

Encoding certainty in bump attractors

Sam Carroll · Krešimir Josić · Zachary P. Kilpatrick

Received: date / Accepted: date

Abstract Persistent activity in neuronal populations has been shown to represent the spatial position of remembered stimuli. Networks that support bump attractors are often used to model such persistent activity. Such models usually exhibit translational symmetry. Thus activity bumps are neutrally stable, and perturbations in position do not decay away. We extend previous work on bump attractors by constructing model networks capable of encoding the certainty or salience of a stimulus stored in memory. Such networks support bumps that are not only neutrally stable to perturbations in position, but also perturbations in amplitude. Possible bump solutions then lie on a two-dimensional attractor, determined by a continuum of positions and amplitudes. Such an attractor requires precisely balancing the strength of recurrent synaptic connections. The amplitude of activity bumps represents certainty, and is determined by the initial input to the system. Moreover, bumps with larger amplitudes are more robust to noise, and over time provide a more faithful representation of the stored stimulus. In networks with separate excitatory and inhibitory populations, generating bumps with a continuum of possible amplitudes, requires tuning the strength of inhibition to precisely cancel background excitation.

Keywords excitation-inhibition balance · bump attractor · neural field · spatial working memory

1 Introduction

Neuronal populations in many cortical areas exhibit sustained activity during the delay period in a spatial working memory task (Wang, 2001; Curtis, 2006). Groups of cells responsive to the presence of a stimulus that needs to be stored in memory can remain active after the stimulus is removed (Rao et al, 2000; Vijayraghavan et al, 2007). Which subset of neurons is active depends on the spatial location of the cue (Funahashi et al, 1989). Such sustained activity has been observed in prefrontal cortex (Goldman-Rakic, 1995), parietal cortex (Pesaran et al, 2002), as well as superior colliculus (Basso and Wurtz, 1997).

Such persistent elevation in firing rates is captured in model networks by “bumps” of activity. The peaks of these activity bumps represent the remembered location of the cue (Compte et al, 2000; Durstewitz et al, 2000; Gutkin et al, 2001). Maintaining a stable activity bump during the delay is hence crucial for representing the remembered cue (Brody et al, 2003). The recurrent architecture of the local

S.R. Carroll
Department of Mathematics, University of Houston,
Houston TX 77204 USA
E-mail: srcarroll314@gmail.com

K. Josić
Department of Mathematics, University of Houston,
Houston TX 77204 USA
E-mail: josic@math.uh.edu

Z.P. Kilpatrick
Department of Mathematics, University of Houston,
Houston TX 77204 USA
E-mail: zpkilpat@math.uh.edu

neuronal networks appears to play a crucial role in maintaining such selective activation (Constantinidis and Wang, 2004). Tuned excitatory neurons reciprocally connect to one another with both fast and slow synapses (Wang, 2001). In addition, inhibitory cells broadly project back to the rest of the network keeping spatial tuning sharp (Rao et al, 2000). Understanding how synaptic architecture can be tuned to produce reliable bumps is essential for understanding the mechanism behind spatial working memory.

Models that can store a continuous range of spatial locations typically possess solutions that are neutrally stable (Amari, 1977; Seung, 1996; Brody et al, 2003). Due to the neutral stability of such attractors, perturbations that change the location of a bump of activity do not decay away (Amari, 1977; Camperi and Wang, 1998; Compte et al, 2000). Aside from experimentally-introduced distractors in spatial working memory experiments (Miller et al, 1996), cue memories can also be degraded by internal variability within cortical networks (Faisal et al, 2008). Stochastic models show that such variability causes bump attractors to wander diffusively, due to their inherent neutral stability (Camperi and Wang, 1998; Compte et al, 2000; Laing and Chow, 2001; Kilpatrick and Ermentrout, 2013). Psychophysical studies show that errors made recalling remembered spatial locations scale roughly linearly with delay time, suggesting the remembered location may diffuse in time (White et al, 1994; Ploner et al, 1998). Also, heterogeneities in the spatial structure of the underlying neuronal network can further degrade the relation between the stored memory and the initial cue (Seung, 1996; Renart et al, 2003; Itskov et al, 2011; Hansel and Mato, 2013). One solution to this problem is to structure the spatial arrangement of excitatory synapses (Kilpatrick and Ermentrout, 2013; Kilpatrick et al, 2013) to make networks robust to dynamic and static parametric perturbations. Thus, the spatial organization of synaptic architecture can play a major role in accurately encoding stimuli for future recall.

To explore the relation between network architecture and the neural computation underlying working memory, we consider bump attractor networks capable of encoding cue *certainty*. We define *certainty* as the likelihood that the presented cue was faithfully communicated to the network generating delay period activity. A number of experiments have shown that the certainty of a decision can be encoded by the instantaneous firing rates of neurons in medial temporal cortex (Shadlen and Newsome, 1998; Gold and Shadlen, 2002; Beck et al, 2008; Kiani and Shadlen, 2009). In this way, the activity of a network can represent the encoded signal as well as the likelihood that the encoded signal accurately represents reality (Zemel et al, 1998). We introduce this notion here in the context of spatial working memory. Recordings from superior colliculus by Basso and Wurtz (1997) suggest that increased uncertainty in a remembered cue position is represented by lower neural activity during the delay period. Conversely, work by Meyer et al (2011) shows that training in a spatial working memory task that leads to improved performance is also accompanied by a rise in delay period firing rates. These observations suggest that certainty about stimulus location in a spatial working memory task may be represented by the level of neural activity during the delay period.

Like cue position, the degree of certainty in a signal is an analog quantity. Phenomenological models of decision making in the presence of two alternatives also frequently exhibit line attractors. In such models the state along the line attractor represents the likelihood, or certainty, that one of the choices is correct (Bogacz et al, 2006). This can be accomplished in more biophysically realistic models by choosing synaptic time constants to match the slope of the input-output relationship of a firing rate model (Wang, 2002). Precisely balancing the rate of feedback excitation with the timescale of synaptic decay leads to a model that behaves as a pure integrator. In the absence of external inputs, networks that behave as pure integrators can store the value of a continuous variable (Goldman et al, 2003). Similar tuning can also be accomplished in mutually inhibitory rate models for parametric working memory (Machens et al, 2005; Polk et al, 2012).

We build on these ideas to study spatial working memory networks that can encode certainty. Typically, bump attractor networks only possess a single stable bump amplitude at each orientation (Amari, 1977; Ermentrout, 1998). We explore networks that support a continuum of bump amplitudes at each orientation. These are dynamical systems that contain two-dimensional attractor surfaces: One dimension corresponded to the amplitude, and the other the position of the bump. The system therefore exhibits a plane attractor. To accomplish this, excitation and inhibition must be balanced, and the shape of the synaptic input to output firing rate function chosen appropriately (Amari, 1977). Namely, there is a monotonic relationship between the total synaptic excitation and total synaptic inhibition that must be maintained to represent a continuum of possible bump amplitudes.

Including a certainty code in networks for bump attractors has several consequences. First, the strength of the original input can be encoded in the amplitude of the bump. Bumps with larger ampli-

tudes stay closer to their original position when the effects of noise are considered during the storage period. However, since certainty is encoded in the amplitude of the bump, memory of the original certainty can also be degraded by dynamic noise. Also, arbitrarily weak inputs can still be stored as a bump attractor, which is not the case in networks that support a single stable bump amplitude (Amari, 1977; Ermentrout, 1998). Using a stochastic neural field model of a bump attractor network with noise, we can develop explicit formulae for most of these results using asymptotic methods.

2 Bump attractor networks

Bump attractor networks were originally developed as general models of recurrent neuronal circuits that can support spatiotemporal patterns of activity (Wilson and Cowan, 1973; Amari, 1977). Since then they have been used to represent activity subserving spatial working memory (Camperi and Wang, 1998; Compte et al, 2000) and visual orientation processing (Ben-Yishai et al, 1995). We consider a spatially organized neural field model where the positions of neurons correspond to their preferred stimulus orientation. We focus on a ring architecture, but we believe these ideas will extend to more general models.

Consider a single population which incorporates local excitation and broadly tuned inhibition (Amari, 1977; Ben-Yishai et al, 1995; Ermentrout, 1998)

$$\frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\pi}^{\pi} w(x, y) f(u(y, t)) dy + I(x, t). \quad (2.1)$$

Here $u(x, t)$ is the total synaptic input to spatial location $x \in [-\pi, \pi]$ at time t . The integral kernel, $w(x, y)$, represents the synaptic feedback from the whole network which encodes the strength of connections from y to x . For a translationally symmetric synaptic weight function, $w(x, y) = \bar{w}(x - y)$, it can be shown that bump solutions will be neutrally stable to translations in position (Amari, 1977; Ermentrout, 1998; Veltz and Faugeras, 2010). This relies on the assumption of a spatial homogeneity in the net excitability of any particular neuron in the network (Renart et al, 2003; Kilpatrick and Ermentrout, 2013). For simplicity, we use a unimodal synaptic weight function

$$w(x, y) = w_0 + w_1 \cos(x - y), \quad (2.2)$$

where w_0 represents the amplitude of broadly tuned inhibition, and w_1 represents the amplitude of locally tuned excitation. Results on the existence and stability of stationary bump solutions can easily be extended to synaptic weights with many more modes (Veltz and Faugeras, 2010). However, the anatomical structure of recurrent connectivity is not known at such a fine level of detail (Rao et al, 2000). We therefore use canonical functions that represent the short range excitation and broad inhibition known to exist.

The nonlinearity, f , is the firing rate function which maps the synaptic inputs, u , to a resulting fraction of active neurons (or probability of activation of a single neuron). Typically, f is a saturating, non-negative function (Coombes and Owen, 2004; Bressloff, 2012). In this study, we consider a piecewise linear firing rate function of the form (Hansel and Sompolinsky, 1998; Pinto and Ermentrout, 2001a; Kilpatrick and Bressloff, 2010)

$$f(u) = \begin{cases} 0, & \text{if } u < \theta \\ s(u - \theta), & \text{if } \theta \leq u \leq \frac{1}{s} + \theta \\ 1, & \text{if } u > \frac{1}{s} + \theta, \end{cases} \quad (2.3)$$

where s is the gain parameter, and θ the threshold. This choice will allow for a straightforward construction of a network capable of storing a continuum of bump amplitudes. In typical analyses of the spatiotemporal dynamics of neural fields, the parameters in Eq. (2.3) are chosen so that the underlying space-clamped system is bistable (Hansel and Sompolinsky, 1998; Pinto and Ermentrout, 2001a). Since we assume the population u is quite large, we make the assumption $\theta = 0$ throughout this study, so arbitrarily weak inputs always activate a small fraction of the population (Hansel and Sompolinsky, 1998).

In addition, we will analyze a two population network containing separate excitatory and inhibitory populations. The model we employ ignores inhibitory-inhibitory interactions – the small proportion of

inhibitory-inhibitory synaptic connections observed in prefrontal cortex is not expected to alter our results substantially (Somogyi et al, 1998). Thus, we consider the system of integro-differential equations (Pinto and Ermentrout, 2001b)

$$\begin{aligned}\frac{\partial u(x, t)}{\partial t} &= -u(x, t) + \int_{-\pi}^{\pi} w_{ee}(x, y) f(u(y, t)) dy - \int_{-\pi}^{\pi} w_{ie}(x, y) v(y, t) dy + I(x, t) \\ \tau \frac{\partial v(x, t)}{\partial t} &= -v(x, t) + \int_{-\pi}^{\pi} w_{ei}(x, y) f(u(y, t)) dy,\end{aligned}\quad (2.4)$$

where $u(x, t)$ is the total synaptic input to the excitatory network and $v(x, t)$ is the total synaptic input to the inhibitory network. The integral kernel, w_{ee} , is the synaptic strength of the excitatory network onto itself, w_{ei} is the strength of the excitatory network onto the inhibitory network, and w_{ie} is the strength of the inhibitory network onto the excitatory network. Additionally, τ is the inhibitory time constant which denotes the speed at which inhibition acts on the excitatory population. We will consider weight functions of the form

$$\begin{aligned}w_{ee}(x) &= \bar{w}_{ee}(1 + \cos x) \\ w_{ei}(x) &= \bar{w}_{ei}(1 + \cos x) \\ w_{ie}(x) &= \bar{w}_{ie}\end{aligned}\quad (2.5)$$

where $\bar{w}_{ee}, \bar{w}_{ei}, \bar{w}_{ie} > 0$. These functions are non-negative. Constant inhibition was chosen both to represent the broader tuning of inhibition compared to excitation, and to ease mathematical analysis.

