Interareal coupling reduces encoding variability in multi-area models of spatial working memory

Zachary P. Kilpatrick *

Department of Mathematics, University of Houston, Houston, TX, USA

Persistent activity observed during delayed-response tasks for spatial working memory (Funahashi et al., 1989) has commonly been modeled by recurrent networks whose dynamics is described as a *bump attractor* (Compte et al., 2000). We examine the effects of interareal architecture on the dynamics of bump attractors in stochastic neural fields. Lateral inhibitory synaptic structure in each area sustains stationary bumps in the absence of noise. Introducing noise causes bumps in individual areas to wander as a Brownian walk. However, coupling multiple areas together can help reduce the variability of the bump's position in each area. To examine this quantitatively, we approximate the position of the bump in each area using a small noise expansion that also assumes weak amplitude interareal projections. Our asymptotic results show the motion of the bumps in each area can be approximated as a multivariate Ornstein–Uhlenbeck process. This shows reciprocal coupling between areas can always reduce variability, if sufficiently strong, even if one area contains much more noise than the other. However, when noise is correlated between areas, the variability-reducing effect of interareal coupling is diminished. Our results suggest that distributing spatial working memory representations across multiple, reciprocally-coupled brain areas can lead to noise cancelation that ultimately improves encoding.

Keywords: neural field, bump attractor, spatial working memory, correlations, noise cancelation

INTRODUCTION

Persistent spiking activity has been experimentally observed in prefrontal cortex (Funahashi et al., 1989; Miller et al., 1996), parietal cortex (Colby et al., 1996; Pesaran et al., 2002), superior colliculus (Basso and Wurtz, 1997), caudate nucleus (Hikosaka et al., 1989; Levy et al., 1997), and globus pallidus (Mushiake and Strick, 1995; McNab and Klingberg, 2008) during the retention interval of visuospatial working memory tasks. Often, the subject must remember a cue's location for several seconds (Funahashi et al., 1989). Delay period neurons persistently fire in response to a preferred cue orientation as described by a bell-shaped tuning curve. Networks of these neurons, with recurrent excitation between similarly tuned neurons and broadly tuned feedback inhibition, can generate spatially localized "bumps." The position of these bumps encodes the remembered location of the cue (Compte et al., 2000).

Dynamic variability can degrade the accuracy of working memory over time though. Fluctuations in membrane voltage and synaptic conductance can lead to spontaneous spike or failure events at the edge of the bump, causing the bump to wander diffusively (Compte et al., 2000; Laing and Chow, 2001). Bump attractor networks are particularly prone to such diffusive error because bump positions lie on a line attractor where each location is neutrally stable (Amari, 1977). Interestingly, psychophysical data demonstrates spatial working memory error does scale linearly with delay time, suggesting the underlying process that degrades memory is diffusive (White et al., 1994;

Ploner et al., 1998). Much theoretical work has examined network properties that might limit memory degradation. Several computational studies have explored networks built from bistable neuronal units, which sustain persistent states that are less susceptible to noise (Camperi and Wang, 1998; Koulakov et al., 2002; Goldman et al., 2003). In addition, synaptic facilitation has been shown to slow the drift of bump position due to internal variability (Itskov et al., 2011). Synaptic plasticity has also be shown to reduce diffusion of bumps in (Hansel and Mato, 2013). Finally, spatially heterogeneous recurrent excitation can reduce wandering of bumps quantizing the line attractor by stabilizing a finite set of bump locations (Kilpatrick and Ermentrout, 2013; Kilpatrick et al., 2013).

Complementary to these possibilities, we propose that interareal coupling across multiple areas of cortex may reduce error in working memory recall generated by dynamic fluctuations. Multiple representations of spatial working memory have been identified in different cortical areas (Colby et al., 1996). This distributed representation makes working memory information readily available for motor (Owen et al., 1996) and decision-making (Curtis and Lee, 2010) tasks. In addition, this redundancy may serve to reduce degrading effects of noise. It is known that several areas involved in oculomotor delayed response tasks are reciprocally coupled to one another (Constantinidis and Wang, 2004; Curtis, 2006). We presume the representation of a spatial working memory in a single area takes the form of a bump in a recurrently coupled neural field. Projections between

areas share information about bump position across the multi-area network. Recently, (Folias and Ermentrout, 2011) showed several novel activity patterns emerge when considering neural fields with multiple areas. In addition, recent analyses of spatiotemporal dynamics of perceptual rivalry have exploited dual population neural field models, where activity in each area represents a single percept (Kilpatrick and Bressloff, 2010; Bressloff and Webber, 2012b). In this study, we focus on activity patterns where bumps in each area have positions that remain close.

Our study mostly focuses on a dual area model of spatial working memory, where each area provides a replicate representation of the presented cue. We begin by demonstrating the neutral stability of the bump position in each area, in the absence of noise and interareal projections. Upon including noise and interareal projections, we use a small-noise expansion to derive an effective stochastic differential equation for the position of the bump in each area. The effective system is a multivariate Ornstein–Uhlenbeck process, which we can analyze using diagonalization. The variance of this stochastic process decreases as the strength of connections between areas increases. Variance reduction relies on cancelations arising due to averaging noise between both areas. Thus, when noise is strongly correlated between areas, the effect of interareal coupling is negligible. Lastly, we show this analysis extends to the case of N (more than two) areas and that for sufficiently strong interareal connections, variance scales as $1/N$.

MATERIALS AND METHODS

DUAL AREA MODEL OF SPATIAL WORKING MEMORY

We consider a recurrently coupled model commonly used for spatial working memory (Camperi and Wang, 1998; Ermentrout, 1998) and visual processing (Ben-Yishai et al., 1995). GABAergic inhibition (Gupta et al., 2000) typically acts faster than excitatory NMDAR kinetics (Clements et al., 1992), and we assume excitatory synapses contain a mixture of AMPA and NMDA components. Thus, we make the assumption that inhibition is slaved to excitation as in (Amari, 1977). We can then describe average activity $u_1(x, t)$ and $u_2(x, t)$ of neurons in either area by the system (Ben-Yishai et al., 1995; Folias and Ermentrout, 2011; Kilpatrick and Ermentrout, 2013)

$$\tau \mathrm{d}u_1(x, t) = \left[-u_1 + w_{11} * f(u_1) + \varepsilon^{1/2} w_{12} * f(u_2)\right] \mathrm{d}t$$
$$+ \varepsilon^{1/2} \mathrm{d}W_1(x, t), \tag{1a}$$

$$\tau \mathrm{d}u_2(x, t) = \left[-u_2 + w_{22} * f(u_2) + \varepsilon^{1/2} w_{21} * f(u_1)\right] \mathrm{d}t$$
$$+ \varepsilon^{1/2} \mathrm{d}W_2(x, t), \tag{1b}$$

where the effects of synaptic architecture are described by the convolution

$$w_{jk} * f(u_k) = \int_{-\pi}^{\pi} w_{jk}(x - y) f(u_k(y, t)) \mathrm{d}y, \tag{2}$$

for $j, k = 1, 2$, so the case $j = k$ describes recurrent synaptic connections within a area and $j \neq k$ describes synaptic connections

between areas (interareal). Several fMRI and electrode recordings have revealed correlations between activity in multiple cortical areas during spatial working memory tasks (Constantinidis and Wang, 2004; Curtis, 2006), such as parietal and prefrontal cortex (Chafee and Goldman-Rakic, 1998). However, it seems the strength of these correlations is often not on the order of the activity itself (di Pellegrino and Wise, 1993). For this reason, we presume the strength of interareal connections is weak $0 \leq \varepsilon^{1/2} \ll 1$. Note, we could choose to make them a different magnitude than the noise, but for analytical convenience, we choose interareal connection and noise magnitude to be roughly the same. Analysis could still be performed in other cases, but it would simply be more complicated. By setting $\tau = 1$, we can assume that time evolves on units of the excitatory synaptic time constant, which we presume to be roughly 10 ms (Häusser and Roth, 1997). The function $w_{jk}(x - y)$ describes the strength (amplitude of w_{jk}) and net polarity (sign of w_{jk}) of synaptic interactions from neurons with stimulus preference y to those with preference x. Following previous studies, we presume the modulation of the recurrent synaptic strength is given by the cosine

$$w_{jj}(x - y) = w(x - y) = \cos(x - y), \quad j = 1, 2, \tag{3}$$

so neurons with similar orientation preference excite one another and those with dissimilar orientation preference disynaptically inhibit one another (Ben-Yishai et al., 1995; Ferster and Miller, 2000). Lateral inhibitory type network architectures are supported by anatomical studies of the delay period neurons in prefrontal cortex (Goldman-Rakic, 1995). Our general analysis will apply to any even symmetric function of the distance $x - y$, but we typically compute things using (Equation 3) since it eases calculations. Finally, synaptic connections from area k to j are specified by the weight function $w_{jk}(x - y)$, and we typically take this to be the function

$$w_{jk}(x - y) = E_j + M_j \cos(x - y), \quad k \neq j \tag{4}$$

where E_j and M_j specify the strength of baseline excitation and modulation projecting to the jth area.

Output firing rates are given by taking the gain function $f(u)$ of the synaptic input, which we usually proscribe to be (Wilson and Cowan, 1973)

$$f(u) = \frac{1}{1 + e^{-\gamma(u - \theta)}},$$

and often take the high gain limit $\gamma \to \infty$ for analytical convenience, so (Amari, 1977)

$$f(u) = H(u - \theta) = \begin{cases} 0 : u < \theta, \\ 1 : u \geq \theta. \end{cases} \tag{5}$$

Effects of noise are described by the small amplitude ($0 \leq \varepsilon \ll 1$) stochastic processes $\varepsilon^{1/2} W_j(x, t)$ that are white in time and correlated in space so that $\langle \mathrm{d}W_j(x, t) \rangle = 0$ and

$$\langle \mathrm{d}W_j(x, t) \mathrm{d}W_j(y, s) \rangle = C_j(x - y) \delta(t - s) \mathrm{d}t \mathrm{d}s,$$

$$\langle \mathrm{d}W_j(x, t) \mathrm{d}W_k(y, s) \rangle = C_c(x - y) \delta(t - s) \mathrm{d}t \mathrm{d}s,$$

describing both local and shared noise in either area, $j = 1, 2$ with $j \neq k$. For simplicity, we assume the local spatial correlations have a cosine profile $C_j(x) = c_j \cos(x)$. We also typically assume the correlated noise component has cosine profile so $C_c(x) = c_c \cos(x)$. Therefore, in the limit $c_c \to 0$, there are no interareal noise correlations, and in the limit $c_c \to \min(c_1, c_2)$, noise in each area is maximally correlated. For instance, when $c_1 = c_2 = c_c = 1$, noise in each area is drawn from the same process.

