and Tolias, A. S. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82(1):235–248.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303.

Freeman, W. J., Holmes, M. D., Burke, B. C., and Vanhatalo, S. (2003). Spatial spectra of scalp EEG and EMG from awake humans. *Clinical Neurophysiology*, 114(6):1053–1068.

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61(2):331–349.

Garcia-Ojalvo, J. and Sancho, J. M. (1999). *Noise in Spatially Extended Systems*. Springer.

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419.

Gong, P., Nikolaev, A. R., and van Leeuwen, C. (2003). Scale-invariant fluctuations of the dynamical synchronization in human brain electrical activity. *Neuroscience Letters*, 336(1):33–36.

Gong, P. and Robinson, P. A. (2012). Dynamic pattern formation and collisions in networks of excitable elements. *Physical Review E*, 85:055101(R).

Hamzehpour, H. and Sahimi, M. (2006). Generation of long-range correlations in large systems as an optimization problem. *Physical Review E*, 73:056121.

Henery, R. J. (1971). The generalized langevin equation and the fluctuation-dissipation theorems. *Journal of Physics A: General Physics*, 4:685–694.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kilpatrick, Z. P., Ermentrout, B., and Doiron, B. (2013). Optimizing working memory with heterogeneity of recurrent cortical excitation. *Journal of Neuroscience*, 33(48):18999–19011.

Leopold, D. A., Murayama, Y., and Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cerebral Cortex*, 13(4):422–433.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, 21(4):1370–1377.

Makse, H. A., Havlin, S., Schwartz, M., and Stanley, H. E. (1996). Method for generating long-range correlations for large systems. *Physical Review E*, 53:5445–5449.

McLaughlin, D., Shapley, R., Shelley, M., and Wielaard, D. J. (2000). A neuronal network model of macaque primary visual cortex (V1): orientation selectivity and dynamics in the input layer 4Cα. *Proceedings of the National Academy of Sciences*, 97(14):8087–8092.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7:663–678.

Metzler, R. and Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77.

Miller, P. (2006). Analysis of spike statistics in neuronal systems with continuous attractors or multiple, discrete attractor states. *Neural Computation*, 18(6):1268–1317.

Miri, A., Daie, K., Arrenberg, A. B., Baier, H., Aksay, E., and Tank, D. W. (2011). Spatial gradients and multidimensional dynamics in a neural integrator circuit. *Nature Neuroscience*, 14(9):1150–1159.

Ploner, C. J., Gaymard, B., Rivaud, S., Agid, Y., and Pierrot-Deseilligny, C. (1998). Temporal limits of spatial working memory in humans. *European Journal of Neuroscience*, 10(2):794–797.

Polk, A., Litwin-Kumar, A., and Doiron, B. (2012). Correlated neural variability in persistent state networks. *Proceedings of the National Academy of Sciences*, 109(16):6295–6300.

Porrà, J. M., Wang, K.-G., and Masoliver, J. (1996). Generalized langevin equations: Anomalous diffusion and probability distributions. *Physical Review E*, 53:5872–5881.

Pouget, A., Zhang, K., Deneve, S., and Latham, P. E. (1998). Statistically efficient estimation using population coding. *Neural Computation*, 10(2):373–401.

Renart, A., Song, P., and Wang, X.-J. (2003). Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron*, 38(3):473–485.

Roxin, A., Hakim, V., and Brunel, N. (2008). The statistics of repeating patterns of

cortical activity can be reproduced by a model network of stochastic binary neurons. *Journal of Neuroscience*, 28(42):10734–10745.

Seung, H. S., Lee, D. D., Reis, B. Y., and Tank, D. W. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron*, 26(1):259–271.

Smith, M. A. and Sommer, M. A. (2013). Spatial and temporal scales of neuronal correlation in visual area V4. *Journal of Neuroscience*, 33(12):5422–5432.

Sokolov, I. M. (2012). Models of anomalous diffusion in crowded environments. *Soft Matter*, 8(35):9043–9052.

Takeda, K. and Funahashi, S. (2004). Population vector analysis of primate prefrontal activity during spatial working memory. *Cerebral Cortex*, 14(12):1328–1339.

Teich, M. C., Johnson, D. H., Kumar, A. R., and Turcott, R. G. (1990). Rate fluctuations and fractional power-law noise recorded from cells in the lower auditory pathway of the cat. *Hearing Research*, 46(1):41–52.

Viñales, A. D. and Despósito, M. A. (2007). Anomalous diffusion induced by a mittag-leffler correlated noise. *Physical Review E*, 75:042102.

Wallace, E., Benayoun, M., van Drongelen, W., and Cowan, J. D. (2011). Emergent oscillations in networks of stochastic spiking neurons. *PLoS ONE*, 6(5):e14804.

Wang, K. G. and Tokuyama, M. (1999). Nonequilibrium statistical description of anomalous diffusion. *Physica A: Statistical Mechanics and its Applications*, 265(3):341 – 351.

Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, 24(8):455–463.

Werner, G. (2010). Fractals in the nervous system: conceptual implications for theoretical neuroscience. *Frontiers in Physiology*, 1:15.

White, J. M., Sparks, D. L., and Stanford, T. R. (1994). Saccades to remembered target locations: an analysis of systematic and variable errors. *Vision Research*, 34(1):79–92.

Wimmer, K., Nykamp, D. Q., Constantinidis, C., and Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*, 17(3):431–439.