We note that in the limit of fast inhibition, $\tau \rightarrow 0$, Eq. (2.4) reduces to

$$\begin{aligned}\frac{\partial u(x, t)}{\partial t} &= -u(x, t) + (w_{ee}(x) - w_{ie}(x) * w_{ei}(x)) * f(u(x, t)) + I(x, t) \\ v(x, t) &= \int_{-\pi}^{\pi} w_{ei}(x, y) f(u(y, t)) dy.\end{aligned}\quad (2.6)$$

Here the first equation is equivalent to the single population case in Eq. (2.1), and the second equation has no impact on stability. Therefore, the study of a single population can provide insight into the behavior of the two population network.

We first describe the general procedure for constructing stationary bump solutions with arbitrary firing rate and weight functions in the network (2.1). In the absence of external input, we look for the stationary bump solutions, $u(x, t) = U(x)$, by plugging into Eq. (2.1) and obtaining

$$U(x) = \int_{-\pi}^{\pi} w(x, y) f(U(y)) dy.\quad (2.7)$$

Since $U(x)$ must be periodic we expand it in a Fourier series

$$U(x) = \sum_{k=0}^N A_k \cos(kx) + \sum_{l=1}^N B_l \sin(lx)\quad (2.8)$$

where N is the maximal mode. The assumption of there being a finite number of terms in the Fourier expansion for $U(x)$ relies on the weight function $w(x, y)$ having a finite Fourier expansion. This is reasonable since most typical smooth weight functions can be well approximated by a few terms in a Fourier series (Veltz and Faugeras, 2010). Doing so, allows us to always construct solvable systems for the coefficients of the bump and its stability. In the most general case for a spatially homogeneous weight kernel we write

$$w(x, y) = \bar{w}(x - y) = \sum_{k=0}^N w_k \cos(k(x - y)) = \sum_{k=0}^N w_k [\cos(kx) \cos(ky) + \sin(kx) \sin(ky)],\quad (2.9)$$

so that

$$A_k = w_k \int_{-\pi}^{\pi} \cos(kx) f(U(x)) dx, \quad B_l = w_l \int_{-\pi}^{\pi} \sin(lx) f(U(x)) dx.\quad (2.10)$$

Since the system is translationally symmetric, solutions centered at any position imply a translated solution of that same shape exists. In addition, the system (2.1) will be reflection symmetric as well (Amari, 1977). With this in mind, we look solely for even solutions (Amari, 1977; Veltz and Faugeras, 2010), so that $B_l = 0$ for all l , and (2.8) becomes

$$U(x) = \sum_{k=0}^N A_k \cos(kx). \quad (2.11)$$

By requiring self-consistency we have that

$$A_k = w_k \int_{-\pi}^{\pi} \cos(kx) f\left(\sum_{k=0}^N A_k \cos(kx)\right) dx. \quad (2.12)$$

In general, numerical methods must be used to solve for the coefficients. However for particular functions the solutions can be found analytically as we show in subsequent sections. In addition, we can typically compute the spectrum of the linear system governing perturbations of the bump (Coombes and Owen, 2004; Folias and Bressloff, 2004; Veltz and Faugeras, 2010), which yields relationships between the eigenvalues determining bump stability and parameters of w and f . It is then straightforward to tune parameters to attain neutral stability along the two eigendirections of interest: one corresponding to translations of the bump (position) and the other corresponding to expansions/contractions of the bump (amplitude). We start by demonstrating this in a single population model.

3 Neutral stability in a single population

We first derive conditions for a network that supports bump solutions with a continuum of possible amplitudes. We find that the parameters of the network must be tuned precisely. We find that recurrent excitation must be inversely proportional to the gain of the firing rate function. The resulting network supports a continuum of bump amplitudes, and is capable of encoding certainty in the level of initial activation. This initial activation is controlled by both the duration of cue exposure, and the contrast (intensity) of the cue. That is, longer cue times and/or higher contrast cue lead to higher initial amplitude, corresponding to higher certainty. Finally, we examine the dynamics of bumps during the delay period using a stochastic neural field equation with additive white noise. We find that the spatial diffusion of the bump depends on its amplitude, and stronger initial activation results in more stable bumps. This has interesting implications for working memory: The greater the initial activation, the greater the certainty, and the better the stimulus is remembered during the delay period.

3.1 Stationary Bump Solution

First we construct stationary bump solutions in the absence of external input and noise by plugging the stationary solution $u(x, t) = U(x)$ into Eq. (2.1) assuming a unimodal weight function given in Eq. (2.2). We obtain

$$U(x) = \int_{-\pi}^{\pi} (w_0 + w_1 \cos(x - y)) f(U(y)) dy. \quad (3.1)$$

We then write,

$$U(x) = A_0 + \sum_{k=1}^N A_k \cos(xk) + \sum_{k=1}^N B_k \sin(xk), \quad (3.2)$$

where N is the maximum mode of $U(x)$, and find that

$$\begin{aligned} A_0 &= w_0 \int_{-\pi}^{\pi} f(U(y)) dy, \\ A_1 &= w_1 \int_{-\pi}^{\pi} \cos y f(U(y)) dy, \\ B_1 &= w_1 \int_{-\pi}^{\pi} \sin y f(U(y)) dy, \end{aligned} \quad (3.3)$$

and $A_k = B_k = 0$ for $k \neq 0, 1$. In particular we would like to restrict our analysis to even stationary bump solutions, and we thus set $B_1 = 0$. Since $f(U(x))$ is then even, the last integral in (3.3) will be zero, making the equation for B_1 self-consistent.

Given the piecewise linear firing rate function in Eq. (2.3), it is useful to set certain conditions on the stationary bump solution $U(x)$. To start, we consider bump solutions where $0 \leq U(x) \leq \frac{1}{s}$. In other words, we will consider strictly positive bumps $U(x)$ whose peaks lie below the saturating threshold of (2.3). We will now show that excitation must properly balance inhibition as well as the gain of the firing rate function in order to attain a line of neutrally stable bump amplitudes. When $0 \leq U(x) \leq \frac{1}{s}$, Eq. (3.3) becomes

$$\begin{aligned} A_0 &= 2\pi w_0 s A_0 \\ A_1 &= \pi w_1 s A_1 \end{aligned} \tag{3.4}$$

where $|A_0| \geq A_1$, so that $U(x) > 0$ for all x . We additionally require that $A_1 > 0$, so that the peak of the bump corresponds to the stored spatial position. To ensure that a continuum of values of (A_0, A_1) exist that solve Eq. (3.4) in this case, we require that $s = \frac{1}{\pi w_1}$ and $w_1 = 2w_0$. Thus, excitation w_1 must properly balance inhibition and the tuning of the firing rate gain s must be inversely proportional to excitation for proper parameter balance.

Next, we will examine solutions that obey the restriction $U(x) \leq \frac{1}{s}$, but have some values at which $U(x) < 0$. This corresponds to bumps in which only a portion of the population is active. Again, we will show that the recurrent excitation in the network must be properly balanced by the gain of the system to yield a continuum of allowable bump amplitudes. From our analysis in section 2, we can conclude that if the synaptic weight function w is unimodal, then $U(x)$ may be unimodal as well and the solution will exhibit two roots, $U(\pm a) = 0$. Then the system of equations (3.3) becomes

$$\begin{aligned} A_0 &= 2s w_0 [a A_0 + A_1 \sin a] \\ A_1 &= s w_1 [A_0 \sin a + a A_1]. \end{aligned} \tag{3.5}$$

Requiring self-consistency implies that either $A_1 = A_0$ or $A_0 = 0$ and $A_1 < \frac{1}{s}$. Therefore, we must either have that

$$U(x) = A(1 + \cos x), \tag{3.6}$$

or

$$U(x) = A \cos x. \tag{3.7}$$

Since equation (3.6) satisfies the more restrictive condition $0 \leq U(x) \leq \frac{1}{s}$ that we have already considered, we will focus on bump solutions of the form given by Eq. (3.7). To obtain solutions of this form, we must have $w_0 = 0$. Now using Eq. (3.3) we find that

$$A = \begin{cases} s A w_1 \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{1}{sA} \right) - \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right] + 2w_1 \sqrt{1 - \left(\frac{1}{sA} \right)^2}, & \text{for } sA > 1, \\ s A w_1 \frac{\pi}{2}, & \text{for } sA \leq 1, \end{cases}$$

which requires that $w_1 = \frac{2}{\pi s}$. This means that recurrent excitation must be inversely proportional to the gain of the system. When $sA \leq 1$ we have a continuum of solutions for $A \in [0, \frac{1}{s}]$ that are all stationary solutions. When $sA > 1$ there is no solution for A for the given fixed value of s . Thus the only fixed points are in the interval $[0, \frac{1}{s}]$ as illustrated in Fig. 1A. As we will show in the next section, any amplitude above the threshold will decay back down to the boundary of the line attractor (red dashed plot in Fig. 1A). We also illustrate in Fig. 1B, and show in the stability analysis, that any phase shifted bump will be neutrally stable as well. Thus we have a two-dimensional attractor surface on the closed disc of radius $\frac{1}{s}$ each point of which corresponds to a bump solution.

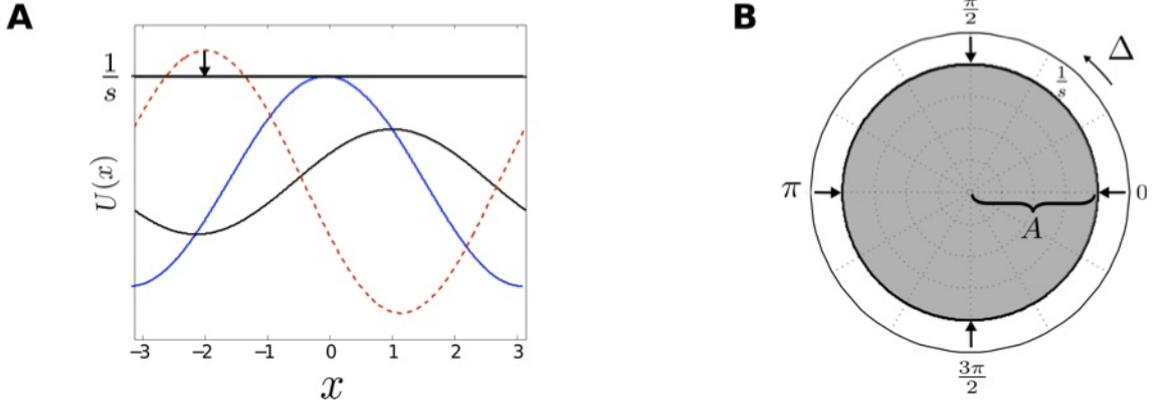


Fig. 1 Stationary Bump Solutions. **(A)** Blue and black indicate neutrally stable solutions for $U(x)$ while red indicates unstable solutions that are attracted to the boundary of the line attractor. **(B)** Polar plot of neutrally stable region, with the amplitude A as the radial parameter and Δ as phase parameter. The grey area shows where solutions of the form $A \cos(x - \Delta)$ are neutrally stable.