MULTIPLE-AREA MODEL OF SPATIAL WORKING MEMORY

To incorporate the effects of many coupled, redundant areas encoding a spatial working memory, we consider a model with N areas and arbitrary synaptic architecture, given by

$$\tau \mathrm{d}u_j(x, t) = \left[-u_j + \varepsilon^{1/2} \sum_{k=1}^{N} w_{jk} * f(u_k) \right] \mathrm{d}t + \varepsilon^{1/2} \mathrm{d}W_j(x, t) \tag{6}$$

where u_j represents neural activity in the jth area where $j = 1, \ldots, N$. As before, we set $\tau = 1$, so each time unit corresponds to the roughly 10 ms timescale of excitatory synaptic conductance. The weight function $w_{jk}(x - y)$ represents the connection from neurons in area k with cue preference y to neurons in area j with cue preference x as described by (Equation 2). For comparison with numerical simulations, we take weight functions to be the cosines (Equation 3) and (Equation 4) and the firing rate function to be Heaviside (Equation 5). As in the dual area model, noises $W_j(x, t)$ are white in time and correlated in space so that $\langle \mathrm{d}W_j(x, t) \rangle = 0$ and

$$\langle \mathrm{d}W_j(x, t) \mathrm{d}W_k(y, s) \rangle = C_{jk}(x - y)\delta(t - s)\mathrm{d}t\mathrm{d}s,$$

with $j, k = 1, \ldots, N$, where local noise correlations are described when $j = k$ and noise correlations between areas are described when $j \neq k$. For comparison with numerical simulations, we consider $C_{jj}(x) = \cos(x)$ and $C_{jk}(x) = c_c \cos(x)$ for all $j \neq k$.

NUMERICAL SIMULATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

The spatially extended model (Equation 1) was simulated using an Euler–Maruyama method with a timestep 10^{-4}, using Riemann integration on the convolution term with 2000 spatial grid points. To compute and compare the variances $\langle \Delta_1(t)^2 \rangle$ for the dual and multiple area model, we simulated the system 5000 times. The position of the bump Δ_j at each timestep, in each simulation, was determined by the position x in each area j at which the maximal value of $u_j(x, t)$ was attained. The variance was then computed at each timepoint and compared to our asymptotic calculations.

RESULTS

We will now study how interareal architecture affect the dynamics of bumps in multiple area stochastic neural fields. To start, we demonstrate that in the absence of reciprocal connectivity between areas bump attractors exist that are neutrally stable to perturbations that change their position, which has long been known (Amari, 1977; Camperi and Wang, 1998; Ermentrout,

1998). Introducing weak interareal connectivity can decrease the variability in bump position because noise that moves bumps in the opposite direction is canceled due to an attractive force introduced by connectivity. Perturbations that push bumps in the same direction are still integrated, so bumps wander due to dynamic fluctuations, but their effective variance is smaller than it would be without interareal synaptic connections. In the presence of noise correlations between areas, effects of noise cancelation are weaker since stochastic forcing in each area is increasingly similar. Our asymptotic analysis is able to explain all of this with its resulting multivariate Ornstein–Uhlenbeck process.

BUMPS IN THE NOISE-FREE SYSTEM

To begin, we seek stationary solutions to Equation (1) in the absence interareal connections and noise ($\varepsilon \to 0$). Similar analyses have been carried out for bumps in single area populations (Ermentrout, 1998; Hansel and Sompolinsky, 1998). For this study, we assume recurrent connections are identical in all areas ($w_{jj} = w$). Relaxing this assumption slightly does not dramatically alter our results. Note first stationary solutions take the form $(u_1(x, t), u_2(x, t)) = (U_1(x), U_2(x))$. In the absence of any interareal connections, we would not necessarily expect the peaks of these bumps to be at the same location. However, translation invariance of the system (Equation 1) allows us to set the center of both bumps to be $x = 0$ to ease calculations. The stationary bump solutions then satisfy the system

$$U_1 = w * f(U_1), \qquad U_2 = w * f(U_2), \tag{7}$$

so the shape of each bump is only determined by the local connections w. For w given by Equation (3), since $U_1(x)$ and $U_2(x)$ are assumed to be peaked at $x = 0$, then by also assuming even symmetric solutions, we find

$$U_1(x) = \int_{-\pi}^{\pi} \cos y f(U_1(y))\mathrm{d}y \cos x,$$

$$U_2(x) = \int_{-\pi}^{\pi} \cos y f(U_2(y))\mathrm{d}y \cos x, \tag{8}$$

where we use $\cos(x - y) = \cos x \cos y + \sin x \sin y$. We can more easily compute the precise shape of these bumps in case of a Heaviside firing rate function (Equation 5). There is then an identical active region of each bump such that $U_1(x) > \theta$ and $U_2(x) > \theta$ when $x \in (-a, a)$, so the Equation (8) become $U_1(x) = U_2(x) = 2 \sin a \cos x$. Applying self-consistency, $U_1(\pm a) = U_2(\pm a) = \theta$, we can generate an implicit equation for the half-widths of the bumps a given by $2 \sin a \cos a = \sin(2a) = \theta$. Solving this explicitly for a, we find two solutions on $a \in [0, \pi]$: $a_u = \frac{1}{2} \sin^{-1} \theta$ and $a_s = \frac{\pi}{2} - \frac{1}{2} \sin^{-1} \theta$. Only the bump associated with a_s is stable.

The bumps (Equation 7) are neutrally stable to perturbations in both directions, which can lead to encoding error once the effects of dynamic fluctuations are considered (Kilpatrick et al., 2013). Since the two areas are uncoupled, examining bumps' stability can be reduced to studying each bump's stability individually (see Kilpatrick and Ermentrout, 2013 for details). Translating

a bump by a scaling of the spatial derivative $U'(x)$, we find $u_j(x, t) = U_j(x) + \varepsilon^{1/2} U_j'(x) e^{\lambda t}$ is associated with a zero eigenvalue ($\lambda = 0$), corresponding to neutral stability. To see this, we plug it into the corresponding bump equation of Equation (1) in the absence of noise and interareal connections and examine the linearization

$$\lambda U_j'(x) = -U_j'(x) + \int_{-\pi}^{\pi} w(x - y) f'(U_j(y)) U_j'(y) \mathrm{d}y. \quad (9)$$

Note, in the limit of infinite gain $\gamma \to \infty$, a sigmoid f becomes the Heaviside (Equation 5), and

$$f'(U(x)) = \frac{\mathrm{d}H(U(x))}{\mathrm{d}U} = \frac{\delta(x - a)}{|U'(a)|} + \frac{\delta(x + a)}{|U'(a)|},$$

in the sense of the distributions. Equation (9) still hold in this case. Differentiating (Equation 7), and integrating by parts, we find

$$\begin{aligned} -U_1' + w * [f'(U_1)U_1'] &= 0, \\ -U_2' + w * [f'(U_2)U_2'] &= 0, \end{aligned} \quad (10)$$

where the boundary terms vanish due to periodicity of the domain $[-\pi, \pi]$. Thus, the right hand side of Equation (9) vanishes, and $\lambda = 0$ is the only eigenvalue corresponding to translating perturbations. Thus, either bump (in area 1 or 2) is neutrally stable to perturbations that shifts its position in either direction (rightwards or leftwards), since the bump in each area experiences no force from the other bump.

This changes when we consider the effect of interareal connectivity. Once the two areas of Equation (1) are reciprocally coupled, bumps are stable to perturbations that translate them in opposite directions of one another (see **Figure 1**). Interareal connections act as a restoring force between the two positions of each bump. We will demonstrate this in the subsequent section by deriving a linear stochastic system for the position of either bump in the presence of small noise and weak interareal connectivity. The restorative nature of interareal connectivity is revealed by the negative eigenvalue associated with the interaction matrix (Equation 15) of our stochastic system, as shown in Equation (18).

NOISE-INDUCED WANDERING OF BUMPS

Now we consider the effects of small noise on the position of bumps in the presence of weak interareal connections. We start by presuming noise generates two distinct effects in the bumps (see **Figure 2**). First, noise causes both bumps to wander away from their initial positions, while still being pulled back into place by the bump in the other area. Bump position in areas 1 and 2 will be described by the time-varying stochastic variables $\Delta_1(t)$ and $\Delta_2(t)$. Second, noise causes fluctuations in the shape of both bumps, described by a correction Φ_j. To account for this, we consider the ansatz

$$\begin{aligned} u_1 &= U_1(x - \Delta_1(t)) + \varepsilon^{1/2} \Phi_1(x - \Delta_1(t), t) + \cdots \\ u_2 &= U_2(x - \Delta_2(t)) + \varepsilon^{1/2} \Phi_2(x - \Delta_2(t), t) + \cdots \end{aligned} \quad (11)$$
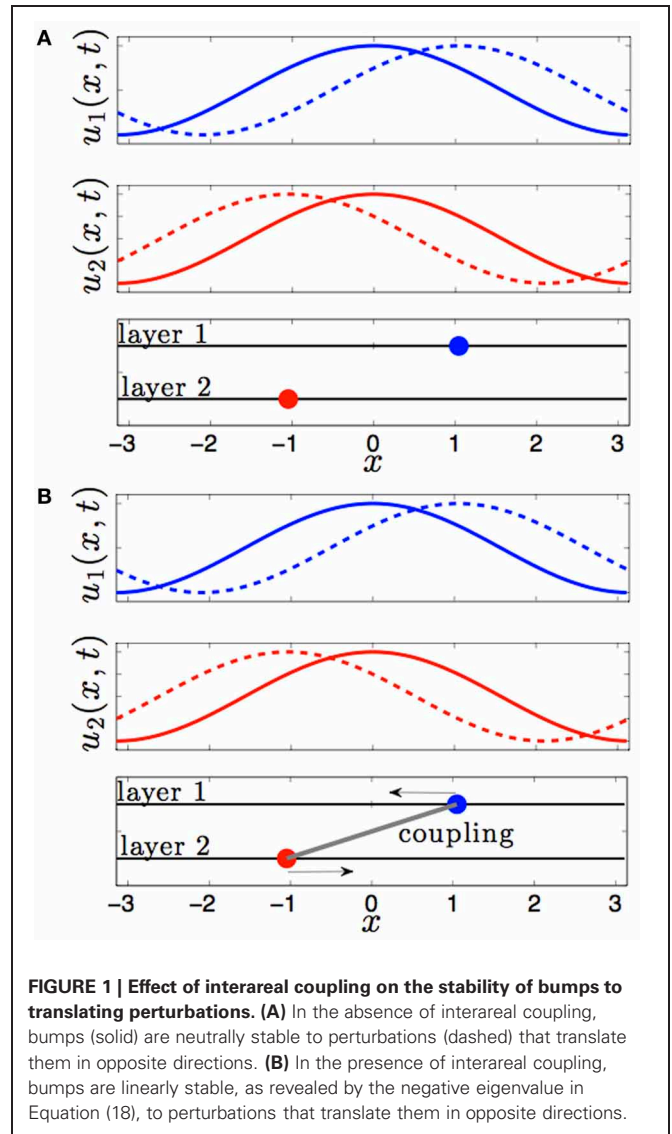

FIGURE 1 | Effect of interareal coupling on the stability of bumps to translating perturbations. (A) In the absence of interareal coupling, bumps (solid) are neutrally stable to perturbations (dashed) that translate them in opposite directions. **(B)** In the presence of interareal coupling, bumps are linearly stable, as revealed by the negative eigenvalue in Equation (18), to perturbations that translate them in opposite directions.