Wixted, J. T. and Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, 2(6):409–415.

Wong, R. (1975). On laplace transforms near the origin. *Mathematics of Computation*, 29(130):573–576.

Wu, S., Hamaguchi, K., and Amari, S. (2008). Dynamics and computation of continuous attractors. *Neural Computation*, 20(4):994?025.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *Journal of Neuroscience*, 16(6):2112–2126.

Appendix

A conductance based model

A two-dimensional conductance-based spiking neural circuit with external noisy input is used to demonstrate the generality of the subdiffusive dynamics. The model uses integrate-and-fire neurons whose dynamics is described as follows,

$$\frac{dV_{ij}}{dt} = -g_L(V_{ij} - V_R) - g_{ij}^E(s)(V_{ij} - V_E) - g_{ij}^I(s)(V_{ij} - V_I) + \varepsilon_1 \xi_{ij}(t) + \varepsilon_2 I_{\text{ext}},$$

where $V_{ij}(t)$ is the membrane potential of an ij-th neuron in a two-dimensional neural circuit, and g_{ij}^E and g_{ij}^I are the excitatory and the inhibitory conductances, respectively. The unit of the voltage is normalized such that the firing threshold is $V_{\text{th}} = 1$ and the reset potential is $V_{\text{reset}} = 0$. A refractory period of $\tau_{\text{ref}} = 2$ ms follows immediately after the reset, that is, set $V_{ij}(t) = 0$ for $t \in (T_{ij}^l, T_{ij}^l + \tau_{\text{ref}}]$, where T_{ij}^l is the time of an l-th spike of a neuron satisfying $V_{ij}(T_{ji}^l) = 1$. The leak conductance is $g_L = 0.05\ \text{ms}^{-1}$, and the resting, the excitatory, and the inhibitory reversal potentials are $V_R = 0$, $V_E = 14/3$, and $V_I = -2/3$, respectively. For conversion to physiological units, see McLaughlin et al. (2000). The conductances are given by:

$$g_{ij}^\lambda(t) = W_\lambda \sum_{kl} W_{ij,kl}^\lambda s_{kl}(t),$$

where $\lambda \in \{E, I\}$, $s_{ij}(t)$ is the synaptic activation of the ij-th neuron, and $W_{ij,kl}^\lambda$ is the excitatory or the inhibitory coupling between the ij-th and the kl-th neurons. The coupling strengths are $W_E = 13.85$ and $W_I = 8.46$ respectively. Similar to the three-

state model, we consider distance-dependent coupling with local excitation and lateral inhibition:

$$W^{\lambda}_{ij,kl} = \frac{1}{\pi\sigma^2_{\lambda}} \exp\left(-\frac{|r_{ij} - r_{kl}|^2}{\sigma^2_{\lambda}}\right),$$

where $r_{ij} = (i, j)$ is a position on the network, and $\sigma_E = 1.94$ and $\sigma_I = 7.06$ are the ranges of the excitatory and the inhibitory coupling respectively. The synaptic activation $s_{ij}(t)$ follows,

$$\tau\frac{ds_{ij}}{dt} = -s_{ij} + \sum_m \delta(t - T^m_{ij}),$$

where $\tau = 5$ ms is the time scale of the synaptic delay, T^m_{ij} is the time of an m-th spike of a neuron satisfying $V_{ij}(T^m_{ij}) = 1$ and $\delta(t)$ is the Dirac delta function. Lastly, $\xi_{ij}(t)$ is a standard Gaussian white noise with zero-mean and $I_{\text{ext}}(\bar{r}, t)$ is the spatio-temporally correlated external input extracted from the three-state model with $N = 8$, where $\bar{r}$ denotes co-moving frames. The strengths of the noises are scaled by $\varepsilon_1 = 0.003$ and $\varepsilon_2 = 0.001$, respectively.

Calculating Bayes factor

The Bayes factor is defined as, $B = \frac{P(D|M_1)}{P(D|M_0)}$, where $P(D|M)$ denotes the marginal likelihood for data D under model M. The marginal likelihood is calculated as, $P(D|M) = \int P(\theta|M)P(D|\theta, M)d\theta$, where $P(\theta|M)$ is the prior distribution and $P(D|\theta, M)$ is the likelihood function. We assume that an observed data point (t_i, d_i) is given by $d_i = f(t_i|M) + e_i$, where $f(t|M_0) = At$ and $f(t|M_1) = At^{\beta}$, and the error term e_i follows a Gaussian distribution with zero mean and a standard deviation of λ. We use the residual standard deviation of the curve fitting as estimations for the standard devi-

ation λ, which is sufficient for the purpose of model comparison. A suitable choice of the prior distribution is Jeffrey prior, defined as, $P(\theta|M) = \sqrt{\det J}$, where $J(\theta|M)$ is the Fisher information matrix, $J_{ij}(\theta|M) = E[(\partial_{\theta_i} \log f)(\partial_{\theta_j} \log f)]$. For our case with $\theta = (A, \beta)$ and $f(x|\theta) = At^{\beta}$, the prior is found to be, $P(A, \beta) \sim \frac{1}{A} \frac{1}{\beta+1}$. This is an improper distribution, which we normalize over a suitable parameter ranges $A \in [0.1, 10]$ and $\beta \in [0, 2]$. The ranges are considered suitable as the likelihood function is close to zero for parameter values outside these ranges. We also note that the result of the model comparison does not depend sensitively on the choice of the prior distribution.