3.2 Stability of Bump Solution

To show that the solutions we described above are neutrally stable, we carry out a linear stability analysis of the stationary bump solutions $U(x)$ of Eq. (2.1). In particular, we study the temporal evolution of small, smooth, separable perturbations $e^{\lambda t} \psi(x)$ by plugging in the linear expansion

$$u(x, t) = U(x) + \psi(x)e^{\lambda t} \quad (3.8)$$

where $\|\psi(x)\| \ll 1$. Since $\psi(x)$ must be periodic, we expand it in a Fourier series

$$\psi(x) = \sum_{k=1}^N \mathcal{A}_k \cos kx + \sum_{k=1}^N \mathcal{B}_k \sin kx. \quad (3.9)$$

The perturbed bump solution given in Eq. (3.8) can then be plugged into Eq. (2.1) and linearized to yield the general spectral problem (Ermentrout, 1998; Veltz and Faugeras, 2010)

$$(\lambda + 1)\psi(x) = \int_{-\pi}^{\pi} w(x, y) f'(U(y)) \psi(y) dy \quad (3.10)$$

for the stability of the bump. The associated coefficients of the expansion in Eq. (3.9) are then determined by the linear system

$$\mathcal{A}_k = w_k \int_{-\pi}^{\pi} \cos(kx) f'(U(x)) \psi(x) dx, \quad \mathcal{B}_l = w_l \int_{-\pi}^{\pi} \sin(lx) f'(U(x)) \psi(x) dx, \quad (3.11)$$

where $k, l = 1, \dots, N$. Solutions of this system, along with the associated λ , are eigensolutions of Eq. (3.10). We can directly compute the eigenvalues associated with the stability of bumps in the case of the weight function in Eq. (2.2) so that

$$(\lambda + 1)\psi(x) = w_1 \int_{-\pi}^{\pi} \cos(x - y) f'(U(y)) \psi(y) dy. \quad (3.12)$$

Analyzing solutions (λ, ψ) of Eq. (3.12) is equivalent to determining the elements of the spectrum of the linear system in the vicinity of the bump. We are mainly interested in the point spectrum of the linear operator in Eq. (3.12), since the sign of the real part of λ for these solutions will determine the associated stability of stationary bump solutions (see (Coombes and Owen, 2004; Veltz and Faugeras, 2010) for detailed discussions of the partitioning of spectra in neural field models). In particular, we

examine the stability of stationary bump solutions of the form $U(x) = A \cos x$ when the firing rate function has the form given in Eq. (2.3). Hence,

$$(\lambda + 1)\mathcal{A}_1 = \begin{cases} \mathcal{A}_1 s w_1 \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{1}{sA} \right) - \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right], & \text{for } sA > 1, \\ \mathcal{A}_1 s w_1 \frac{\pi}{2}, & \text{for } sA \leq 1, \end{cases} \quad (3.13)$$

$$(\lambda + 1)\mathcal{B}_1 = \begin{cases} \mathcal{B}_1 s w_1 \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{1}{sA} \right) + \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right], & \text{for } sA > 1, \\ \mathcal{B}_1 s w_1 \frac{\pi}{2}, & \text{for } sA \leq 1, \end{cases} \quad (3.14)$$

and $\mathcal{A}_k = \mathcal{B}_k = 0$ for $k \neq 1$ and $\mathcal{A}_0 = 0$. Now, bump solutions of Eq. (2.1) will be neutrally stable to both even and odd perturbations when parameters in Eqs.(3.13-3.14) are such that some solutions have $\text{Re}\lambda = 0$ and others have $\text{Re}\lambda < 0$.

We show that for the conditions we derived in the previous section, when $A \in [0, \frac{1}{s}]$, this is the case. It is easy to see that if $A \leq \frac{1}{s}$, we obtain $\lambda = 0$ corresponding to both even and odd perturbations by requiring that $w_1 = \frac{2}{\pi s}$ as was the condition for finding the stationary amplitudes in section 3.1. When $A > \frac{1}{s}$, we find

$$\lambda_o = \frac{2}{\pi} \left(-\cos^{-1} \left(\frac{1}{sA} \right) - \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right) < 0, \quad (3.15)$$

$$\lambda_e = \frac{2}{\pi} \left(-\cos^{-1} \left(\frac{1}{sA} \right) + \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right) < 0, \quad (3.16)$$

so that the bump is linearly stable in this region. Thus if the amplitude of the bump exceeds this threshold, it will decay back down to $\frac{1}{s}$. Therefore, the saturating threshold $1/s$ sets an upper limit on the certainty that can be encoded by the amplitude A . The two-dimensional surface of hence consists of neutrally stable bumps to which all solutions are attracted.

3.3 Integrating Input

In an oculomotor delayed response task, an observer is presented with a spatial cue, for a time period T_0 , during which the position of the cue must be remembered. During this period, the cue location is encoded in the network activity by integration of the stimulus. By the end of the presentation a bump of activity arises setting the initial condition for the solution during the delay period. We next propose a way of determining how the cue is integrated by the network. More specifically we study how the amplitude of the bump evolves during this integration period and how it evolves during the delay period, once the cue disappears. We assume the stimulus current has the form

$$I(t) = I_0(t) (H(t) - H(t - T_0)) \cos x, \quad (3.17)$$

so that input starts at $t = 0$, ends at $t = T_0$ and has magnitude $I_0(t)$. We write the solution to Eq. (2.1) in the form

$$u(x, t) = A(t) \cos x.$$

Substituting this into Eq. (2.1) we find

$$\dot{A}(t) = \begin{cases} I_0(t), & \text{if } T_0 \leq t_{max}, \\ g(t), & \text{if } T_0 > t_{max}, \end{cases} \quad (3.18)$$

where $g(t)$ is the piecewise function

$$g(t) = \begin{cases} I_0(t), & \text{if } t < t_{max}, \\ I_0(t) - sA(t)w_1 \left[\cos^{-1} \left(\frac{1}{sA(t)} \right) + \frac{1}{sA(t)} \sqrt{1 - \left(\frac{1}{sA(t)} \right)^2} \right] + 2w_1 \sqrt{1 - \left(\frac{1}{sA(t)} \right)^2}, & \text{if } t > t_{max}, \end{cases} \quad (3.19)$$

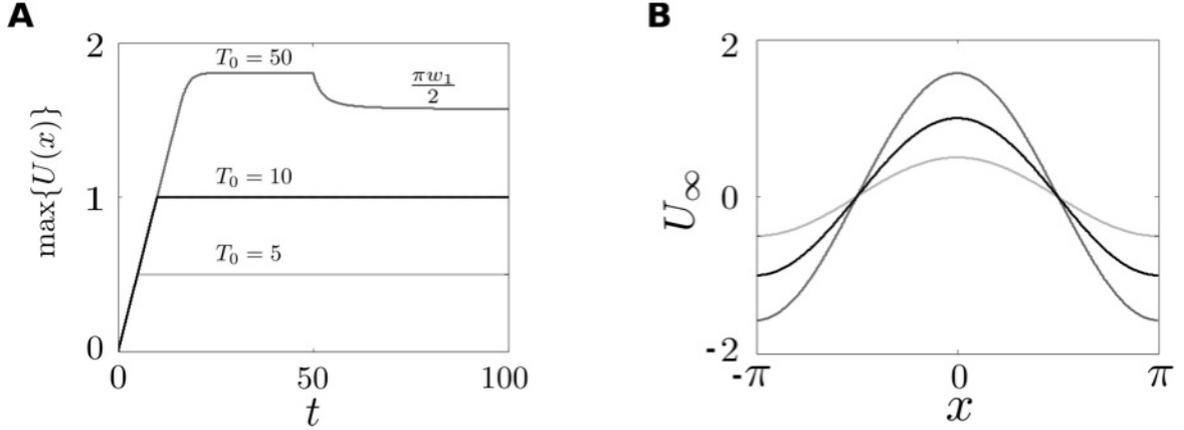


Fig. 2 The integration of constant external input $I(t) = I_0$ lasting for T_0 time units by the single population network (2.1). **(A)** The amplitude of the bump in response to a current injection of the form (3.17). **(B)** Bump solution as $t \rightarrow \infty$.

and t_{max} is the time at which $A(t) = \frac{1}{s}$. We see that, in the line attractor region, the network simply integrates the amplitude of $I(t)$. Numerical methods must be employed to solve for $A(t)$ when it exceeds this value. However for amplitudes beyond this range, solutions are attracted toward the boundary of the line attractor, and hence

$$\lim_{t \rightarrow \infty} A(t) = \begin{cases} \int_0^{T_0} I_0(t), & \text{if } T_0 \leq t_{max}, \\ \frac{\pi w_1}{2}, & \text{if } T_0 > t_{max}. \end{cases} \quad (3.20)$$

Typically, in an experiment, an observer would be exposed to the spatial cue for a short time, T_0 . Furthermore the cue would be at full contrast instantly. We therefore set $I_0(t) \equiv I_0$ to be constant, and find from Eq. (3.18) that if $T_0 \leq t_{max}$ then

$$A(t) = \begin{cases} I_0 t, & \text{for } t \leq T_0, \\ I_0 T_0, & \text{for } t > T_0. \end{cases} \quad (3.21)$$

It is easy to find that

$$t_{max} = \frac{1}{sI_0}$$

by solving $A(t_{max}) = \frac{1}{s}$.

In general, the integration of the spatial cue has a very important consequence in terms of encoding certainty. As mentioned above, experimental evidence suggests that higher uncertainty is preceded by lower neuronal activity and vice versa. Thus we show here that in our model, longer cue exposure leads to greater bump amplitudes which we interpret as greater initial certainty. Additionally, greater cue contrast, I_0 , will also lead to larger bump amplitudes. Therefore we can reach the same bump amplitudes for shorter cue times by increasing the contrast of the cue. It is also important to note that the amplitude eventually saturates to a maximum value, even in the presence of infinite cue time (see Fig. 2A). Therefore we always have an upper bound on the amount of transient initial certainty, corresponding to transient high values of bump amplitude. Once the cue is turned off the amplitude relaxes to the boundary of the line attractor set by the saturating threshold $1/s$ and corresponding to the maximal long-term certainty of the system.

3.4 Diffusion of Bump in a Single Population Network

Cortical neurons *in vivo* typically have high variability in their spike train output (Softky and Koch, 1993), arising from channel noise (White et al, 2000) as well as a high level of background synaptic input not linked to a circuit's immediate task (Faisal et al, 2008). Effects of these fluctuations are typically

incorporated into neural field models by considering finite sized corrections to the mean field (Ginzburg and Sompolinsky, 1994; El Boustani and Destexhe, 2009; Bressloff, 2009). Truncating stochastic terms to linear order then yields Langevin equations that can be analyzed using asymptotic techniques for stochastic partial differential equations (Hutt et al, 2008; Bressloff, 2012). Here, we consider a phenomenological model that incorporates fluctuations into a neural field model, which has recently been used to study the effects of noise on spatiotemporal patterns (Hutt et al, 2008; Bressloff and Webber, 2012; Kilpatrick and Ermentrout, 2013)

$$d\mathcal{U}(x, t) = \left[-\mathcal{U}(x, t) + \int_{-\pi}^{\pi} w(x-y)f(\mathcal{U}(y, t))dy \right] dt + \varepsilon^{1/2}dW(x, t), \quad (3.22)$$

where

$$\langle dW(w, t) \rangle = 0, \quad \langle dW(x, t)dW(y, s) \rangle = C(x-y)\delta(t-s)dtds, \quad (3.23)$$

and $\varepsilon \ll 1$ is the amplitude of noise. In our analysis, we use an ansatz originally used to study the effects of noise on wave propagation in stochastic reaction-diffusion equations (Armero et al, 1998). In line with previous studies (Armero et al, 1998; Bressloff and Webber, 2012; Kilpatrick and Ermentrout, 2013), we assume that the noise term leads to diffusion of the bump's position. In addition, due to the additional neutrally stable direction associated with the bump's amplitude, we assume the amplitude will diffuse in response to noise too. With this in mind we can express the solution, \mathcal{U} , as the sum of a fixed bump profile, U , shifted in its phase by $\Delta(t)$ (which represents the remembered position), increased/decreased in amplitude by amount $\xi(t)$, and higher order time-dependent fluctuations $\varepsilon^{1/2}\Phi + \varepsilon\Phi_1 + \varepsilon^{3/2}\Phi_2 + \dots$ in the profile of the bump. Hence, we write

$$\mathcal{U}(x, t) = (1 + \xi_A(t))U(x - \Delta(t)) + \varepsilon^{1/2}\Phi(x - \Delta(t), t), \quad (3.24)$$

where, for convenience, we use the normalized stochastic variable

$$\xi_A(t) = \frac{\xi(t)}{A_0}. \quad (3.25)$$

Plugging this into Eq. (3.22) we get

$$\begin{aligned} d\Phi(x, t) = & \varepsilon^{-1/2}U'(x)d\Delta(t) + \mathcal{L}\Phi(x, t) - \varepsilon^{-1/2}d\xi_A(t)U(x) \\ & + \varepsilon^{-1/2}\xi_A(t)\mathcal{L}U(x) + \varepsilon^{-1/2}\xi_A(t)U'(x)d\Delta(t) + dW(x, t), \end{aligned} \quad (3.26)$$

where

$$\mathcal{L}p(x) = -p(x) + \int_{-\pi}^{\pi} w(x-y)f'(U(y))p(y)dy. \quad (3.27)$$