Armero et al. (1998) originally developed this approach to analyze of front propagation in stochastic PDE models. In stochastic neural fields, it has been modified to analyze wave propagation (Bressloff and Webber, 2012a) and bump wandering (Kilpatrick and Ermentrout, 2013). Plugging the ansatz (Equation 11) into the system (Equation 1) and expanding in powers of $\varepsilon^{1/2}$, we find that at $\mathcal{O}(1)$, we have the bump solution (Equation 7). Proceeding to $\mathcal{O}(\varepsilon^{1/2})$, we find

$$\mathrm{d}\boldsymbol{\Phi} - \mathcal{L}\boldsymbol{\Phi} = \begin{pmatrix} \varepsilon^{-1/2}\dot{\Delta}_1 U_1' + \mathrm{d}W_1 \\ \varepsilon^{-1/2}\dot{\Delta}_2 U_2' + \mathrm{d}W_2 \end{pmatrix} + \mathcal{K}(x, t), \quad (12)$$

where $\mathcal{K}(x, t)$ is the 2×1 vector function

$$\mathcal{K}(x, t) = \begin{pmatrix} w_{12} * [f(U_2) + f'(U_2)U_2' \cdot (\Delta_2 - \Delta_1)] \, \mathrm{d}t \\ w_{21} * [f(U_1) + f'(U_1)U_1' \cdot (\Delta_1 - \Delta_2)] \, \mathrm{d}t \end{pmatrix}$$
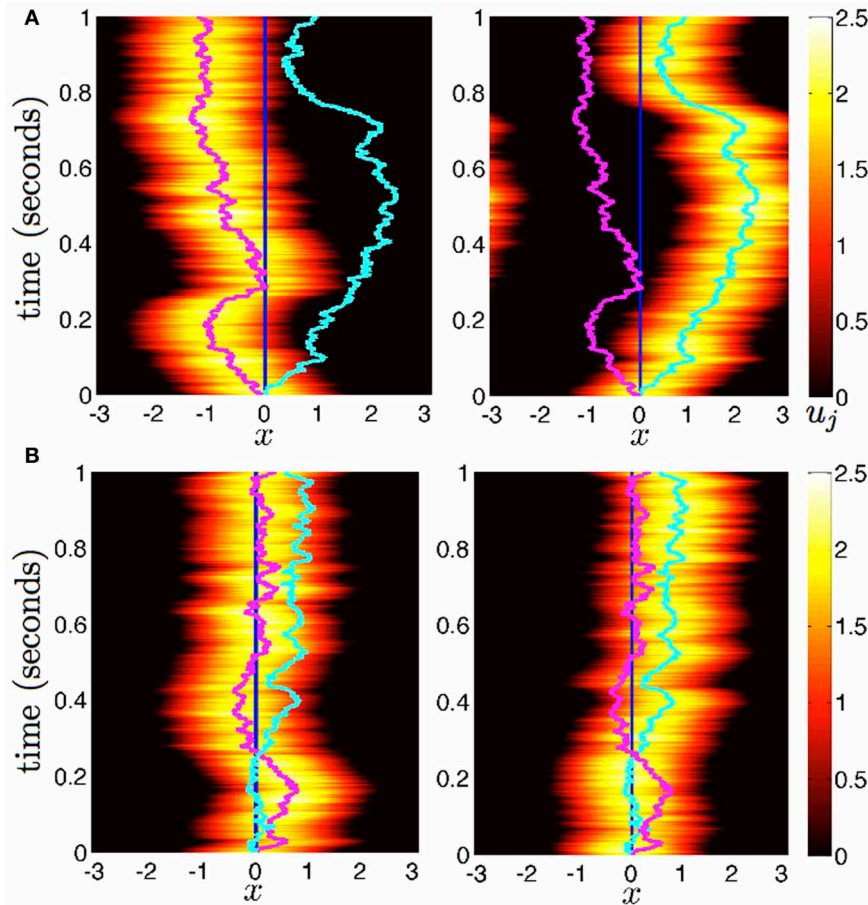
FIGURE 2 | Diffusion of bumps in the dual area stochastic neural field (Equation 1). (A) Without interareal connections ($w_{12} = w_{21} \equiv 0$), each bump executes Brownian motion about the domain, due to stochastic forces. **(B)** In the presence of interareal connections $\sqrt{\varepsilon}w_{12}(x) = \sqrt{\varepsilon}w_{21}(x) = 0.01(\cos(x) + 1)$, the position of bump 1 (magenta) is attracted to the position of bump 2 (cyan) and vice versa. Due to the reversion of each bump to the position of the other, both bumps effectively wander the domain less. Local connectivity is described by the cosine (Equation 3); the firing rate function is Equation (5). Other parameters are threshold $\theta = 0.5$ and noise amplitude $\varepsilon = 0.025$.

$\Phi = (\Phi_1(x, t), \Phi_2(x, t))^T$; and $\mathcal{L}$ is the linear operator

$$\mathcal{L}\mathbf{u} = \begin{pmatrix} -u(x) + w(x) * [f'(U_1(x))u(x)] \\ -v(x) + w(x) * [f'(U_2(x))v(x)] \end{pmatrix}$$

for any vector $\mathbf{u} = (u(x)\ v(x))^T$ of integrable functions. Note that the nullspace of $\mathcal{L}$ includes the vectors $(U_1', 0)^T$ and $(0, U_2')^T$, due to Equation (10). The last terms in the right hand side vector of Equation (12) arise due to interareal connections. We have linearized them under the assumption $|\Delta_1 - \Delta_2|$ remains small, so

$$f(U_j(x + \Delta_k - \Delta_j)) \approx f(U_j(x))$$
$$+ f'(U_j(x))U_j'(x) \cdot (\Delta_k - \Delta_j),$$

where $j = 1, 2$ and $k \neq j$. To make sure that a solution to Equation (12) exists, we require the right hand side is orthogonal to all elements of the null space of the adjoint $\mathcal{L}^*$, which is defined

$$\int_{-\pi}^{\pi} \mathbf{p}^T \mathcal{L}\mathbf{u}\,dx = \int_{-\pi}^{\pi} \mathbf{u}^T \mathcal{L}^*\mathbf{p}\,dx,$$

for any integrable vector $\mathbf{p} = (p(x)\ q(x))^T$. It then follows

$$\mathcal{L}^*\mathbf{p} = \begin{pmatrix} -p(x) + f'(U_1(x))[w(x) * p(x)] \\ -q(x) + f'(U_2(x))[w(x) * q(x)] \end{pmatrix}. \tag{13}$$

We can show that the nullspace of $\mathcal{L}^*$ contains the vector $\mathbf{f}_1 = (f'(U_1)U_1', 0)^T$ by plugging it into Equation (13) to yield

$$\mathcal{L}^*\mathbf{f}_1 = \begin{pmatrix} -f'(U_1)U_1' + f'(U_1)[w * [f'(U_1)U_1']] \\ 0 \end{pmatrix} = \mathbf{0}$$

where $\mathbf{0} = (0, 0)^T$ and we use Equation (10). We can also show the nullspace of $\mathcal{L}^*$ contains $\mathbf{f}_2 = (0, f'(U_2)U_2')^T$ in the same way. Thus, we can ensure Equation (12) has a solution by taking the inner product of both sides of Equation (12) with the two null vectors to yield

$$\langle f'(U_1)U_1', \varepsilon^{-1/2}\dot{\Delta}_1 U_1' + \mathrm{d}W_1$$
$$+ w_{12} * \left[f(U_2) + f'(U_2)U_2' \cdot (\Delta_2 - \Delta_1) \right] \mathrm{d}t \rangle = 0$$
$$\langle f'(U_2)U_2', \varepsilon^{-1/2}\dot{\Delta}_2 U_2' + \mathrm{d}W_2$$
$$+ w_{21} * \left[f(U_1) + f'(U_1)U_1' \cdot (\Delta_1 - \Delta_2) \right] \mathrm{d}t \rangle = 0,$$

where we define the inner product $\langle u, v \rangle = \int_{-\pi}^{\pi} u(x)v(x)\mathrm{d}x$. Therefore, the stochastic vector $\mathbf{\Delta}(t) = (\Delta_1(t), \Delta_2(t))^T$ obeys the multivariate Ornstein–Uhlenbeck process

$$\mathrm{d}\mathbf{\Delta}(t) = \mathbf{K}\mathbf{\Delta}(t)\mathrm{d}t + \mathrm{d}\mathbf{W}(t) \qquad (14)$$

where effects of interareal connections are described by the matrix

$$\mathbf{K} = \begin{pmatrix} -\kappa_1 & \kappa_1 \\ \kappa_2 & -\kappa_2 \end{pmatrix}, \qquad (15)$$

with

$$\kappa_1 = \frac{\langle f'(U_1)U_1', \varepsilon^{1/2}w_{12} * \left[f'(U_2)U_2' \right] \rangle}{\langle f'(U_1)U_1', U_1' \rangle},$$
$$\kappa_2 = \frac{\langle f'(U_2)U_2', \varepsilon^{1/2}w_{21} * \left[f'(U_1)U_1' \right] \rangle}{\langle f'(U_2)U_2', U_2' \rangle}, \qquad (16)$$

and $(w_{12} * f(U_2)) \cdot U_1'$ and $(w_{21} * f(U_1)) \cdot U_2'$ vanish upon integration since they are odd. Noise is described by the vector $\mathrm{d}\mathbf{W}(t) = (\mathrm{d}\mathcal{W}_1, \mathrm{d}\mathcal{W}_2)^T$ with

$$\mathrm{d}\mathcal{W}_1(t) = -\varepsilon^{1/2}\frac{\langle f'(U_1)U_1, \mathrm{d}W_1 \rangle}{\langle f'(U_1)U_1', U_1' \rangle},$$
$$\mathrm{d}\mathcal{W}_2(t) = -\varepsilon^{1/2}\frac{\langle f'(U_2)U_2, \mathrm{d}W_2 \rangle}{\langle f'(U_2)U_2', U_2' \rangle}.$$

The white noise term $\mathbf{W}$ has zero mean $\langle \mathbf{W}(t) \rangle = \mathbf{0}$ and variance described by pure diffusion so $\langle \mathbf{W}(t)\mathbf{W}^T(t) \rangle = \mathbf{D}t$ with

$$\mathbf{D} = \begin{pmatrix} D_1 & D_c \\ D_c & D_2 \end{pmatrix} \qquad (17)$$

where the associated diffusion coefficients of the variance are

$$D_1 = \varepsilon\frac{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F_1(x)F_1(y)C_1(x-y)\mathrm{d}x\mathrm{d}y}{\left[\int_{-\pi}^{\pi} F_1(x)U_1'(x)\mathrm{d}x \right]},$$
$$D_2 = \varepsilon\frac{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F_2(x)F_2(y)C_2(x-y)\mathrm{d}x\mathrm{d}y}{\left[\int_{-\pi}^{\pi} F_2(x)U_2'(x)\mathrm{d}x \right]}.$$

where $F_j(x) = f'(U_j(x))U_j'(x)$ and covariance is described by the coefficient

$$D_c = \varepsilon\frac{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F_1(x)f'(U_2(y))F_2(y)C_c(x-y)\mathrm{d}x\mathrm{d}y}{\left[\int_{-\pi}^{\pi} F_1(x)U_1'(x)\mathrm{d}x \right]\left[\int_{-\pi}^{\pi} F_2(x)U_2'(x)\mathrm{d}x \right]}.$$

In the next section, we analyze this stochastic system (Equation 14), showing how coupling between areas can reduce the variability of the bump positions $\Delta_1(t)$ and $\Delta_2(t)$.

EFFECT OF COUPLING ON BUMP POSITION VARIANCE

To analyze the Ornstein–Uhlenbeck process (Equation 14), we start by diagonalizing the matrix $\mathbf{K} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ using the eigenvalue decomposition

$$\mathbf{\Lambda} = \begin{pmatrix} 0 & 0 \\ 0 & -\kappa_1 - \kappa_2 \end{pmatrix},$$
$$\mathbf{V} = \frac{1}{\kappa_1 + \kappa_2}\begin{pmatrix} 1 & \kappa_1 \\ 1 & -\kappa_2 \end{pmatrix}, \qquad (18)$$
$$\mathbf{V}^{-1} = \begin{pmatrix} \kappa_2 & \kappa_1 \\ 1 & -1 \end{pmatrix},$$

such that $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues; columns of $\mathbf{V}$ are right eigenvectors; and rows of $\mathbf{V}^{-1}$ are left eigenvectors. Eigenvalues λ_1, λ_2 and eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ inform us of the effect of interareal coupling on linear stability. The eigenvalue $\lambda_1 = 0$ corresponds to the neutral stability of the positions $(\Delta_1, \Delta_2)^T$ to translations in the same direction $\mathbf{v}_1 = (1, 1)^T$. The negative eigenvalue $\lambda_2 = -(\kappa_1 + \kappa_2)$ corresponds to the linear stability introduced by interareal connections. The positions $(\Delta_1, \Delta_2)^T$ revert to one another when perturbations translate them in opposite directions $\mathbf{v}_2 = (\kappa_1, -\kappa_2)^T$.

Diagonalizing $\mathbf{K} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ using Equation (18), we can compute the mean and variance of the vector $\mathbf{\Delta}(t)$ given by Equation (14). First, note that the mean $\langle \mathbf{\Delta}(t) \rangle = e^{\mathbf{K}t}\mathbf{\Delta}(0)$ (Gardiner, 2003), which we can compute

$$\langle \mathbf{\Delta} \rangle = \begin{pmatrix} (\kappa_2 + \kappa_1 e^{\lambda_2 t})\Delta_1(0) + (\kappa_1 - \kappa_1 e^{\lambda_2 t})\Delta_2(0) \\ (\kappa_2 - \kappa_2 e^{\lambda_2 t})\Delta_1(0) + (\kappa_1 + \kappa_2 e^{\lambda_2 t})\Delta_2(0) \end{pmatrix}$$

using the diagonalization $e^{\mathbf{K}t} = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^{-1}$. Since $\lambda_2 = -(\kappa_1 + \kappa_2) < 0$,

$$\lim_{t \to \infty} \langle \mathbf{\Delta}(t) \rangle = [\kappa_2\Delta_1(0) + \kappa_1\Delta_2(0)]\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thus, the means of $\Delta_1(t)$ and $\Delta_2(t)$ always relax to the same position in long time, due to the linear stability introduced by connections between areas. Under the assumption they both begin at $\Delta_1(0) = \Delta_2(0) = 0$, the covariance matrix is given (Gardiner, 2003)

$$\langle \mathbf{\Delta}(t)\mathbf{\Delta}^T(t) \rangle = \int_0^t e^{\mathbf{K}(t-s)}\mathbf{D}e^{\mathbf{K}^T(t-s)}\mathrm{d}s, \qquad (19)$$

where $\mathbf{D}$ is the covariance coefficient matrix of the white noise vector $\mathbf{W}(t)$ given by Equation (17). To compute Equation (19), we additionally need the diagonalization $\mathbf{K}^T = (\mathbf{V}^{-1})^T \Lambda \mathbf{V}^T$, so $e^{\mathbf{K}^T t} = (\mathbf{V}^{-1})^T e^{\Lambda t} \mathbf{V}^T$. After multiplying and integrating (Equation 19), we find the elements of the covariance matrix

$$\langle \mathbf{\Delta}(t) \mathbf{\Delta}^T(t) \rangle = \begin{pmatrix} \langle \Delta_1(t)^2 \rangle & \langle \Delta_1(t)\Delta_2(t) \rangle \\ \langle \Delta_1(t)\Delta_2(t) \rangle & \langle \Delta_2(t)^2 \rangle \end{pmatrix}$$

are

$$\langle \Delta_1(t)^2 \rangle = D_+ t + 2\kappa_1 r_1(t) + \frac{\kappa_1}{\kappa_2} r_2(t) \qquad (20)$$

$$\langle \Delta_2(t)^2 \rangle = D_+ t - 2\kappa_2 r_1(t) + \frac{\kappa_2}{\kappa_1} r_2(t) \qquad (21)$$

$$\langle \Delta_1(t)\Delta_2(t) \rangle = D_+ t + (\kappa_1 - \kappa_2) r_1(t) - r_2(t)$$

where the effective diffusion coefficients are

$$D_+ = \frac{\kappa_2^2 D_1 + 2\kappa_1 \kappa_2 D_c + \kappa_1^2 D_2}{(\kappa_1 + \kappa_2)^2}, \qquad (22)$$

$$D_r = \frac{\kappa_2 D_1 - \kappa_1 D_2 + (\kappa_1 - \kappa_2) D_c}{(\kappa_1 + \kappa_2)^2}, \qquad (23)$$

$$D_- = \frac{D_1 - 2D_c + D_2}{(\kappa_1 + \kappa_2)^2}, \qquad (24)$$

so that D_+ and D_- are variances of noises occurring along the eigendirections $\mathbf{v}_1$ and $\mathbf{v}_2$. The functions $r_1(t)$, $r_2(t)$ are exponentially saturating

$$r_1(t) = \frac{D_r}{\kappa_1 + \kappa_2} \left[1 - e^{-(\kappa_1 + \kappa_2)t} \right],$$

$$r_2(t) = \frac{\kappa_1 \kappa_2 D_-}{2(\kappa_1 + \kappa_2)} \left[1 - e^{-2(\kappa_1 + \kappa_2)t} \right].$$

The main quantities of interest to us are the variances (Equation 20) and (Equation 21) with which we can make a few observations concerning the effect of interareal connections on the variance of bump positions.

First, note the long term variance of either bump's position $\Delta_1(t)$ and $\Delta_2(t)$ will be the same, described by the averaged diffusion coefficient D_+, since

$$\lim_{t \to \infty} \langle \Delta_1(t)^2 \rangle = \lim_{t \to \infty} \langle \Delta_2(t)^2 \rangle = D_+ t. \qquad (25)$$

As the effective coupling strengths κ_j are increased, we can expect the variances $\langle \Delta_j(t)^2 \rangle$ approach these limits at faster rates since other portions of the variance decay at a rate proportional to $|\lambda_2| = \kappa_1 + \kappa_2$.

Next, we study the case, across all times t, where connections between areas are the same ($w_{12} \equiv w_{21} = w_r$) and noise

within areas is identical ($D_1 \equiv D_2 = D_l$), the mean reversion rates will be the same ($\kappa_1 = \kappa_2 = \kappa$) and terms in Equation (23) cancel so $D_r = 0$. Thus, the variances will be identical ($\langle \Delta_1(t)^2 \rangle = \langle \Delta_2(t)^2 \rangle = \langle \Delta(t)^2 \rangle$) and

$$\langle \Delta(t)^2 \rangle = \frac{D_l + D_c}{2} t + \frac{D_l - D_c}{8\kappa} \left[1 - e^{-4\kappa t} \right].$$

This demonstrates the way in which correlated noise (D_c) contributes to the variance. When noise within each area is shared ($D_c \to D_l$), there is no benefit to interareal coupling and $\langle \Delta(t)^2 \rangle = D_l t$ (see Kilpatrick and Ermentrout, 2013). However, when any noise is not shared between areas ($D_c < D_l$), variance can be reduced by increasing coupling strength κ between areas. The variance $\langle \Delta(t)^2 \rangle$ is monotone decreasing in κ since

$$\frac{\partial}{\partial \kappa} \langle \Delta(t)^2 \rangle = \frac{D_l - D_c}{8} \frac{(1 + 4\kappa t)e^{-4\kappa t} - 1}{\kappa^2} \le 0.$$

Inequality holds because $(1 + 4\kappa t) \le e^{4\kappa t}$ is ensured by the Taylor series expansion of $e^{4\kappa t}$ when $\kappa t > 0$.

Thus, variance is minimized in the limit

$$\lim_{\kappa \to \infty} \langle \Delta(t)^2 \rangle = \frac{D_l + D_c}{2} t. \qquad (26)$$

Therefore, strengthening interareal connections in *both* directions reduces the variance in bump position. On the other hand, in the limit of no interareal connections, we find $\lim_{\kappa \to 0} \langle \Delta(t)^2 \rangle = D_l t$, and the variance in a bump's position is determined entirely by local sources of noise.

Returning to asymmetric connectivity ($\kappa_1 \ne \kappa_2$), we consider the case of feedforward connectivity from area 1 to 2 ($w_{12} \equiv 0$), $\kappa_1 = 0$, so $D_+ = D_1$ and the formulas for the variances reduce to

$$\langle \Delta_1(t)^2 \rangle = D_1 t,$$

$$\langle \Delta_2(t)^2 \rangle = D_1 t + \frac{2(D_1 - D_c)}{\kappa_2} \left[1 - e^{-\kappa_2 t} \right]$$

$$+ \frac{D_1 - 2D_c + D_2}{2\kappa_2} \left[1 - e^{-2\kappa_2 t} \right],$$

so the pure diffusive term of both variances is wholly determined by the local noise of area 1. Then, only the position of the bump in area 2 possesses additional mean-reverting fluctuations in its position, which arise from local sources of noise that force it away from the position of the bump in area 1. In this situation, the variance of the bump in area 2's position is minimized when

$$\lim_{\kappa_2 \to \infty} \langle \Delta_1(t)^2 \rangle = \lim_{\kappa_2 \to \infty} \langle \Delta_2(t)^2 \rangle = D_1 t.$$

Comparing this with Equation (26) we see that, since $D_c \le D_1$, the variances $\langle \Delta_j(t)^2 \rangle$ will always be higher in this case than in

the case of very strong reciprocal coupling between both areas. Averaging information and noise between both areas decreases positional variance as opposed to one area simply receiving noise and information from another. Similar results have been recently identified in the context of studying synchrony of reciprocally coupled noisy oscillators (Ly and Ermentrout, 2010).

One important caveat is that if area 1 has more noise than area 2, the weighting of reciprocal connectivity, κ_1 and κ_2, should be balanced to minimize the variance. If the average diffusion coefficient D_+ is weighted too heavily with the area having the larger variance, the area with less intrinsic noise can end up noisier than it would be without reciprocal connectivity. To see this in the extreme case feedforward coupling, note that if $D_2 < D_1$, then $D_2 t < D_1 t < \langle \Delta_2(t)^2 \rangle$. Thus, the variance of $\Delta_2(t)$ increases as opposed to the uncoupled case where $\langle \Delta_2(t)^2 \rangle = D_2 t$.

We now derive the optimal weighting of κ_1 and κ_2 to minimize the long term variance (Equation 25) for general asymmetric connectivity, in the absence of correlated noise $D_c = 0$. To do so, we fix κ_2 and find the κ_1 that minimizes D_+, which happens to be

$$\kappa_1 = \kappa_2 \frac{D_1}{D_2}.$$

Thus, for identical noise $D_1 = D_2$, setting $\kappa_1 = \kappa_2$ minimizes D_+. For much stronger noise in area 2 ($D_2 \gg D_1$), κ_1 should be made relatively small. In the case of noise correlations between areas ($D_c > 0$), the optimal value of κ_1 that minimizes (Equation 25) is

$$\kappa_1 = \kappa_2 \frac{D_1 - D_c}{D_2 - D_c}.$$

CALCULATING THE STOCHASTIC MOTION OF BUMPS

We now compute the effective variances (Equation 20) and (Equation 21), considering the specific case of Heaviside firing rate functions (Equation 5), cosine synaptic weights (Equation 3) and (Equation 4). Doing so, we can compare our asymptotic results to those computed from numerical simulations. We compute the mean reversion terms κ_1 and κ_2 by noting the spatial derivative of each bump will be $U_1'(x) = U_2'(x) = -2\sin a \sin x$ and the null vector components are

$$f'(U_j(x))U_j'(x) = \delta(x + a) - \delta(x - a).$$

for $j = 1, 2$. Plugging these formulae into Equation (16), we find $\kappa_1 = \varepsilon^{1/2} M_1$ and $\kappa_2 = \varepsilon^{1/2} M_2$.