In the case of the weight function given by Eq. (2.2) tuned so that the amplitude of the bump is neutrally stable, we also have that $\mathcal{L}U(x) = 0$. Then Eq. (3.26) can be rewritten as

$$d\Phi(x, t) = \varepsilon^{-1/2}U'(x)(1 + \xi_A(t))d\Delta(t) + \mathcal{L}\Phi(x, t) - \varepsilon^{-1/2}d\xi_A(t)U(x) + dW(x, t). \quad (3.28)$$

We can ensure that a bounded solution exists by requiring that the inhomogeneous part of Eq. (3.28) be orthogonal to all elements of the nullspace of the adjoint operator \mathcal{L}^* , (Bressloff, 2001; Kilpatrick and Ermentrout, 2013), where

$$\mathcal{L}^*q(x) = -q(x) + f'(U(x)) \int_{-\pi}^{\pi} w(x-y)q(y)dy \quad (3.29)$$

Therefore, the equation defining the elements $\varphi(x)$ of the nullspace of \mathcal{L}^* is

$$\varphi(x) = f'(U(x)) \int_{-\pi}^{\pi} w(x-y)q(y)dy. \quad (3.30)$$

To identify the nullspace elements of \mathcal{L}^* , recall that we have required neutral stability ($\lambda = 0$) with respect to the linear operator \mathcal{L} , defined in Eq. (3.27), for an odd $\phi_o(x)$ and even $\phi_e(x)$ eigenfunction, so

$$\phi_j(x) = \int_{-\pi}^{\pi} w(x-y)f'(U(y))\phi_j(y)dy, \quad j = o, e. \quad (3.31)$$

Setting $\varphi_j(x) = f'(U(x))\phi_j(x)$ ($j = o, e$) in Eq. (3.30), we have

$$\varphi_j(x) = f'(U(x))\phi_j(x) = f'(U(x)) \int_{-\pi}^{\pi} w(x-y)f'(U(y))\phi_j(y)dy, \quad j = o, e, \quad (3.32)$$

which holds according to Eq. (3.31). Thus, there are two functions that span the nullspace of \mathcal{L}^* : one even function, $\varphi_e(x)$, and one odd function, $\varphi_o(x)$. Taking the inner product of both sides of Eq. (3.28) with respect to each of these functions yields the following equations:

$$\begin{aligned} \int_{-\pi}^{\pi} \varphi_o(x) \left[U'(x)(1 - \xi_A(t))d\Delta(t) + \varepsilon^{1/2}dW(x, t) \right] dx &= 0, \\ \int_{-\pi}^{\pi} \varphi_e(x) \left[-U(x)d\xi_A(t) + \varepsilon^{1/2}dW(x, t) \right] dx &= 0, \end{aligned} \quad (3.33)$$

since $U(x)$ is even and $U'(x)$ is odd. Solving for $d\Delta(t)$ and $d\xi_A(t)$, we find that

$$d\Delta(t) = -\frac{\varepsilon^{1/2}}{1 + \xi_A(t)} \frac{\int_{-\pi}^{\pi} \varphi_o(x)dW(x, t)dx}{\int_{-\pi}^{\pi} \varphi_o(x)U'(x)dx}, \quad (3.34)$$

$$d\xi_A(t) = \varepsilon^{1/2} \frac{\int_{-\pi}^{\pi} \varphi_e(x)dW(x, t)dx}{\int_{-\pi}^{\pi} \varphi_e(x)U(x)dx}, \quad (3.35)$$

and we see that the stochastic variable $\Delta(t)$ depends on $\xi(t)$. Therefore $\Delta(t)$ will not undergo linear diffusion.

We proceed by first computing the distribution of $\xi(t)$, then and use it to find the distribution of $\Delta(t)$. Since $\langle \xi(t) \rangle = 0$ (the additive noise we apply is white in time), computing the variance of $\xi(t)$ we find that it evolves according to pure diffusion since

$$\begin{aligned} \langle \xi(t)^2 \rangle &= \varepsilon A_0^2 \frac{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \varphi_e(x)\varphi_e(y)\langle W(x, t)W(y, t) \rangle dx dy}{\left[\int_{-\pi}^{\pi} \varphi_e(x)U(x)dx \right]^2} t, \\ &= D_{\xi}(\varepsilon)t \end{aligned} \quad (3.36)$$

Using Eq. (3.23) we write the diffusion coefficient as

$$D_{\xi}(\varepsilon) = \varepsilon A_0^2 \frac{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \varphi_e(x)\varphi_e(y)C(x-y)dx dy}{\left[\int_{-\pi}^{\pi} \varphi_e(x)U(x)dx \right]^2}. \quad (3.37)$$

Note that since $\mathcal{L}^*\varphi_e(x) = \mathcal{L}^*\varphi_o(x) = 0$, we use Eq. (3.29) and the general weight function given by Eq. (2.9) to find that

$$\begin{aligned} \varphi_o(x) &= f'(U(x)) \sum_{k=1}^N \mathcal{S}_k \sin(kx), \\ \varphi_e(x) &= f'(U(x)) \sum_{k=1}^N \mathcal{C}_k \cos(kx), \end{aligned} \quad (3.38)$$

where

$$\begin{aligned} \mathcal{S}_k &= w_k \int_{-\pi}^{\pi} \sin(kx)\varphi_o(x), \\ \mathcal{C}_k &= w_k \int_{-\pi}^{\pi} \cos(kx)\varphi_e(x). \end{aligned} \quad (3.39)$$

This system can in general be solved using methods of linear algebra, (Veltz and Faugeras, 2010; Kilpatrick and Ermentrout, 2013). For the weight function $w(x, y) = w_1 \cos(x - y)$ the system simplifies to

$$\begin{aligned} \varphi_o(x) &= w_1 \mathcal{S} f'(A_0 \cos x) \sin x, \\ \varphi_e(x) &= w_1 \mathcal{C} f'(A_0 \cos x) \cos x. \end{aligned} \quad (3.40)$$

Therefore, we can solve Eq. (3.36) by first computing the integral in the denominator as

$$\int_{-\pi}^{\pi} \varphi_e(x)U(x)dx = w_1 \mathcal{C} A_0 \int_{-\pi}^{\pi} f'(A_0 \cos x) \cos^2 x dx = \frac{\pi}{2} s w_1 \mathcal{C} A_0, \quad (3.41)$$

where we impose the condition for neutral stability, $s = \frac{2}{\pi w_1}$. Thus, Eq. (3.36) becomes

$$D_{\xi}(\varepsilon) = \varepsilon \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cos x \cos y f'(U(x)) f'(U(y)) C(x-y) dx dy. \quad (3.42)$$

We can approximate $\Delta(t)$ by expanding the $\xi_A(t)$ dependent term in Eq. (3.34) to second order, so

$$d\Delta(t) = -\varepsilon^{1/2} (1 - \xi_A(t) + \xi_A(t)^2) \frac{\int_{-\pi}^{\pi} \varphi_o(x) dW(x, t) dx}{\int_{-\pi}^{\pi} \varphi_o(x) U'(x) dx}, \quad (3.43)$$

where we know that $\langle \xi_A(t) \rangle = 0$ and $\langle \xi_A(t)^2 \rangle = \frac{1}{A_0^2} \langle \xi(t)^2 \rangle$. For short timescales we can approximate $\Delta(t)$ as a pure diffusion process by ignoring the $\xi_A(t)$ terms. We proceed similarly to find $D_{\xi}(\varepsilon)$ for unimodal solutions $U(x)$. Noting that $U'(x) = -A_0 \sin x$, we find that

$$D_{\Delta}(\varepsilon) = \frac{\varepsilon}{A_0^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sin x \sin y f'(U(x)) f'(U(y)) C(x-y) dx dy. \quad (3.44)$$

Thus the diffusion coefficient decreases as the inverse square of the initial bump amplitude. Since we interpreted this amplitude as a measure of certainty, this implies that the greater the certainty of the stored position the less the position diffuses during the delay period. We have shown that the initial amplitude can be controlled by the duration and contrast at which the spatial cue presented to an observer. Therefore longer exposure times along with contrast will determine the accuracy of recall.

3.5 Calculating effective stochastic motion of bumps

We now compute the diffusion coefficients using cosine shaped spatial correlations, $C(x-y) = \cos(x-y)$. Under this assumption, we find that for $A_0 \in [0, \frac{1}{s}]$

$$D_{\xi}(\varepsilon) = \varepsilon, \quad D_{\Delta}(\varepsilon) = \frac{\varepsilon}{A_0^2}.$$

As demonstrated by comparing single realizations of the stochastic equation (3.34) in Fig. 3A, bumps with initially smaller amplitudes ($A_0 = \pi/16$) diffuse more than bumps with larger amplitudes ($A = \pi/4$). In fact, we can see that both the relationship predicted by Eq. (3.45) and simulations show that the diffusion of the bump's position decreases as the initial amplitude increases (Fig. 3B). Theory and simulations agree well, for noise amplitude of $\varepsilon = 0.001$. To account for dynamics occurring for larger values of ε or longer timescales, we consider nonlinearities in Eq. (3.34). Additionally we must pay special attention to deriving the effective stochastic differential equation for $\xi(t)$, as a linear truncation becomes insufficient.

To derive a more accurate approximation of the variances of $\Delta(t)$ and $\xi(t)$ in the case of cosine shaped spatial noise correlations, we propose a more precise ansatz for the stochastic motion of the bump

$$\mathcal{U}(x, t) = U_0(x) + A_1(t) \cos x + A_2(t) \sin x \quad (3.45)$$

where both $A_1(t), A_2(t)$ are stochastic variables. Using the trigonometric identity for the sum of a sine and a cosine and the initial condition, $U_0(x) = A_0 \cos(x)$, we can write

$$\mathcal{U}(x, t) = \sqrt{(A_0 + A_1(t))^2 + A_2(t)^2} \cos(x - \Delta(t)) \quad (3.46)$$

where $\Delta(t) = \tan^{-1} \left(\frac{A_2(t)}{A_0 + A_1(t)} \right)$. We use the equality, $\xi(t) = \sqrt{(A_0 + A_1(t))^2 + A_2(t)^2} - A_0$, to track the stochastic variable that measures the displacement of the amplitude from its initial point. Therefore as long as this effective amplitude is in the interval $[0, \frac{1}{s}]$, then

$$\mathcal{U}(x, t) = \int_{-\pi}^{\pi} w(x-y) f(\mathcal{U}(y, t)) dy, \quad (3.47)$$

is satisfied by the condition of neutral stability. Therefore Eq. (3.22) simplifies to

$$dA_1(t) \cos x + dA_2(t) \sin x = \varepsilon^{1/2}(dW_1(t) \cos x + dW_2(t) \sin x), \quad (3.48)$$

which we rewrite as

$$\begin{aligned} dA_1(t) &= \varepsilon^{1/2}dW_1(t), \\ dA_2(t) &= \varepsilon^{1/2}dW_2(t). \end{aligned} \quad (3.49)$$

We see that this is equivalent to a 2D diffusion process with initial condition, $(A_1(0), A_2(0)) = (0, 0)$ and

$$\begin{aligned} \langle A_1(t)^2 \rangle &= \langle A_2(t)^2 \rangle = \varepsilon t, \\ \langle A_1(t) \rangle &= \langle A_2(t) \rangle = 0. \end{aligned} \quad (3.50)$$

Therefore to compute the variance of the amplitude we must compute

$$\langle \xi(t)^2 \rangle - \langle \xi(t) \rangle^2 = A_0^2 + 2\varepsilon t - \langle \sqrt{(A_0 + A_1(t))^2 + A_2(t)^2} \rangle^2, \quad (3.51)$$

where the last term must be determined using Monte Carlo simulations. Again, we obtain the relationship that $\Delta(t)$ decreases with increasing initial bump amplitude, A_0 . As shown in Fig. (4B) the Monte Carlo simulation of the stochastic process $\Delta(t)$ agrees well with the full system. Additionally, we see that the variance of $\Delta(t)$ indeed does decrease with initial amplitude, A_0 . Thus, the certainty of bumps relates to how sensitive they will be to stochastic fluctuations. More initial certainty (higher A_0) translates to a lower diffusion coefficient across a broad range of noise amplitudes.