We first consider the case of uncorrelated noise between areas, so $c_c \equiv 0$, meaning $D_c = 0$. We can compute the diffusion coefficients associated with the local noise in each area assuming cosine spatial correlations

$$D_1 = \frac{c_1 \varepsilon}{2 + 2\sqrt{1 - \theta^2}}, \quad D_2 = \frac{c_2 \varepsilon}{2 + 2\sqrt{1 - \theta^2}}. \quad (27)$$

We can then compute Equations (20) and (21) directly, for the case of no noise correlations between areas, by plugging in Equation (27).

For symmetric connections between areas, $\kappa = \varepsilon^{1/2} M_1 = \varepsilon^{1/2} M_2$, as well as identical noise, $c_1 = c_2 = 1$, we have $\langle \Delta_1(t)^2 \rangle = \langle \Delta_2(t)^2 \rangle = \langle \Delta(t)^2 \rangle$ and
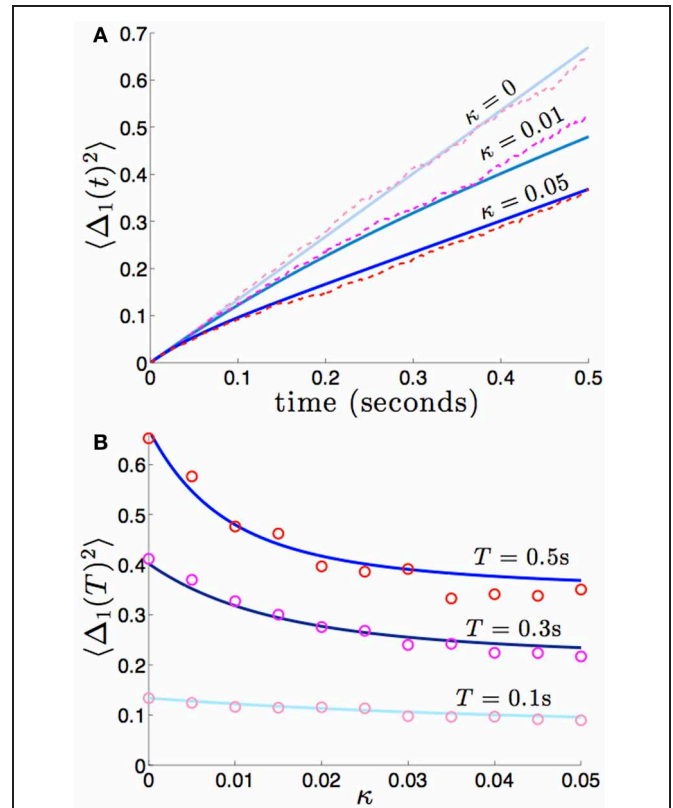
FIGURE 3 | Variance in the position of bumps as computed numerically (red shades) and from theory (blue shades) using Equation (28). Coupling between areas is symmetric $\sqrt{\varepsilon} w_{12}(x) = \sqrt{\varepsilon} w_{21}(x) = \kappa(\cos(x) + 1)$, so $\langle \Delta_1(t)^2 \rangle = \langle \Delta_2(t)^2 \rangle$, and there is no shared noise ($c_c = 0$). **(A)** The increase in variance is slower for stronger amplitudes of interareal coupling κ. Notice variance climbs sublinearly for $\kappa > 0$, due to the mean-reversion caused by coupling. **(B)** Variance drops considerably more over low values of κ that over high values. Other constituent functions and parameters are the same as in **Figure 2**.

$$\langle \Delta(t)^2 \rangle = \frac{\varepsilon t}{4(1 + \sqrt{1 - \theta^2})} + \frac{\varepsilon}{16(1 + \sqrt{1 - \theta^2})\kappa} \left[1 - e^{-4\kappa t} \right]. \quad (28)$$

We compare the formula (28) to results we obtain from numerical simulations in **Figure 3**, finding our asymptotic formula (28) matches quite well. In addition, we compare our results for general (possibly asymmetric) reciprocal connectivity to results from numerical simulations in **Figure 4**. We also show in **Figure 5**, as predicted, when κ_2 is held fixed, there is a finite optimal value of κ_1 that minimizes variance $\langle \Delta_1(t)^2 \rangle$. Therefore, reciprocal connectivity in multi-area networks should be balanced, in order to minimize positional variance of the stored bump.

Next, we consider the case of correlated noise between areas, so $c_c > 0$, meaning $D_c > 0$. In this case, the covariance terms in D_+ and D_- are non-zero. We can thus compute the diffusion coefficient associated with correlated noise

$$D_c = \frac{c_c \varepsilon}{2 + 2\sqrt{1 - \theta^2}}.$$
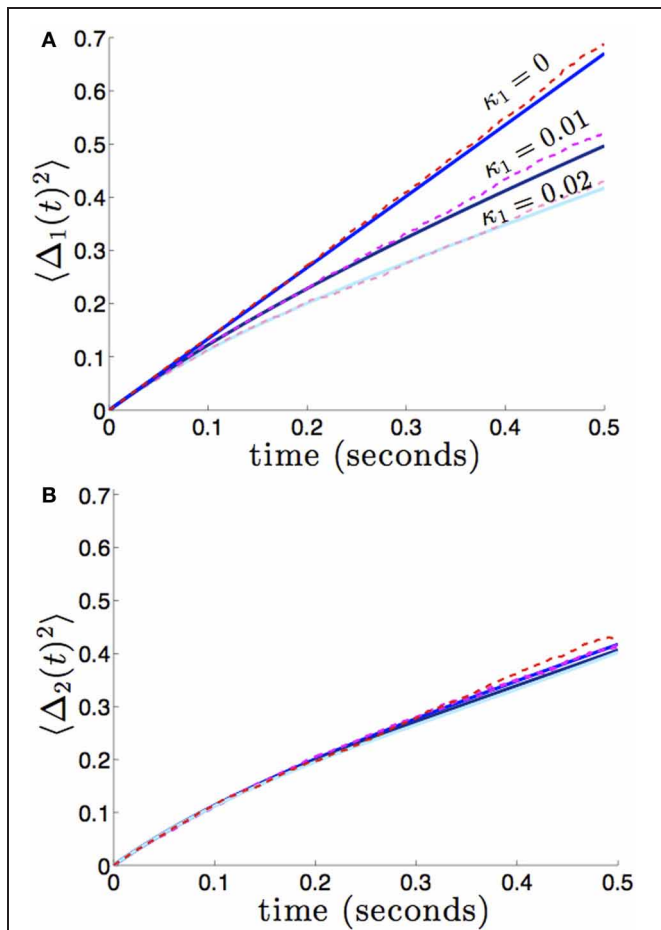
FIGURE 4 | Variance in the position of bumps as it depends on asymmetric reciprocal connectivity ($\kappa_1 \neq \kappa_2$) when noise in each area is independent and identical ($c_1 = c_2 = 1$). Fixing $\kappa_2 = 0.02$ and varying κ_1, we find **(A)** the variance $\langle \Delta_1(t)^2 \rangle$ of bump 1 decreases as coupling from area 2 to 1 (κ_1) increases; **(B)** variance $\langle \Delta_2(t)^2 \rangle$ of bump 2 remain relatively unchanged. Other constituent functions and parameters are the same as in **Figure 2**.
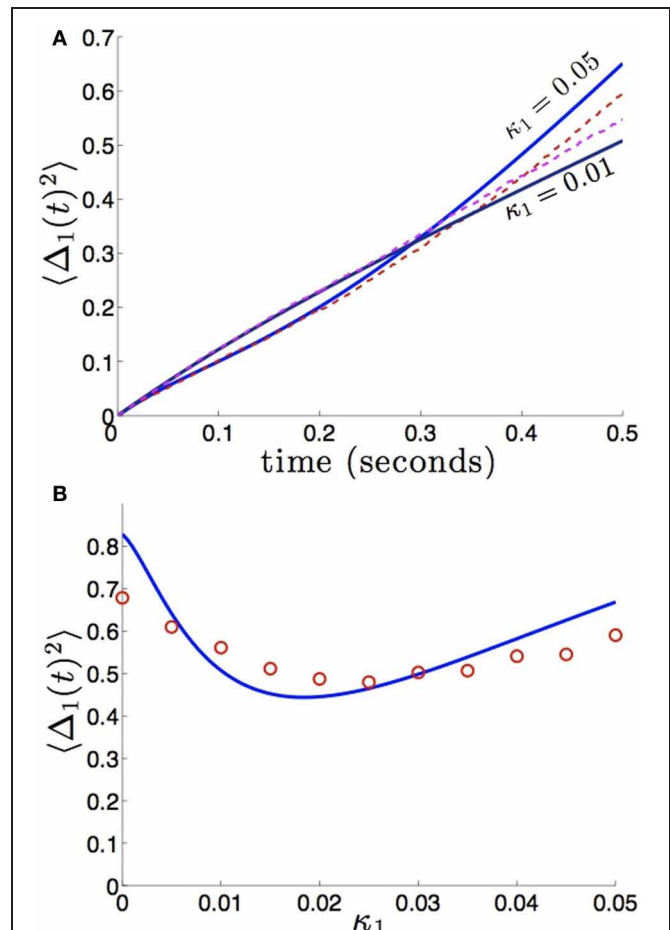
FIGURE 5 | Bump position variance depends non-monotonically on asymmetric connectivity strength. (A) For $\kappa_2 = 0.01$ and high enough values of coupling ($\kappa_1 = 0.05$), variance $\langle \Delta_1(t)^2 \rangle$ scales more quickly than for symmetric coupling ($\kappa_1 = 0.01$). Layer 1 is being sourced by the noisier area 2. **(B)** Non-monotonic dependence of variance $\langle \Delta_1(t)^2 \rangle$ on projection strength from area 2 to area 1 κ_1 is shown for fixed time $T = 50$ and $\kappa_2 = 0.01$ fixed. Amplitude of noise in area 2 is twice that of area 1 ($c_1 = 1$ and $c_2 = 2$). Other constituent functions and parameters are the same as in **Figure 2**.

In the case of symmetric connections between areas, $\kappa = \varepsilon^{1/2} M_1 = \varepsilon^{1/2} M_2$, and identical internal noise, $c_1 = c_2 = 1$, we have $\langle \Delta_1(t)^2 \rangle = \langle \Delta_2(t)^2 \rangle = \langle \Delta(t)^2 \rangle$ and

$$\langle \Delta(t)^2 \rangle = \frac{(1 + c_c)\varepsilon}{4\left(1 + \sqrt{1 - \theta^2}\right)}t + \frac{(1 - c_c)\varepsilon}{16\left(1 + \sqrt{1 - \theta^2}\right)\kappa}\left[1 - e^{-4\kappa t}\right], \tag{29}$$

which reflects the fact that interareal connections do not reduce variability as much when there are strong noise correlations c_c between areas. We demonstrate the accuracy of the theoretical calculation (Equation 29) as compared to numerical simulations in **Figure 6**. Numerical simulations also reveal the fact that stronger noise correlations between areas diminish the effectiveness of interareal connections at reducing bump position variance.

REDUCTION OF BUMP WANDERING IN MULTIPLE AREAS

We now examine the effect of interareal connections in networks with more than two areas using the system (Equation 6). As with the dual area network without noise or interareal connectivity, stationary bump solutions take the form $(u_1, \ldots, u_N) = (U_1(x), \ldots, U_N(x))$, and translation invariance let us to set all bump peaks to be located at $x = 0$ so

$$U_j = w * f(U_j), \quad j = 1, \ldots, N. \tag{30}$$

As before, we presume $w_{jj} = w$, and relaxing this assumption does not dramatically alter our results. Linear stability analysis of bumps proceeds along similar lines to the dual area network, so we omit those calculations and summarize the results. In the absence of interareal connections, each bump is neutrally stable to

perturbation in either direction. In the presence of interareal connections, all bumps are only neutrally stable to translations that move them all in the same direction. Therefore, networks with more areas provide more perturbation cancelations.

To study how noise and interareal connections affect the trajectory of bump positions, we again note noise causes all bumps to wander away from their initial position, while being pulled back into place by projections from other areas (see **Figure 7**). The position of the bump in area j is described by the stochastic variable Δ_j. Noise also causes fluctuations in the shape

FIGURE 6 | Variance in the position of bumps as noise correlation between areas is increased. Numerically computed variance (red shades) match theoretical curves from Equation (29) (blue shades), very well. Reciprocal connectivity reduces variability the most when there is no correlated noise ($c_c = 0$) between areas. As the shared noise between areas increased is amplitude ($c_c = 0.5, 1$), the advantage of reciprocal connectivity is diminished. When $c_c = 1$ changing κ does not affect the variance $\langle \Delta(t)^2 \rangle$ (see formula (29) in the limit $c_c \to 1$). Other constituent functions and parameters are the same as in **Figure 2**.

of both bumps, which is described by the correction term Φ_j. Therefore, we presume the resulting state of the system satisfies the ansatz

$$u_j = U_j(x - \Delta_j(t)) + \varepsilon^{1/2} \Phi_j(x - \Delta_j(t), t) + \cdots,$$

where $j = 1, \ldots, N$. Plugging this ansatz into Equation (6) and expanding in powers of $\varepsilon^{1/2}$, we find that at $\mathcal{O}(1)$, we simply have the system of Equation (30) for the bump solutions. Proceeding to $\mathcal{O}(\varepsilon^{1/2})$, we find

$$d\Phi - \mathcal{L}\Phi = \mathcal{K}(x, t) + \begin{pmatrix} \varepsilon^{-1/2} \dot{\Delta}_1 U_1' + dW_1 \\ \vdots \\ \varepsilon^{-1/2} \dot{\Delta}_j U_j' + dW_j \\ \vdots \\ \varepsilon^{-1/2} \dot{\Delta}_N U_N' + dW_N \end{pmatrix} \quad (31)$$
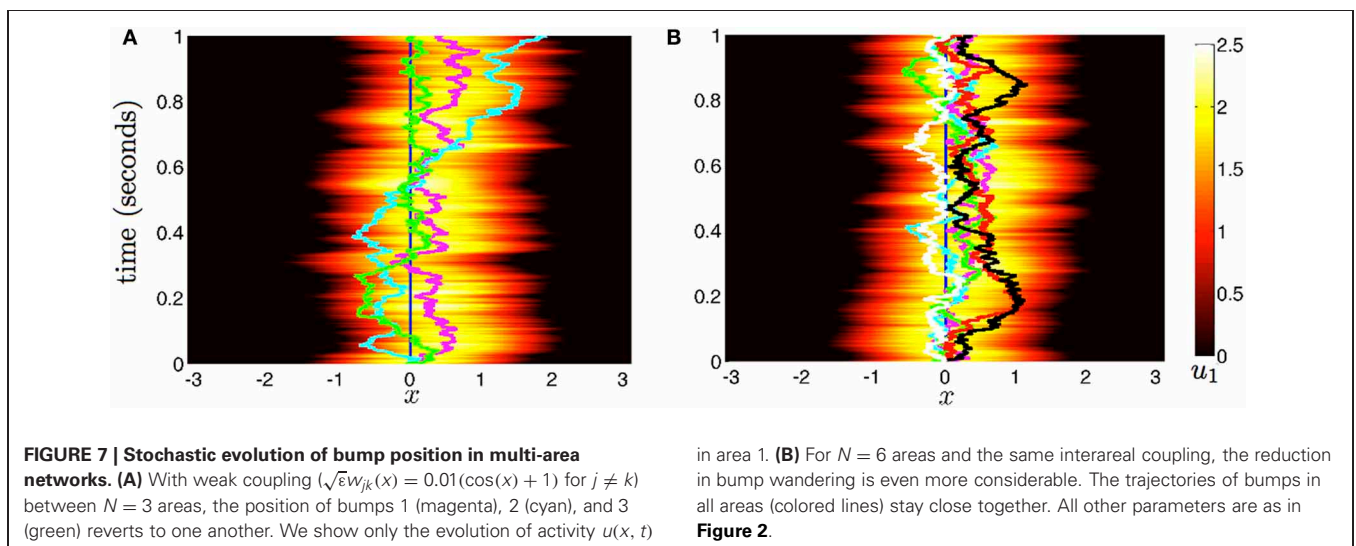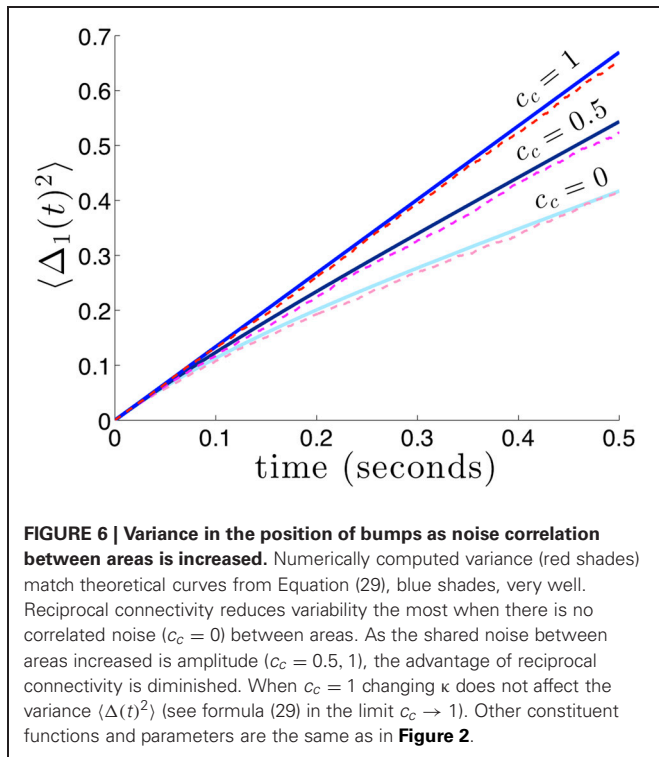
where $\mathcal{K}(x, t)$ is an $N \times 1$ vector whose jth entry is

$$\mathcal{K}_j = \sum_{k \neq j} w_{jk} * \left[f(U_k) + f'(U_k) U_k' \cdot (\Delta_k - \Delta_j) \right] dt;$$

$\Phi = (\Phi_1(x, t), \cdots, \Phi_N(x, t))^T$; and $\mathcal{L}$ is the linear operator

$$\mathcal{L}\Psi = \begin{pmatrix} -\Psi_1(x) + w * [f'(U_1(x))\Psi_1(x)] \\ \vdots \\ -\Psi_N(x) + w * [f'(U_N)\Psi_N(x)] \end{pmatrix}$$

for any integrable vector $\Psi = (\Psi_1(x), \ldots, \Psi_N(x))^T$. The nullspace of $\mathcal{L}$ is spanned by the vectors $(U_1', 0, \ldots, 0)^T$; $(0, U_2', 0, \ldots, 0)^T$; $\ldots$; and $(0, \ldots, 0, U_N')^T$, which can be seen

FIGURE 7 | Stochastic evolution of bump position in multi-area networks. (A) With weak coupling ($\sqrt{\varepsilon} w_{jk}(x) = 0.01(\cos(x) + 1)$ for $j \neq k$) between $N = 3$ areas, the position of bumps 1 (magenta), 2 (cyan), and 3 (green) reverts to one another. We show only the evolution of activity $u(x, t)$ in area 1. **(B)** For $N = 6$ areas and the same interareal coupling, the reduction in bump wandering is even more considerable. The trajectories of bumps in all areas (colored lines) stay close together. All other parameters are as in **Figure 2**.

by differentiating (Equation 30). The last terms on the right hand side of Equation (31) arise due to interareal connections. We have linearized them under the assumption that $|\Delta_k - \Delta_j|$ remains small for all j, k. To ensure a solution to Equation (31), we require the right hand side is orthogonal to all elements of the null space of the adjoint operator $\mathcal{L}^*$. The adjoint is defined with respect to the inner product

$$\int_{-\pi}^{\pi} \mathbf{\Upsilon}^T \mathcal{L} \mathbf{\Psi} \, dx = \int_{-\pi}^{\pi} \mathbf{\Psi}^T \mathcal{L}^* \mathbf{\Upsilon} \, dx$$

where $\mathbf{\Upsilon} = (\Upsilon_1(x), \ldots, \Upsilon_N(x))^T$ is integrable. It then follows

$$\mathcal{L}^* \mathbf{\Upsilon} = \begin{pmatrix} -\Upsilon_1(x) + f'(U_1(x))[w * \Upsilon_1] \\ \vdots \\ -\Upsilon_N(x) + f'(U_N)[w * \Upsilon_N] \end{pmatrix}.$$

The nullspace of $\mathcal{L}^*$ contains the vectors $(f'(U_1)U_1', 0, \ldots, 0)^T$; $(0, f'(U_2)U_2', 0, \ldots 0)^T$; $\ldots$; and $(0, \ldots, 0, f'(U_N)U_N')$, which can be shown by applying $\mathcal{L}^*$ to them and using the formula generated by differentiating (Equation 30). Thus, to be sure (Equation 31) has a solution, we take the inner product of both sides of the equation with all N null vectors and isolate $d\Delta_j$ terms to yield the multivariate Ornstein–Uhlenbeck process

$$d\mathbf{\Delta}(t) = \mathbf{K}\mathbf{\Delta}(t)dt + d\mathbf{W}(t), \qquad (32)$$

where effects of interareal connections are described by the matrix $\mathbf{K} \in \mathbb{R}^{N \times N}$ where the diagonal and off-diagonal entries are given

$$\mathbf{K}_{jj} = -\sum_{k \neq j} \kappa_{jk}, \qquad \mathbf{K}_{jk} = \kappa_{jk}$$

for $j = 1, \ldots, N$ and $k \neq j$, where

$$\kappa_{jk} = \frac{\langle f'(U_j)U_j', \varepsilon^{1/2} w_{jk} * [f'(U_k)U_k'] \rangle}{\langle f'(U_j)U_j', U_j' \rangle},$$

and we have used the fact that $w_{jk} * f(U_k) \cdot U_j'$ is an odd function for all j, k, so they vanish on integration. Stochastic forces are described by the vector

$$d\mathbf{W}(t) = \begin{pmatrix} d\mathcal{W}_1(t) \\ \vdots \\ d\mathcal{W}_N(t) \end{pmatrix},$$

$$d\mathcal{W}_j(t) = -\varepsilon^{1/2} \frac{\langle f'(U_j)U_j', dW_j \rangle}{\langle f'(U_j)U_j', U_j' \rangle}.$$

The white noise vector $\mathbf{W}(t)$ has zero mean $\langle \mathbf{W}(t) \rangle = \mathbf{0}$, and covariance matrix $\langle \mathbf{W}(t)\mathbf{W}^T(t) \rangle = \mathbf{D}t$ where associated coefficients of the matrix $\mathbf{D}$ are

$$D_{jj} = \varepsilon \frac{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F_j(x) F_j(y) C_j(x - y) dx dy}{\left[\int_{-\pi}^{\pi} F_j(x) U_j'(x) dx \right]^2}.$$

where $F_j(x) = f'(U_j(x))U_j'(x)$, which describe the variance within an area and

$$D_{jk} = \varepsilon \frac{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F_j(x) F_k(y) C_{jk}(x - y) dx dy}{\left[\int_{-\pi}^{\pi} F_j(x) U_j'(x) dx \right] \left[\int_{-\pi}^{\pi} F_k(x) U_k'(x) dx \right]},$$

which describes covariance between areas. Since correlations are symmetric $C_{jk}(x) = C_{kj}(x)$ for all j, k, then $D_{jk} = D_{kj}$ for all j, k.