4 Obtaining neutral stability through excitatory-inhibitory balance

So far we have considered a network described by Eq. (2.2) which lumps excitatory and inhibitory cells into a single population. In doing so, we were able to restrict the set of parameters to generate a network which supported a set of bump solutions with a continuum of amplitudes. We can perform a

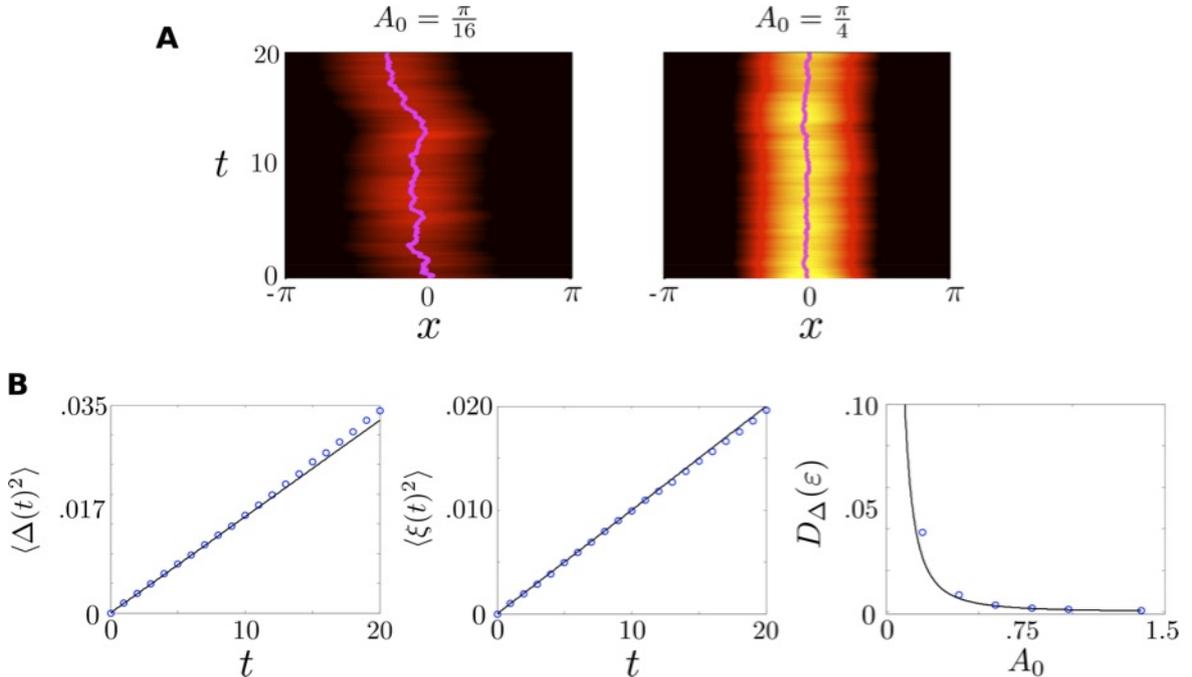


Fig. 3 Diffusion of Bump Solution. **(A)** Space-time plot of bump during the delay period for both low and high initial amplitude. The bump's position diffuses more for smaller initial bump amplitude. **(B)** Variance of the bump center, $\Delta(t)$, and bump amplitude $\xi(t)$ as a function of time as well as the the diffusion coefficient for $\Delta(t)$ as a function of initial amplitude, A_0 . Other parameters used: $\varepsilon = .001$, $s = 2/\pi$, $w_1 = 1$. For the variance plots, $A_0 = \pi/4$

similar analysis in a network with two separate populations for excitatory and inhibitory cells, but our restrictions on parameters simply involve more conditions. As explained in section 2, the single population network is equivalent to the two population network in the limit of fast inhibition, $\tau \rightarrow 0$. However, as inhibition becomes slower, $\tau > 0$, it is no longer clear that stability in the single population implies stability in the two population network. We will see that the conditions for finding a continuum of fixed points are exactly the same as in the single population, since the stationary solutions do not depend on τ . However the equations used in the stability analysis do depend on τ . Excessively slow inhibition can destabilize stationary bump solutions. Perturbations that translate the position of the bump are always neutrally stable, due to the underlying translation symmetry of the network (Bressloff, 2001; Kilpatrick and Ermentrout, 2013). However, generating a network with bumps that are neutrally stable to even symmetric perturbations depends on conditions that relate to the speed of inhibition.

4.1 Stationary Bump Solution

We now study the excitatory-inhibitory network defined by Eq. (2.4) with synaptic weights determined by the functions in Eq. (2.5). First we look for even stationary bump solutions of the form

$$\begin{aligned} u(x, t) &= U(x) = A_0 + A_1 \cos x, \\ v(x, t) &= V(x) = M_0 + M_1 \cos x. \end{aligned} \quad (4.1)$$

Note this ansatz implies time-derivatives $u_t = v_t = 0$ in Eq. (2.4). Thus, by substituting in $V(x)$ equation into the u equation we can generate the single equation

$$U(x) = (w_{ee}(x) - w_{ei}(x) * w_{ie}(x)) * f(U(x)), \quad (4.2)$$

where $f(x) * g(x) = \int_{-\pi}^{\pi} f(x-y)g(y)dy$. Therefore, stationary solutions to Eq. (2.4) are the same as stationary solutions to Eq. (2.1), under the requirement that the assignment of the effective synaptic weight function

$$w(x) = w_{ee}(x) - w_{ei}(x) * w_{ie}(x) = \bar{w}_{ee} - 2\pi\bar{w}_{ei}\bar{w}_{ie} + \bar{w}_{ee} \cos x. \quad (4.3)$$

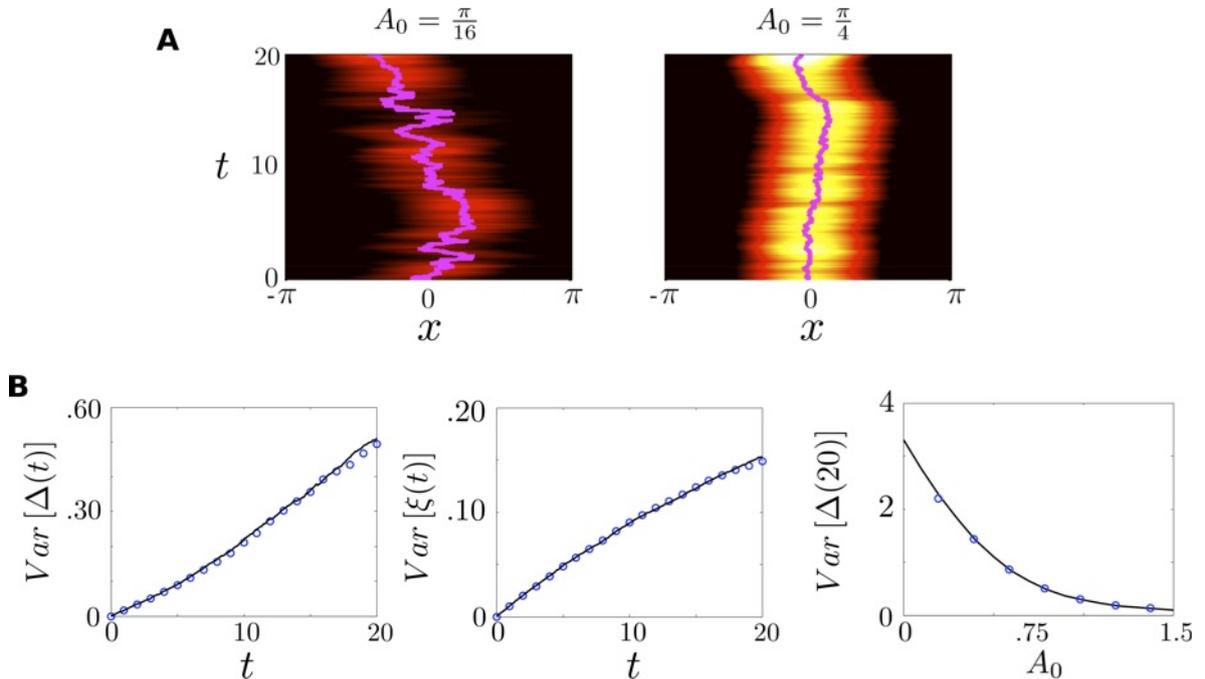


Fig. 4 Diffusion of Bump Solution. **(A)** Space-time plot of bump during the delay period for both low and high initial amplitude. The bump's position diffuses more for smaller initial bump amplitude. **(B)** Variance of the bump center, $\Delta(t)$, and bump amplitude $\xi(t)$ as a function of time as well as the variance for $\Delta(t)$, evaluated at $t = 20$, as a function of initial amplitude, A_0 . Other parameters used: $\varepsilon = .01$, $s = 2/\pi$, $w_1 = 1$. For the variance plots, $A_0 = \pi/4$.

Note that Eq. (4.3) is equivalent to Eq. (2.2) by setting $w_0 = \bar{w}_{ee} - 2\pi\bar{w}_{ei}\bar{w}_{ie}$ and $w_1 = \bar{w}_{ee}$. Therefore, under an appropriate change of variables, solving Eq. (4.2) is equivalent to solving Eq. (3.1). Therefore, our results concerning the existence of a continuum of amplitudes concerning Eq. (3.1) should hold here as well. This means that in order to obtain a line attractor of bump amplitudes, we must have that $A_0 = 0$ and $\bar{w}_{ee} = 2\pi\bar{w}_{ei}\bar{w}_{ie}$ (i.e $w_0 = 0$). However, we can still have $M_0 \neq 0$. Additionally, analogous to the single network in Eq. (3.1), we must require that $\bar{w}_{ee} = \frac{2}{\pi s}$. Again, we have

$$A = \begin{cases} sA\bar{w}_{ee} \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{1}{sA} \right) - \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right] + 2\bar{w}_{ee} \sqrt{1 - \left(\frac{1}{sA} \right)^2}, & \text{for } sA > 1, \\ sA\bar{w}_{ee} \frac{\pi}{2}, & \text{for } sA \leq 1, \end{cases} \quad (4.4)$$

and, for the v equation

$$V(x) = \bar{w}_{ei} \int_{-\pi}^{\pi} (1 + \cos(x-y)) f(U(y)) dy,$$

so that

$$M_0 = \begin{cases} 2s\bar{w}_{ei}A \left[1 - \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right] + 2\bar{w}_{ei} \cos^{-1} \left(\frac{1}{sA} \right), & \text{for } sA > 1, \\ 2s\bar{w}_{ei}A, & \text{for } sA \leq 1, \end{cases}$$

$$M_1 = \begin{cases} sA\bar{w}_{ei} \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{1}{sA} \right) - \frac{1}{sA} \sqrt{1 - \left(\frac{1}{sA} \right)^2} \right] + 2\bar{w}_{ei} \sqrt{1 - \left(\frac{1}{sA} \right)^2}, & \text{for } sA > 1, \\ sA\bar{w}_{ei} \frac{\pi}{2}, & \text{for } sA \leq 1. \end{cases} \quad (4.5)$$

Again, we have a continuum of values for $A \in [0, \frac{\pi\bar{w}_{ee}}{2}]$ that are fixed points, and the coefficients for v will depend on A , and upon substituting values for s we obtain

$$M_0 = \frac{4\bar{w}_{ie}}{\pi\bar{w}_{ee}} A, \quad M_1 = \frac{\bar{w}_{ie}}{\bar{w}_{ee}} A. \quad (4.6)$$

To study the way in which the line attractor globally organizes dynamics, we consider effects of breaking this balance condition in two ways: excess excitation or excess inhibition. As we shall see, too much inhibition leads to no stable bump solutions whereas too much excitation leads to only a single stable bump solution. To do this, we define the quantity

$$\bar{w} = \bar{w}_{ee} - 2\pi\bar{w}_{ei}\bar{w}_{ie} \quad (4.7)$$

and simply consider when $\bar{w} > 0$ (excess excitation) and $\bar{w} < 0$ (excess inhibition).

First let $\bar{w} < 0$ (excess inhibition) and consider when $U(x) < \frac{1}{s}$. Then, similar to section 3.1, we find that

$$A_0 = 2s\bar{w} [aA_0 + \sin aA_1]$$

$$A_1 = s\bar{w}_{ee} [\sin aA_0 + aA_1] \quad (4.8)$$

where $a = \cos^{-1} \left(-\frac{A_0}{A_1} \right)$ and $|A_0| \leq |A_1|$. We must consider the cases when $A_0 > 0$, $A_0 < 0$ and $A_0 = 0$. If $A_0 > 0$, then since $0 \leq a \leq \pi$ we know that $\sin a \geq 0$. Also, we impose that $A_1 \geq 0$ so that the peak of the bump corresponds to the remembered location of the stimulus. Then, since $\bar{w} < 0$, Eq. (4.8) implies that A_0 equals something negative, which is a contradiction. Now assume that $A_0 < 0$. Then Eq. (4.8) implies that

$$a \geq \frac{A_1}{|A_0|} \sin a, \quad \text{and} \quad a \geq \frac{|A_0|}{A_1} \sin a,$$

which implies that $|A_0| = A_1$. Then our only choices are $U(x) = A_1(\cos x - 1)$ or $U(x) \equiv 0$. However, if the former were true, then $f(u) \equiv 0$ which forces $U(x) \equiv 0$. Finally it is easy to see that if $A_0 = 0$ then $A_1 = 0$ for $\bar{w} \neq 0$.

Now assume that $\bar{w} > 0$ (excess excitation). In the case $U(x) < \frac{1}{s}$, we find that the only solution is $U(x) \equiv 0$. When $U(x) > \frac{1}{s}$ for some x , then equation (4.8) becomes

$$A_0 = 2s\bar{w} [(a-b)A_0 + (\sin a - \sin b)A_1] + 2\bar{w}b,$$

$$A_1 = s\bar{w}_{ee} [(\sin a - \sin b)A_0 + (a-b)A_1] + 2\bar{w}_{ee} \sin b, \quad (4.9)$$

where $b = \cos^{-1}\left(\frac{1-sA_0}{sA_1}\right)$ such that $U(b) = \frac{1}{s}$. To simplify the analysis, we will let $s = \frac{2}{\pi\bar{w}_{ee}}$ as was the condition for the line attractor. We cannot solve the system in Eq. (4.9) analytically, so we compute the solutions numerically and plot them in Fig. 5. We see that for any given \bar{w} there is either only a single solution or two solutions for $U(x)$, therefore there are no line attractors in this case. This means that the only case that yields a line attractor of bump amplitudes is when $\bar{w} = \bar{w}_{ee} - 2\pi\bar{w}_{ei}\bar{w}_{ie} = 0$, when recurrent excitation is perfectly balanced by feedback inhibition. Thus, by considering separate excitatory and inhibitory populations, we see that we must place additional restrictions on the parameters of our model to attain a continuum of bump amplitudes. In the next section we will compute the stability of these solutions, showing we must consider the timescale τ of inhibition that feeds back upon the excitatory population.

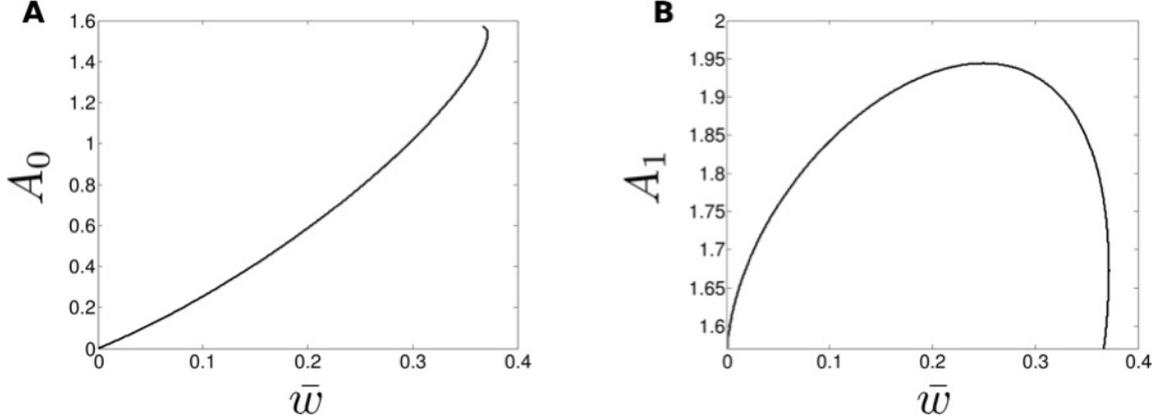


Fig. 5 Amplitudes of the stationary bump solutions $U(x) = A_0 + A_1 \cos x$ for varying \bar{w} values as per equation (4.9). (A) A_0 as a function of \bar{w} . (B) A_1 as a function of \bar{w} . Other values used: $s = \frac{2}{\pi}$, $\bar{w}_{ee} = 1$

4.2 Stability of Bump Solution

We now perform a stability analysis on the fixed bump solution in Eq. (4.1). We consider the set of parameters $A_0 = 0$ and $A_1 = A \in [0, \frac{1}{s}]$ that leads to a line attractor of amplitudes. Similar to section 3.2, we study the temporal evolution of small, smooth, separable perturbations, $e^{\lambda t}\psi(x)$ and $e^{\lambda t}\phi(x)$, by plugging in the linear expansion

$$\begin{aligned} u(x, t) &= U(x) + \psi(x)e^{\lambda t}, \\ v(x, t) &= V(x) + \phi(x)e^{\lambda t}, \end{aligned} \quad (4.10)$$

where $\|\psi(x)\|, \|\phi(x)\| \ll 1$ and since both solutions must be periodic,

$$\begin{aligned} \psi(x) &= \sum_{k=0}^N \mathcal{A}_k \cos(kx) + \sum_{k=1}^N \mathcal{B}_k \sin(kx), \\ \phi(x) &= \sum_{k=0}^N \mathcal{M}_k \cos(kx) + \sum_{k=1}^N \mathcal{N}_k \sin(kx). \end{aligned} \quad (4.11)$$

Plugging the ansatz given by Eq. (4.10) into Eq. (2.4) we obtain

$$\begin{aligned} (\lambda + 1)\psi(x) &= w_{ee} * (f'(U(x))\psi(x)) - w_{ie} * \phi(x), \\ (\tau\lambda + 1)\phi(x) &= w_{ei} * (f'(U(x))\psi(x)). \end{aligned} \quad (4.12)$$

Similar to section 3, we analyze the solutions (λ, ψ, ϕ) to determine the stability of the perturbations by observing the sign of the real part of λ . By self-consistency of (4.11) with (4.12), we see that when using

the weight functions in Eq. (2.5) we have the system

$$\begin{aligned}
(\lambda + 1)\mathcal{A}_0 &= \bar{w}_{ee} \int_{-\pi}^{\pi} (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy - \bar{w}_{ie} \int_{-\pi}^{\pi} (\mathcal{M}_0 + \mathcal{M}_1 \cos y + \mathcal{N}_1 \sin y) dy, \\
(\lambda + 1)\mathcal{A}_1 &= \bar{w}_{ee} \int_{-\pi}^{\pi} \cos y (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy, \\
(\lambda + 1)\mathcal{B}_1 &= \bar{w}_{ee} \int_{-\pi}^{\pi} \sin y (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy, \\
(\tau\lambda + 1)\mathcal{M}_0 &= \bar{w}_{ei} \int_{-\pi}^{\pi} (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy, \\
(\tau\lambda + 1)\mathcal{M}_1 &= \bar{w}_{ei} \int_{-\pi}^{\pi} \cos y (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy, \\
(\tau\lambda + 1)\mathcal{N}_1 &= \bar{w}_{ei} \int_{-\pi}^{\pi} \sin y (\mathcal{A}_0 + \mathcal{A}_1 \cos y + \mathcal{B}_1 \sin y) f'(U(y)) dy,
\end{aligned} \tag{4.13}$$

where $\mathcal{A}_k = \mathcal{B}_k$ for $k \neq 0, 1$. When $\tau \neq 0$, we can compute the integrals and set conditions of the parameters for the line attractor to find that the system in Eq. (4.13) is equivalent to the linear system

$$\lambda \begin{pmatrix} \mathcal{A}_0 \\ \mathcal{A}_1 \\ \mathcal{B}_1 \\ \mathcal{M}_0 \\ \mathcal{M}_1 \\ \mathcal{N}_1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{4}{\pi} & 0 & -2\pi\bar{w}_{ie} & 0 & 0 \\ \frac{4}{\pi} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\pi\bar{w}_{ie}\tau} & \frac{2}{\pi^2\bar{w}_{ie}\tau} & 0 & -\frac{1}{\tau} & 0 & 0 \\ \frac{1}{\pi^2\bar{w}_{ie}\tau} & \frac{1}{2\pi\bar{w}_{ie}\tau} & 0 & 0 & -\frac{1}{\tau} & 0 \\ 0 & 0 & \frac{1}{2\pi\bar{w}_{ie}\tau} & 0 & 0 & -\frac{1}{\tau} \end{pmatrix} \begin{pmatrix} \mathcal{A}_0 \\ \mathcal{A}_1 \\ \mathcal{B}_1 \\ \mathcal{M}_0 \\ \mathcal{M}_1 \\ \mathcal{N}_1 \end{pmatrix}.$$

The associated matrix has the characteristic equation

$$\lambda^2(\tau\lambda + 1)^2 \left(\tau\lambda^2 + (1 - \tau)\lambda + 1 - \frac{16}{\pi^2}\tau \right) = 0 \tag{4.14}$$

from which we obtain only two zero eigenvalues corresponding to odd perturbations $(0, 0, 1, 0, 0, \frac{\bar{w}_{ei}}{\bar{w}_{ee}})$ and even perturbations $(0, 1, 0, \frac{4\bar{w}_{ei}}{\pi\bar{w}_{ee}}, \frac{\bar{w}_{ei}}{\bar{w}_{ee}}, 0)$. Thus we see that obtaining a zero eigenvalue associated with even perturbations does not depend on the speed of inhibition, τ . However, neutral stability still does depend on τ , as it is possible that other eigenvalues associated with even perturbations may have positive real part. Looking at the other eigenvalues, we have two negative ones, $\lambda_- = -\frac{1}{\tau}$, corresponding to perturbations in \mathcal{M}_1 and \mathcal{N}_1 . Therefore, if we only perturb the inhibitory network, then solutions will be attracted back toward the fixed bump solutions. The final two eigenvalues can be analyzed by examining

$$\lambda_{\pm} = \frac{1}{2} \left(1 - \frac{1}{\tau} \right) \pm \frac{1}{2\tau} \sqrt{\left(1 + \frac{64}{\pi^2} \right) \tau^2 - 6\tau + 1}. \tag{4.15}$$

Note that if $\tau = \frac{\pi^2}{16}$, then we obtain one more zero eigenvalue, however the corresponding eigenvector is the zero vector. We plot the two eigenvalues determined by Eq. (4.15) in Fig. 7, showing λ_- is negative for all τ whereas λ_+ is only negative when $\tau < \frac{\pi^2}{16}$. Thus we can ensure that all eigenvalues are zero or have negative real part as long as $\tau < \frac{\pi^2}{16}$. Note that for a certain range of values for τ we have a non-zero imaginary component in λ , however both eigenvalues have negative real part so that oscillatory instabilities never arise.

When $\tau = 0$, then the u equation is equivalent to the single population by letting $w_1 = w_{ee}$. The analysis we performed on the system (2.1) then applies, and we can derive the same conditions for neutral stability. Thus, we conclude that one of the major differences between networks with one and two populations is that it is possible to destabilize stationary bumps with sufficiently slow inhibition (τ large enough). In the case of two populations, the restrictions required to derive a network possessing a line attractor of bump amplitudes generate a relationship implying a monotone increasing correspondence between the strength of excitation and inhibition.