A detailed analysis of the linear stochastic system (Equation 32) is difficult without some knowledge of the entries κ_{jk}. However, we can make a few general statements. We note that all eigenvalues of $\mathbf{K}$ must have negative real part or be zero, due to the Gerschgorin circle theorem (Feingold and Varga, 1962), which states that all eigenvalues a matrix $\mathbf{K}$ must lie in one of the disks with center K_{jj} and radius $\sum_{k \neq j} |K_{jk}|$. Since $K_{jj} = -\sum_{k \neq j} \kappa_{jk}$ and $K_{jk} = \kappa_{jk}$, then

$$K_{jj} + \sum_{k \neq j} K_{jk} = -\sum_{k \neq j} \kappa_{jk} + \sum_{k \neq j} |\kappa_{jk}| = 0 \qquad (33)$$

is the maximal possible eigenvalue, since $\kappa_{jk} \geq 0$ for all j, k. Therefore, we expect N eigenpairs $\lambda_j, \mathbf{v}_k$ associated with $\mathbf{K}$, where $\lambda_N \leq \lambda_{N-1} \leq \cdots \leq \lambda_2 \leq \lambda_1 = 0$. This means we can perform the diagonalization $\mathbf{K} = \mathbf{V} \Lambda \mathbf{V}^{-1}$, where Λ is the diagonal matrix of eigenvalues; columns of $\mathbf{V}$ are right eigenvectors; and rows of $\mathbf{V}^{-1}$ are left eigenvectors. Therefore, we can decompose the stochastic solution to Equation (32), when $\mathbf{\Delta}(0) = \mathbf{0}$ as

$$\mathbf{\Delta}(t) = \int_0^t e^{\mathbf{K}(t-s)} d\mathbf{W}(s) = \int_0^t \mathbf{V} e^{\Lambda(t-s)} \mathbf{V}^{-1} d\mathbf{W}(s),$$

Thus, as we expect, any stochastic fluctuations in Equation (32) will be integrated or decay over time due to the exponential filters $e^{\lambda_j(t-s)}$. In addition, when $\mathbf{\Delta}(0) = \mathbf{0}$ the covariance matrix can be computed as

$$\langle \mathbf{\Delta}(t)\mathbf{\Delta}^T(t) \rangle = \int_0^t e^{\mathbf{K}(t-s)} \mathbf{D} e^{\mathbf{K}^T(t-s)} ds, \qquad (34)$$

where $\mathbf{D}$ is the matrix of diffusion coefficients for the covariance $\langle \mathbf{W}(t)\mathbf{W}^T(t) \rangle$. We now compute the covariance in the specific case of symmetric connectivity.

In the case of symmetric connectivity between areas, $w_{jk} = w_r$ for all $j \neq k$, so $\kappa_{jk} = \kappa$ for all $j \neq k$. Effects of connectivity

between areas are described by the symmetric matrix

$$\mathbf{K} = \kappa J_N - N\kappa I$$

where J_N is the $N \times N$ matrix of ones and I is the identity. The eigenvalues of J_N are N, with multiplicity one, and zero, with multiplicity $N - 1$. Thus, the largest eigenvalue of $\mathbf{K} = \kappa J_N - N\kappa I$ is $\lambda_1 = 0$ with associated eigenvector $\mathbf{v}_1 = (1, \ldots, 1)^T$. All other eigenvalues are $\lambda_j = -N\kappa$ for $j \geq 2$, with associated eigenvectors $\mathbf{v}_j = \mathbf{e}_1 - \mathbf{e}_j$, where $j = 2, \ldots, N$ and $\mathbf{e}_j$ is the unit vector with a one in the jth row and zeros elsewhere. Our diagonalization of the symmetric matrix $\mathbf{K} = \mathbf{K}^T = \mathbf{V}\Lambda\mathbf{V}^{-1}$ then involves the diagonal matrix Λ of eigenvalues λ_j; the symmetric matrix $\mathbf{V}$ whose columns $\mathbf{v}_j$ are right eigenvectors; and the symmetric matrix $\mathbf{V}^{-1}$ whose rows are left eigenvectors. The matrix $\mathbf{V}^{-1}$ takes the form

$$\mathbf{V}^{-1} = \frac{1}{N} \begin{pmatrix} 1 & 1 & \cdots & \\ 1 & -(N-1) & 1 & \cdots \\ & & \ddots & 1 \\ 1 & \cdots & 1 & -(N-1) \end{pmatrix}.$$

We can thus compute the covariance using the diagonalization $e^{\mathbf{K}t} = e^{\mathbf{K}^T t} = \mathbf{V}e^{\Lambda t}\mathbf{V}^{-1}$. In addition, we will assume each area receives noise with identical statistics ($D_{jj} = D_l$) and there are identical noise correlations between areas ($D_{jk} = D_c$ for $j \neq k$), so $\mathbf{D} = (D_l - D_c)I + D_c J_N$. Multiplying and integrating (Equation 34), we find the diagonal entries (variances) of $\langle \mathbf{\Delta}(t)\mathbf{\Delta}^T(t) \rangle$ are

$$\langle \Delta_j(t)^2 \rangle = \frac{D_l + (N-1)D_c}{N}t + \frac{(N-1)(D_l - D_c)}{2N^2\kappa}\left[1 - e^{-2N\kappa t}\right], \tag{35}$$

and the off-diagonal entries (true covariances) are

$$\langle \Delta_j(t)\Delta_k(t) \rangle = \frac{D_l + (N-1)D_c}{N}t - \frac{(D_l - D_c)}{2N^2\kappa}\left[1 - e^{-2N\kappa t}\right].$$

As revealed by the diffusive term in Equation (35), the system still possesses a rotational symmetry, given by the action of rotating all the bumps in the same direction. Thus, the component of noise in this direction is not damped out by coupling. Thus, note that the long term variance of any bump's position $\Delta_j(t)$ will be approximately described by the averaged diffusion

$$\lim_{t \to \infty} \langle \Delta_j(t)^2 \rangle = \frac{D_l + (N-1)D_c}{N}t.$$

As the strength of coupling κ or number of areas N is increased, the variances $\langle \Delta_j(t)^2 \rangle$ approach this limit at a faster rate, since the other portions of variance decay at a rate proportional to $|\lambda_2| = N\kappa$. Note also that in the limit $D_c \to D_l$, effects of coupling are negligible and the long term variance of each bump is determined by the diffusion introduced by its area's internal noise.
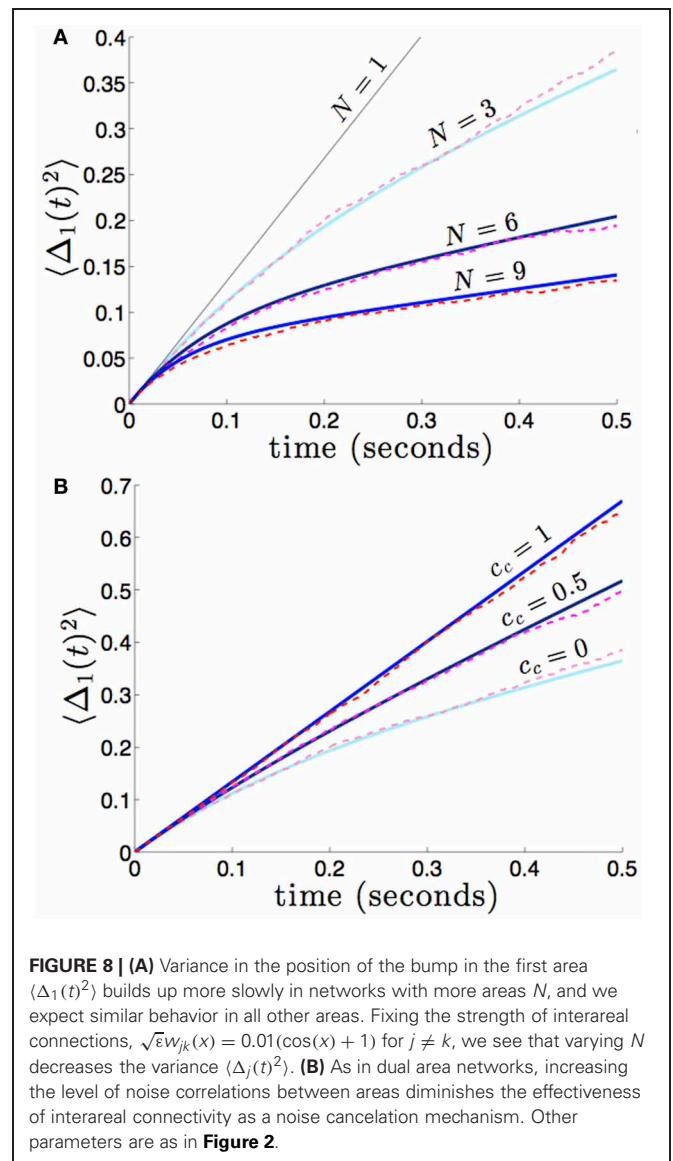
Returning to study the full variance Equation (35) for symmetric coupling and noise, we make a few observations. First, in the limit of purely correlated noise across areas ($D_c \to D_l$), interareal connections have no effect, and $\langle \Delta_j(t)^2 \rangle = D_l t$ for all areas

and arbitrary coupling strength. However, if there is any independent noise in each area ($D_c < D_l$), variance $\langle \Delta_j(t)^2 \rangle$ can always be reduced further by increasing coupling strength or the number of areas since

$$\frac{d}{d\kappa}\langle \Delta_j(t)^2 \rangle = \frac{(N-1)(D_l - D_c)}{2N^2} \times \frac{(1 + 2N\kappa)e^{-2N\kappa t} - 1}{\kappa^2} \leq 0,$$

where inequality $(1 + 2N\kappa t) \geq e^{2N\kappa t}$ holds due to the Taylor expansion of $e^{2N\kappa t}$ when $N\kappa t \geq 0$, and

$$\frac{d}{dN}\langle \Delta_j(t)^2 \rangle = -\frac{D_l - D_c}{N^2}$$
$$+ \frac{D_l - D_c}{2N^3\kappa}\left[2(1 + N\kappa t)e^{-2N\kappa t} - N\right] \leq 0$$

FIGURE 8 | (A) Variance in the position of the bump in the first area $\langle \Delta_1(t)^2 \rangle$ builds up more slowly in networks with more areas N, and we expect similar behavior in all other areas. Fixing the strength of interareal connections, $\sqrt{\varepsilon}w_{jk}(x) = 0.01(\cos(x) + 1)$ for $j \neq k$, we see that varying N decreases the variance $\langle \Delta_j(t)^2 \rangle$. **(B)** As in dual area networks, increasing the level of noise correlations between areas diminishes the effectiveness of interareal connectivity as a noise cancelation mechanism. Other parameters are as in **Figure 2**.

when $N \geq 2$, since $D_l \geq D_c$ and due to the Taylor expansion of $e^{2N\kappa t}$. Note, we have temporarily treated N as a continuous variable. Thus, we know the variance $\langle \Delta_j(t)^2 \rangle$ to decrease with increasing κ and expect it to decrease with increasing N.