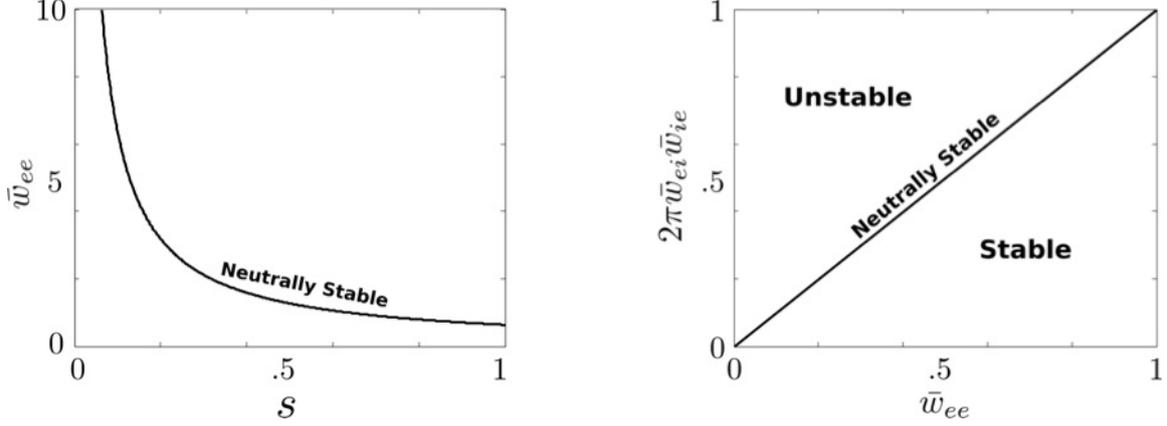


Fig. 6 Stability regions of stationary bumps in two population network, whose amplitudes are determined by equations (4.3), (4.4). Eigenvalues that determine linear stability are given by (4.14). **(A)** The line indicates all the values of \bar{w}_{ee} and s that induce neutral stability in the network. **(B)** The line $\bar{w}_{ee} = 2\pi\bar{w}_{ie}\bar{w}_{ie}$ divides parameter space into unstable region (excess inhibition) and stable region (excess excitation). We achieve neutral stability when excitatory and inhibitory strengths are balanced properly.

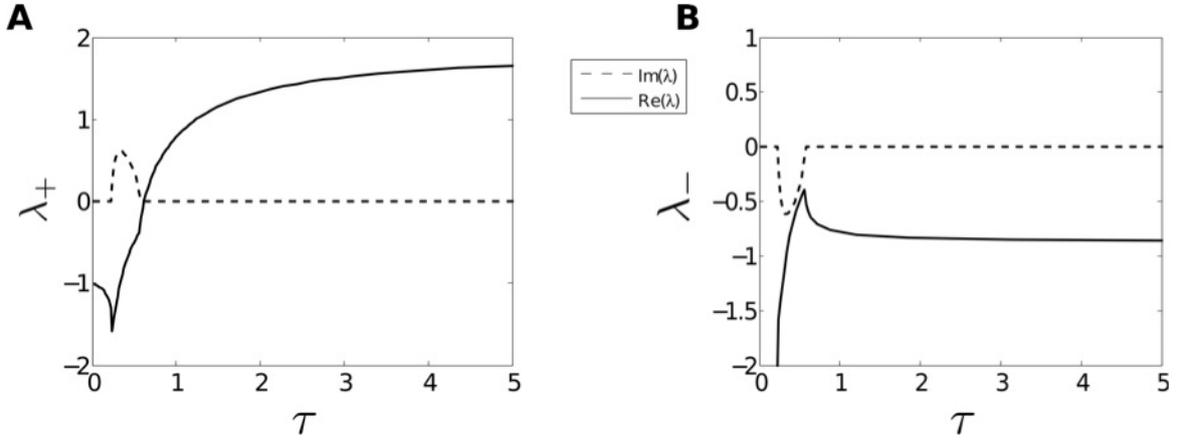


Fig. 7 Dependence of eigenvalues, defined by equation (4.15) plotted against τ for **(A)** λ_+ and **(B)** λ_- .

4.3 Stochastic motion of bumps in the two population network

As in the case of the single population network, we now study how fluctuations affect the motion of bumps in the network (2.4). To do so, we consider a phenomenological model that incorporates noise as an additive term in a Langevin equation

$$\begin{aligned} d\mathcal{U}(x, t) &= \left[-\mathcal{U}(x, t) + \int_{-\pi}^{\pi} w_{ee}(x, y) f(\mathcal{U}(y, t)) dy - \int_{-\pi}^{\pi} w_{ie}(x, y) \mathcal{V}(y, t) dy \right] dt + \varepsilon^{1/2} dW(x, t) \\ \tau d\mathcal{V}(x, t) &= \left[-\mathcal{V}(x, t) + \int_{-\pi}^{\pi} w_{ei}(x, y) f(\mathcal{U}(y, t)) dy \right] dt \end{aligned} \quad (4.16)$$

with

$$\langle dW(x, t) \rangle = 0, \quad \langle dW(x, t) dW(y, s) \rangle = C(x - y) \delta(t - s) dt ds, \quad (4.17)$$

where ε parameterizes the level of noise. Note that, because we could convert the system in Eq. (4.16) into a single second order stochastic differential equation, it would be redundant to include a noise term in the \mathcal{V} equation. Rather than performing an asymptotic analysis on the system (4.16), we now briefly present the results of numerical simulations, showing that bumps diffuse in a similar way to the single population network (see Fig. 8A).

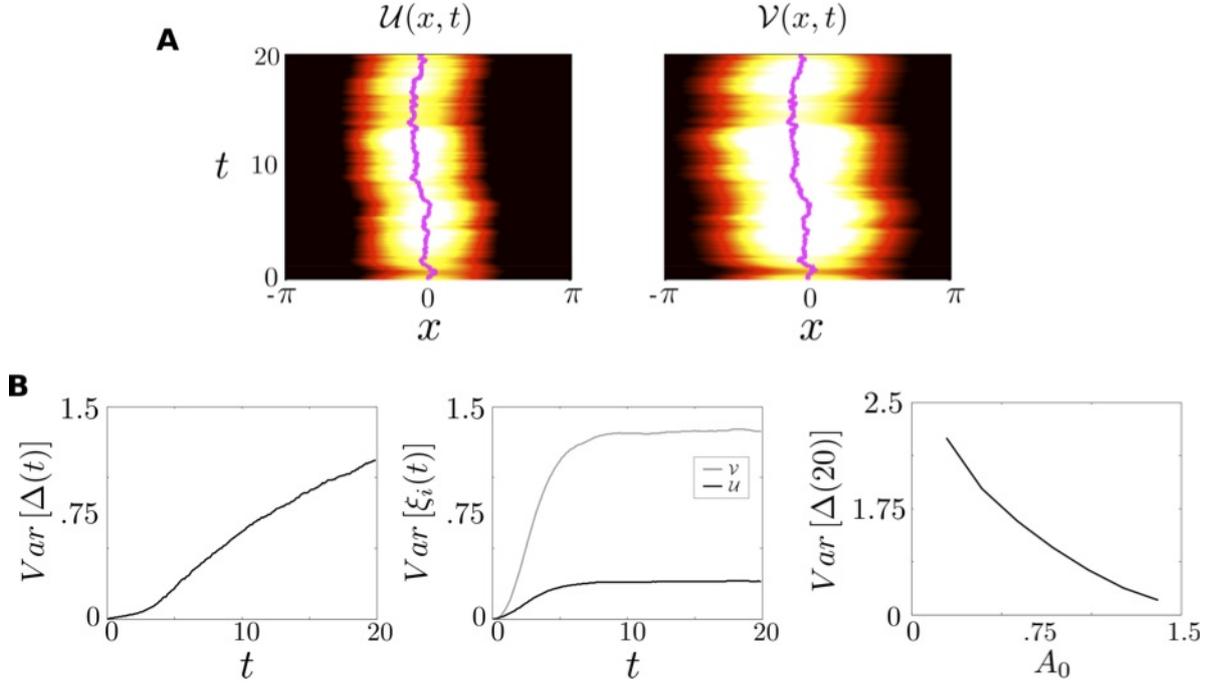


Fig. 8 Diffusion of bump solution in the two population E-I network. **(A)** Space-time plot of bump during the delay period for both the excitatory network, \mathcal{U} , and inhibitory network, \mathcal{V} . **(B)** Variances of the bump center (left) and peak of bump (middle) as well as the variance for $\Delta(t)$, evaluated at $t = 20$, as a function of initial amplitude. Other parameters used: $\varepsilon = .01$, $s = 2/\pi$, $\bar{w}_{ee} = \bar{w}_{ei} = 1$, $\bar{w}_{ie} = 1/2\pi$, $\tau = .4$, $A_1 = \pi/4$

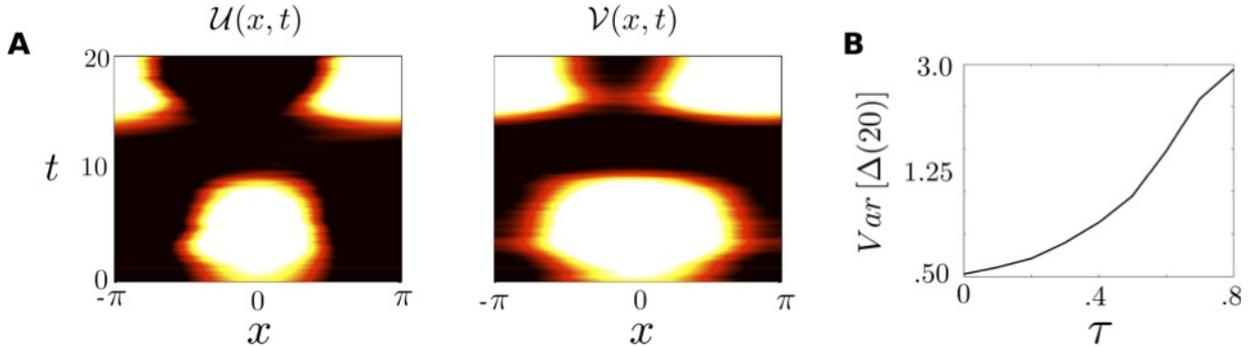


Fig. 9 **(A)** Bump extinction occurring when $\tau = .7$. **(B)** Variance of $\Delta(t)$ at $t = 20$ as a function of τ . Other values used: $\varepsilon = .01$, $s = 2/\pi$, $\bar{w}_{ee} = \bar{w}_{ei} = 1$, $\bar{w}_{ie} = 1/2\pi$, $A_1 = \pi/4$.

First we note that in the limit of fast inhibition, $\tau \rightarrow 0$, Eq. (4.16) is equivalent to Eq. (3.22). Therefore, we would expect identical results for the numerically computed variance of the bump position and amplitude as compared to the single population model in this limit. Moving away from this limit we see slightly different behaviors in the variances of the bump center and maximum as shown in Fig. 8B. First, we see that the variance of the position $\Delta(t)$ of the bump scales more quickly than in the single population network. One reason for this may be that the excitatory and inhibitory populations are not instantaneously coupled to one another, so the transient behavior of the excitatory population involves more activation than in the single population network. Next, we note that the variance of the bump amplitude saturates more quickly than in the single population network. Based on the stability analysis we have carried out, we speculate that this saturation may arise because the initial build up in variance is mostly along weakly stable eigendimensions. After this, variance in the amplitude may continue to build up along the remaining neutrally stable eigendimension, happening at a considerably slower pace. However, fully understanding this behavior will require studying the stochastic system in depth, which we leave for future work. Finally, we still see that the variance of the bump center decreases with increasing

initial amplitude as we saw in the single population. Once again, this implies that a signal that initially possesses more certainty will be more robust to dynamic fluctuations during the storage period.

We showed in section 4.2 that all the remaining nonzero eigenvalues increase towards zero as τ increases for $\tau \leq \pi^2/16$. Thus we would expect that as inhibition becomes slower (τ increases), the variance of the bump’s position would increase due to the stability of the bump to certain odd perturbations becomes weaker. We observe this numerically across a broad range of τ values in Fig. 9B. Essentially, when the inhibitory population does not respond as quickly to transient motion of the excitatory population, noise causes the bump’s position to alter more rapidly. Additionally, for $\tau > \pi/16$, one of the eigenvalues associated with the linear stability of the bump becomes positive. In this case we would expect a total loss of the instantiated bump solution, which we can associated with a total loss of the remembered position, as seen in Fig. 9A with $\tau = .7$. Thus, by considering two separate excitatory and inhibitory populations, we see that the speed of inhibition plays a crucial role in the response of the bump to noise. Previously, Pinto and Ermentrout (2001a) showed that bumps on an unbounded domain are destabilized by oscillatory instabilities for sufficiently slow inhibition. Here, we extend this work by showing non-oscillatory instabilities can occur on the bounded domain of a ring, but these instabilities are still associated with bump extinction.