We can compute the variance $\langle \Delta_j(t)^2 \rangle$ explicitly in the case of Heaviside firing rate functions (Equation 5), cosine synaptic weights (Equation 3) and (Equation 4). With these assumptions, as well as there being identical noise to all areas ($c_{jj} = 1$ for all j, $c_{jk} = c_c$ for $j \neq k$), we find

$$D_l = \frac{\varepsilon}{2 + 2\sqrt{1 - \theta^2}}, \quad D_c = \frac{c_c \varepsilon}{2 + 2\sqrt{1 - \theta^2}},$$

so that

$$\langle \Delta_j(t)^2 \rangle = \frac{(1 + (N-1)c_c)\varepsilon}{2N(1 + \sqrt{1 - \theta^2})} t + \frac{(1 - c_c)\varepsilon}{4N^2\kappa} \left[1 - e^{-2N\kappa t} \right],$$
(36)

which reflects the fact that increasing the number of areas will decrease variability, when noise between areas is not too strongly correlated. We demonstrate the accuracy of this formula (36) in **Figure 8**. In numerical simulations, as predicted by our asymptotic calculations, the variance scales more slowly in time in networks with more areas.

DISCUSSION

We have shown that interareal coupling in multi-area stochastic networks can reduce the diffusive wandering of bumps. Since bump attractors offer a well studied model of persistent activity underlying spatial working memory (Compte et al., 2000), our results provide a novel suggestion for how the memory networks may reduce error. Our calculations have exploited a small noise approximation for the position of the bump in each area (Armero et al., 1998; Bressloff and Webber, 2012a). Assuming connectivity between areas is weak, we have shown the equations describing bump positions reduce to a multivariate Ornstein–Uhlenbeck process. In this formulation, we find interareal connectivity stabilizes all but one eigendirection in the space of bump position movements. Neutral stability does still exist, so stochastic forces that move bumps in all areas in the same direction do not decay away. However, sources of noise that force bumps in opposite directions create bump movements that will decay with time. Thus, interareal connectivity provides a noise cancelation

mechanism that operates by stabilizing the bumps in each area to stochastic forces that push them in opposite directions. (Polk et al., 2012) recently explored noise correlation statistics in persistent state networks that reduce wandering. Our work complements these results by studying synaptic architectures that limit persistent state diffusion.

Storing spatial working memories with neural activity that spans multiple brain areas does serve other purposes than potential noise cancelation. Delayed response tasks that lead to limb motion can generate persistent activity in the parietal cortex (Colby et al., 1996; Pesaran et al., 2002) so that motor responses can be readily executed. In addition, superior colliculus demonstrates sustained activity (Basso and Wurtz, 1997), which is an area also thought to underlie directed behavioral responses. Therefore, activity is distributed between areas providing short term information storage, like prefrontal cortex (Goldman-Rakic, 1995), and those responsible for motor responses and/or behavior. An additional effect of this delegation of activity is that reciprocal connections between areas may provide noise cancelation during the storage period of working memory. However, our work suggests distributing working memory-serving neural activity between areas that receive strongly correlated noise will not provide as effective cancelation.

Our work should be contrasted with several other results concerning the stabilization of networks that encode a continuous variable (Koulakov et al., 2002; Goldman et al., 2003; Cain and Shea-Brown, 2012; Kilpatrick et al., 2013). Pure integrators, which are usually line attractors, are notoriously fragile to parametric perturbations, so (Koulakov et al., 2002) suggested they may be made more robust by considering networks that integrate in discrete bursts, rather than continuously. This can be implemented by considering a population of bistable neural units so that firing rate integration of a stimulus occurs in a stairstep fashion, rather than a ramplike fashion (see Goldman et al., 2003 for example). Related ideas were recently implemented in a bump attractor model of spatial working memory (Kilpatrick et al., 2013), but quantization was implemented with synaptic architecture rather than single neural unit properties. As opposed to the approach of quantizing the space of possible stimulus representations, we have kept the representation space a continuum. Deleterious effects of noise are reduced by considering reciprocal connectivity between encoding areas that redundantly represent the stimulus. Due to noise cancelations, the encoding error of the network decreases as the number of areas is increased.

REFERENCES

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87. doi: 10.1007/BF00337259

Armero, J., Casademunt, J., Ramirez-Piscina, L., and Sancho, J. M. (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

Basso, M. A., and Wurtz, R. H. (1997). Modulation of neuronal activity by target uncertainty. *Nature* 389, 66–69. doi: 10.1038/37975

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848. doi: 10.1073/pnas.92.9.3844

Bressloff, P. C., and Webber, M. A. (2012a). Front propagation in

stochastic neural fields. *SIAM J. Appl. Dyn. Syst.* 11, 708–740. doi: 10.1137/110851031

Bressloff, P. C., and Webber, M. A. (2012b). Neural field model of binocular rivalry waves. *J. Comput. Neurosci.* 32, 233–252. doi: 10.1007/s10827-011-0351-y

Cain, N., and Shea-Brown, E. (2012). Computational models of decision making: integration, stability, and noise. *Curr. Opin.*

Neurobiol. 22, 1047–1053. doi: 10.1016/j.conb.2012.04.013

Camperi, M., and Wang, X. J. (1998). A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* 5, 383–405. doi: 10.1023/A:1008837311948

Chafee, M. V., and Goldman-Rakic, P. S. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during

a spatial working memory task. *J. Neurophysiol.* 79, 2919–2940.

Clements, J. D., Lester, R. A., Tong, G., Jahr, C. E., and Westbrook, G. L. (1992). The time course of glutamate in the synaptic cleft. *Science* 258, 1498–1501. doi: 10.1126/science.1359647

Colby, C. L., Duhamel, J. R., and Goldberg, M. E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.* 76, 2841–2852.

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. *Cereb. Cortex* 10, 910–923. doi: 10.1093/cercor/10.9.910

Constantinidis, C., and Wang, X.-J. (2004). A neural circuit basis for spatial working memory. *Neuroscientist* 10, 553–565. doi: 10.1177/1073858404268742

Curtis, C. E. (2006). Prefrontal and parietal contributions to spatial working memory. *Neuroscience* 139, 173–180. doi: 10.1016/j.neuroscience.2005.04.070

Curtis, C. E., and Lee, D. (2010). Beyond working memory: the role of persistent activity in decision making. *Trends Cogn. Sci.* 14, 216–222. doi: 10.1016/j.tics.2010.03.006

di Pellegrino, G., and Wise, S. P. (1993). Visuospatial versus visuomotor activity in the premotor and prefrontal cortex of a primate. *J. Neurosci.* 13, 1227–1243.

Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Progress Phys.* 61, 353. doi: 10.1088/0034-4885/61/4/002

Feingold, D. G., and Varga, R. S. (1962). Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem. *Pacific J. Math.* 12, 1241–1250.

Ferster, D., and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* 23, 441–471. doi: 10.1146/annurev.neuro.23.1.441

Folias, S. E., and Ermentrout, G. B. (2011). New patterns of activity in a pair of interacting excitatory-inhibitory neural fields. *Phys. Rev. Lett.* 107:228103. doi: 10.1103/PhysRevLett.107.228103

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

Gardiner, C. W. (2003). *Handbook of Stochastic Methods.* New York, NY: Springer.

Goldman, M. S., Levine, J. H., Major, G., Tank, D. W., and Seung, H. S. (2003). Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cereb. Cortex* 13, 1185–1195. doi: 10.1093/cercor/bhg095

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6

Gupta, A., Wang, Y., and Markram, H. (2000). Organizing principles for a diversity of gabaergic interneurons and synapses in the neocortex. *Science* 287, 273–278. doi: 10.1126/science.287.5451.273

Hansel, D., and Mato, G. (2013). Short-term plasticity explains irregular persistent activity in working memory tasks. *J. Neurosci.* 33, 133–149. doi: 10.1523/JNEUROSCI.3455-12.2013

Hansel, D., and Sompolinsky, H. (1998). "Modeling feature selectivity in local cortical circuits," in *Methods in Neuronal Modeling: From Ions to Networks*, chapter 13, eds C. Koch, and I. Segev (Cambridge: MIT), 499–567.

Häusser, M., and Roth, A. (1997). Estimating the time course of the excitatory synaptic conductance in neocortical pyramidal cells using a novel voltage jump method. *J. Neurosci.* 17, 7606–7625.

Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. iii. activities related to expectation of target and reward. *J. Neurophysiol.* 61, 814–832.

Itskov, V., Hansel, D., and Tsodyks, M. (2011). Short-term facilitation may stabilize parametric working memory trace. *Front. Comput. Neurosci.* 5:40. doi: 10.3389/fncom.2011.00040

Kilpatrick, Z. P., and Bressloff, P. C. (2010). Binocular rivalry in a competitive neural network with synaptic depression. *SIAM J. Appl. Dyn. Syst.* 9, 1303–1347. doi: 10.1137/100788872

Kilpatrick, Z. P., and Ermentrout, B. (2013). Wandering bumps in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.* 12, 61–94. doi: 10.1137/120877106

Kilpatrick, Z. P., Ermentrout, B., and Doiron, B. (2013). Optimizing working memory with spatial heterogeneity of recurrent cortical excitation. (Submitted).

Koulakov, A. A., Raghavachari, S., Kepecs, A., and Lisman, J. E. (2002). Model for a robust neural integrator. *Nat. Neurosci.* 5, 775–782. doi: 10.1038/nn893

Laing, C. R., and Chow, C. C. (2001). Stationary bumps in networks of spiking neurons. *Neural. Comput.* 13, 1473–1494. doi: 10.1162/089976601750264974

Levy, R., Friedman, H. R., Davachi, L., and Goldman-Rakic, P. S. (1997). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* 17, 3870–3882.

Ly, C., and Ermentrout, G. B. (2010). Coupling regularizes individual units in noisy populations. *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.* 81(1 Pt 1):011911. doi: 10.1103/PhysRevE.81.011911

McNab, F., and Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107. doi: 10.1038/nn2024

Miller, E. K., Erickson, C. A., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16, 5154–5167.

Mushiake, H., and Strick, P. L. (1995). Pallidal neuron activity during sequential arm movements. *J. Neurophysiol.* 74, 2754–2758.

Owen, A. M., Evans, A. C., and Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cereb. Cortex* 6, 31–38. doi: 10.1093/cercor/6.1.31

Pesaran, B., Pezaris, J. S., Sahani, M., Mitra, P. P., and Andersen, R. A. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat. Neurosci.* 5, 805–811. doi: 10.1038/nn890

Ploner, C. J., Gaymard, B., Rivaud, S., Agid, Y., and Pierrot-Deseilligny, C. (1998). Temporal limits of spatial working memory in humans. *Eur. J. Neurosci.* 10, 794–797. doi: 10.1046/j.1460-9568.1998.00101.x

Polk, A., Litwin-Kumar, A., and Doiron, B. (2012). Correlated neural variability in persistent state networks. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6295–6300. doi: 10.1073/pnas.1121274109

White, J. M., Sparks, D. L., and Stanford, T. R. (1994). Saccades to remembered target locations: an analysis of systematic and variable errors. *Vis. Res.* 34, 79–92. doi: 10.1016/0042-6989(94)90259-3

Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cybern.* 13, 55–80. doi: 10.1007/bf00288786