5 Discussion

We have derived conditions under which networks can support bumps with a continuum of amplitudes. While the location of the bump represents the stored location of a stimulus, its amplitude can represent the certainty of this internal representation. These stationary bump solutions are neutrally stable to perturbations in position as well as amplitude. Our analysis shows that recurrent excitation must be balanced by inhibition to obtain a network that supports such neutrally stable bumps. When considering the effects of noise, we find that the bump diffuses away from its initial position. Not only does the position in orientation space change, but so does the amplitude. Using asymptotic approximations, we can relate the parameters of the model to the effective amount of diffusion the bump will experience. Importantly, bumps with larger initial amplitude diffuse less than bumps with smaller initial amplitudes. Therefore, the amount of certainty initially attached to the stored stimulus determines the fragility of the memory.

We believe this work contributes to the established claims concerning the importance of tuning excitation and inhibition in cortical networks to support flexible computations for cognitive tasks (Brunel and Wang, 2003; Haider et al, 2006; Yizhar et al, 2011). Note that by deriving conditions under which a network supports a neutrally stable line attractor, we tune parameters of the model so its dynamics lie right at a bifurcation, hence moving the system to criticality. Pharmacological manipulations of cortical networks have recently revealed that a precise balance of excitation and inhibition in cortical networks is crucial for criticality, and in this state a network can maximize the range of inputs it can process (Shew et al, 2009). The fact that balancing excitation and inhibition in a network can lead to an increase in the rate of information transfer was originally shown by the work of van Vreeswijk and Sompolinsky (1996). As opposed to the work of Shew et al (2009), which examines the total transient activation of a network in response to stimuli, we are considering the persistent spatially-dependent response of a network encoding cue position. Nonetheless, we note that by balancing excitation and inhibition, we have created networks that can transfer additional information about the certainty of a stimulus. Without this balance, the amplitude of the bump would always relax to a single value.

The networks we considered are not only useful for studying spatial working memory tasks, but could also be used as neural circuit models of decision making (Gold and Shadlen, 2002). Balancing synaptic feedback with the timescale of synaptic decay is a well-established way to obtain model networks capable of performing two-alternative forced choice (2AFC) tasks (Wang, 2002; Bogacz et al, 2006). External inputs can then be slowly integrated along the direction of their bias. In 2AFC tasks, there are only two possible directions (Gold and Shadlen, 2002). In this work, we have developed a model capable of integrating inputs that can be biased to one of a continuum of dimensions around a ring. Thus, we suggest the two-dimensional attractor we have derived in this work may be ideal for integrating information in the presence of a continuum of alternatives. Recent recordings from lateral intraparietal cortex suggest there are neurons whose firing rates climb in correspondence to an animal’s certainty about one of a multitude of decisions (Churchland et al, 2008). Subsequent computational work suggested an entire circuit may

then be capable of sequentially updating a probability function that provides the instantaneous likelihood of each alternative being true, based on accumulated input (Beck et al, 2008). Our model may provide a complementary network implementation of a decision making circuit. Of course, depending on the form of the synaptic weight function, the accumulating persistent activity would depend differently on the timing of inputs. We will explore these issues in future studies.

References

- Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol Cybern* 27(2):77–87, URL <http://dx.doi.org/10.1007/BF00337259>
- Armero J, Casademunt J, Ramirez-Piscina L, Sancho JM (1998) Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise. *Phys Rev E* 58:5494–5500, DOI 10.1103/PhysRevE.58.5494, URL <http://link.aps.org/doi/10.1103/PhysRevE.58.5494>
- Basso MA, Wurtz RH (1997) Modulation of neuronal activity by target uncertainty. *Nature* 389(6646):66–9, DOI 10.1038/37975
- Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A (2008) Probabilistic population codes for bayesian decision making. *Neuron* 60(6):1142–52, DOI 10.1016/j.neuron.2008.09.021
- Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A* 92(9):3844–8
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113(4):700–65, DOI 10.1037/0033-295X.113.4.700
- Bressloff PC (2001) Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D* 155(1-2):83–100
- Bressloff PC (2009) Stochastic neural field theory and the system-size expansion. *SIAM Journal on Applied Mathematics* 70(5):1488–1521, DOI 10.1137/090756971, URL <http://link.aip.org/link/?SMM/70/1488/1>
- Bressloff PC (2012) Spatiotemporal dynamics of continuum neural fields. *J Phys A: Math Theor* 45(3):033,001, URL <http://stacks.iop.org/1751-8121/45/i=3/a=033001>
- Bressloff PC, Webber MA (2012) Front propagation in stochastic neural fields. *SIAM J Appl Dyn Syst* 11(2):708–740
- Brody CD, Romo R, Kepecs A (2003) Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr Opin Neurobiol* 13(2):204–11
- Brunel N, Wang XJ (2003) What determines the frequency of fast network oscillations with irregular neural discharges? i. synaptic dynamics and excitation-inhibition balance. *J Neurophysiol* 90(1):415–30, DOI 10.1152/jn.01095.2002
- Camperi M, Wang XJ (1998) A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J Comput Neurosci* 5(4):383–405
- Churchland AK, Kiani R, Shadlen MN (2008) Decision-making with multiple alternatives. *Nature neuroscience* 11(6):693–702
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10(9):910–23
- Constantinidis C, Wang XJ (2004) A neural circuit basis for spatial working memory. *Neuroscientist* 10(6):553–65, DOI 10.1177/1073858404268742
- Coombes S, Owen MR (2004) Evans functions for integral neural field equations with heaviside firing rate function. *SIAM Journal on Applied Dynamical Systems* 3(4):574–600
- Curtis CE (2006) Prefrontal and parietal contributions to spatial working memory. *Neuroscience* 139(1):173–80, DOI 10.1016/j.neuroscience.2005.04.070
- Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. *Nat Neurosci* 3 Suppl:1184–91, DOI 10.1038/81460
- El Boustani S, Destexhe A (2009) A master equation formalism for macroscopic modeling of asynchronous irregular activity states. *Neural Comput* 21(1):46–100, DOI 10.1162/neco.2009.02-08-710
- Ermentrout B (1998) Neural networks as spatio-temporal pattern-forming systems. *Reports on Progress in Physics* 61(4):353, URL <http://stacks.iop.org/0034-4885/61/i=4/a=002>

- Faisal A, Selen L, Wolpert D (2008) Noise in the nervous system. *Nature Reviews Neuroscience* 9(4):292–303
- Folias SE, Bressloff PC (2004) Breathing pulses in an excitatory neural network. *SIAM Journal on Applied Dynamical Systems* 3(3):378–407
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61(2):331–49
- Ginzburg I, Sompolinsky (1994) Theory of correlations in stochastic neural networks. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 50(4):3171–3191
- Gold JI, Shadlen MN (2002) Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36(2):299–308
- Goldman MS, Levine JH, Major G, Tank DW, Seung Ha (2003) Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral Cortex* 13(11):1185–1195, DOI 10.1093/cercor/bhg095, URL <http://cercor.oxfordjournals.org/content/13/11/1185.abstract>, <http://cercor.oxfordjournals.org/content/13/11/1185.full.pdf+html>
- Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14(3):477–85
- Gutkin B, Laing C, Colby C, Chow C, Ermentrout G (2001) Turning on and off with excitation: The role of spike-timing asynchrony and synchrony in sustained neural activity. *J Comput Neurosci* 11(2):121–134
- Haider B, Duque A, Hasenstaub AR, McCormick DA (2006) Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J Neurosci* 26(17):4535–45, DOI 10.1523/JNEUROSCI.5297-05.2006
- Hansel D, Mato G (2013) Short-term plasticity explains irregular persistent activity in working memory tasks. *J Neurosci* 33(1):133–49, DOI 10.1523/JNEUROSCI.3455-12.2013
- Hansel D, Sompolinsky H (1998) Modeling feature selectivity in local cortical circuits. In: Koch C, Segev I (eds) *Methods in neuronal modeling: From ions to networks*, Cambridge: MIT, chap 13, pp 499–567
- Hutt A, Longtin A, Schimansky-Geier L (2008) Additive noise-induced Turing transitions in spatial systems with application to neural fields and the Swift–Hohenberg equation. *Physica D: Nonlinear Phenomena* 237(6):755–773
- Itskov V, Hansel D, Tsodyks M (2011) Short-term facilitation may stabilize parametric working memory trace. *Front Comput Neurosci* 5:40, DOI 10.3389/fncom.2011.00040
- Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324(5928):759–64, DOI 10.1126/science.1169405
- Kilpatrick ZP, Bressloff PC (2010) Effects of adaptation and synaptic depression on spatiotemporal dynamics of an excitatory neuronal network. *Physica D* 239:547–560, URL <http://dx.doi.org/10.1016/j.physd.2009.06.003>
- Kilpatrick ZP, Ermentrout B (2013) Wandering bumps in stochastic neural fields. *SIAM J Appl Dyn Syst* 12:61–94
- Kilpatrick ZP, Ermentrout B, Doiron B (2013) Optimizing working memory with spatial heterogeneity of recurrent cortical excitation. submitted
- Laing CR, Chow CC (2001) Stationary bumps in networks of spiking neurons. *Neural Comput* 13(7):1473–94, DOI 10.1162/089976601750264974
- Machens C, Romo R, Brody C (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 307(5712):1121–1124
- Meyer T, Qi XL, Stanford TR, Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J Neurosci* 31(17):6266–76, DOI 10.1523/JNEUROSCI.6798-10.2011
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16(16):5154–67
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA (2002) Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci* 5(8):805–11, DOI 10.1038/nn890
- Pinto DJ, Ermentrout GB (2001a) Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses. *SIAM J Appl Math* 62(1):206–225, URL <http://dx.doi.org/10.1137/S0036139900346453>
- Pinto DJ, Ermentrout GB (2001b) Spatially structured activity in synaptically coupled neuronal networks: II. lateral inhibition and standing pulses. *SIAM Journal on Applied Mathematics* 62(1):226–243

- Ploner CJ, Gaymard B, Rivaud S, Agid Y, Pierrot-Deseilligny C (1998) Temporal limits of spatial working memory in humans. *Eur J Neurosci* 10(2):794–7
- Polk A, Litwin-Kumar A, Doiron B (2012) Correlated neural variability in persistent state networks. *Proc Natl Acad Sci U S A* 109(16):6295–300, DOI 10.1073/pnas.1121274109
- Rao SG, Williams GV, Goldman-Rakic PS (2000) Destruction and creation of spatial tuning by disinhibition: Gaba(a) blockade of prefrontal cortical neurons engaged by working memory. *J Neurosci* 20(1):485–94
- Renart A, Song P, Wang XJ (2003) Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron* 38(3):473–85
- Seung HS (1996) How the brain keeps the eyes still. *Proc Natl Acad Sci USA* 93(23):13,339–44
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18(10):3870–96
- Shew WL, Yang H, Petermann T, Roy R, Plenz D (2009) Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *J Neurosci* 29(49):15,595–600, DOI 10.1523/JNEUROSCI.3864-09.2009
- Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J Neurosci* 13(1):334–50
- Somogyi P, Tamás G, Lujan R, Buhl EH (1998) Salient features of synaptic organisation in the cerebral cortex. *Brain Res Brain Res Rev* 26(2-3):113–35
- Veltz R, Faugeras O (2010) Local/global analysis of the stationary solutions of some neural field equations. *SIAM J Appl Dyn Syst* 9:954–998, DOI 10.1137/090773611, URL <http://link.aip.org/link/?SJA/9/954/1>
- Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AFT (2007) Inverted-u dopamine d1 receptor actions on prefrontal neurons engaged in working memory. *Nat Neurosci* 10(3):376–84, DOI 10.1038/nm1846
- van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274(5293):1724–6
- Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* 24(8):455–63
- Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36(5):955–68
- White JA, Rubinstein JT, Kay AR (2000) Channel noise in neurons. *Trends in neurosciences* 23(3):131–137
- White JM, Sparks DL, Stanford TR (1994) Saccades to remembered target locations: an analysis of systematic and variable errors. *Vision Res* 34(1):79–92
- Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol Cybern* 13(2):55–80, URL <http://dx.doi.org/10.1007/BF00288786>
- Yizhar O, Fenno LE, Prigge M, Schneider F, Davidson TJ, O’Shea DJ, Sohal VS, Goshen I, Finkelstein J, Paz JT, Stehfest K, Fudim R, Ramakrishnan C, Huguenard JR, Hegemann P, Deisseroth K (2011) Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* 477(7363):171–8, DOI 10.1038/nature10360
- Zemel RS, Dayan P, Pouget A (1998) Probabilistic interpretation of population codes. *Neural Comput* 10(2):403–30