Coherence influences the dimensionality of communication subspaces

Summary: The brain relies on communication between specialized cortical areas to accomplish complex cognitive tasks. To understand this ability of the brain, we need a deeper understanding of information transfer across cortical areas. There are two leading hypotheses for communication between cortical areas: 1) The communication through coherence (CTC) hypothesis posits that coherent oscillations between source and target populations of neurons are required for information propagation [**1**]. 2) The information transmission via communication subspace (CS) hypothesis, recently introduced by Semedo et al. [**2**], advances the idea that low-dimensional subspaces of population activity are responsible for communication across cortical areas. There is a clear divide between these two mechanisms and the CTC hypothesis, in particular, has been surrounded by considerable skepticism, with many authors reducing oscillations in the cortex to an epiphenomenon (*viz.*, a by-product with no functional role).

Here, we reconcile these two mechanisms of communication through a spectral decomposition of communication subspaces. In our main result, we predict that coherence influences the dimensionality of the CS: Dimensionality is lowest, and the prediction performance is highest, at frequencies that exhibit a peak in coherence between a source and target population (e.g., V1 to V2). We arrive at these results by developing an analytical theory of communication for circuits described by stochastic dynamical systems exhibiting fixed-point solutions. We compute directly (*i.e.*, by a simulation-free method) the predictive performance for the mean-subtracted activity of a target population from that of a source population, and show that our predictions are in agreement with the experimental results by Semedo et al. [**2**]. Then via a band-pass filtered version of the covariance matrix, we arrive at our main result. Hence, our theory makes experimentally-testable predictions of how oscillations influence interareal communication while advancing a new hypothesis for the functional role of oscillatory activity in the brain.

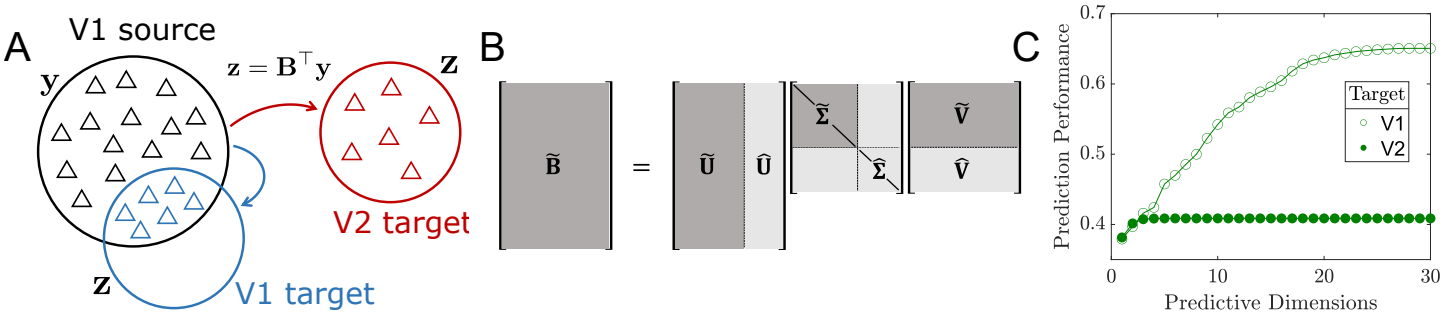

Figure 1: **Communication subspace analysis. A**) A schematic for the model, based on the experimental protocol of Semedo et al. [**2**]. Triangles are individual neurons. Source population ($\mathbf{y}$) is a randomly chosen subset of neurons simulating V1 activity. The target population ($\mathbf{z}$) is a randomly chosen subset of either V1 or V2 neurons. A linear model ($\mathbf{z} = \mathbf{B}^\top \mathbf{y}$) is used to predict the activity of the target from the source. **B**) Reduced-rank-regression. Predictive performance is computed as a function of the n largest principal components (predictive dimensions). **C**) Prediction performance vs. predictive dimensions for a nonlinear recurrent model implementing normalization [**3**]. Smaller dimensionality and performance when the target is V2.

Additional detail: We begin with a brief background on the theory of Stochastic Differential Equations (SDEs) around fixed points. Consider a general SDE representing the time evolution of the activity of neurons in two cortical areas (V1 and V2) of the form,

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{L}\mathrm{d}\mathbf{W} \tag{1}$$

Where the $\mathbf{x} \in \mathbb{R}^n$ is a vector representing the activity of neurons ; $\mathbf{f}(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n$ is a nonlinear function of $\mathbf{x}$ and the input $\mathbf{u}$; and $\mathbf{L}\mathrm{d}\mathbf{W}$ is an additive noisy drive, where $\mathrm{d}\mathbf{W}$ is a vector of mutually-independent Gaussian increments with correlation matrix $\mathbf{D}$. Assuming the system of equations exhibits fixed-point solutions, we can linearize the system about the fixed point to arrive at a wide-sense stationary Gaussian process, $\mathbf{x}(t)$. The power spectral density matrix, $\mathcal{S}(\omega) \in \mathbb{R}^{n \times n}$, is then given by

$$\mathcal{S}(\omega) = \mathbb{E}\left[|\hat{\mathbf{x}}(\omega)|^2\right] = (\iota\omega\mathbf{I} + \mathbf{J})^{-1}\,\mathbf{L}\,\mathbf{D}\,\mathbf{L}^\top\,(-\iota\omega\mathbf{I} + \mathbf{J})^{-\top} \tag{2}$$

where $\mathbf{J}$ is the Jacobian at the fixed point, and $\hat{\mathbf{x}}(\omega) = \mathcal{F}[x(t)]$. The correlation matrix for $\mathbf{x}(t)$, $\mathcal{C}(0)$, can be obtained from the Fokker-Planck formalism as the solution to the Lyapunov equation,

$$\mathbf{J}\,\mathcal{C}(0) + \mathcal{C}(0)\,\mathbf{J}^\top = -\mathbf{L}\,\mathbf{D}\,\mathbf{L}^\top \tag{3}$$

Now, to arrive at an analytical theory of communication subspaces, we follow the experimental protocol of Semedo et al. [**2**] (summarized in Fig. 1a-b), and divide the mean-subtracted responses of the total population of neurons, $\mathbf{x}$, into the mean-subtracted response of the source neurons, $\mathbf{y}$ (V1), and the target neurons, $\mathbf{z}$ (V1 or V2). For a linear readout $\mathbf{z} = \mathbf{B}^\top\mathbf{y}$, minimization of the L2 norm yields $\mathbf{B}_{opt} = \mathbf{C_1}^{-1}\mathbf{C_3}$ with residual mean squared error

$$\epsilon = \mathrm{Tr}(\mathbf{C_2} + \mathbf{B}^\top\mathbf{C_1}\mathbf{B} - 2\mathbf{B}^\top\mathbf{C_3}), \tag{4}$$

where $\mathbf{C_1} = \mathbb{E}[\mathbf{y}\mathbf{y}^\top]$, $\mathbf{C_2} = \mathbb{E}[\mathbf{z}\mathbf{z}^\top]$ and $\mathbf{C_3} = \mathbb{E}[\mathbf{y}\mathbf{z}^\top]$. These matrices can be obtained directly from $\mathcal{C}(0)$.

Given $\mathbf{B}_{opt}$, we use reduced-rank regression (Fig. 1b) to find the dimensionality of the CS within and across areas. We apply our theory to a linear recurrent (LR) circuit, and to a nonlinear recurrent (NR) circuit implementing normalization exactly [**3**] and find that the dimensionality of the subspace is significantly lower for interareal compared to within-area communication (Fig 1c), in agreement with experimental observations [**1**].

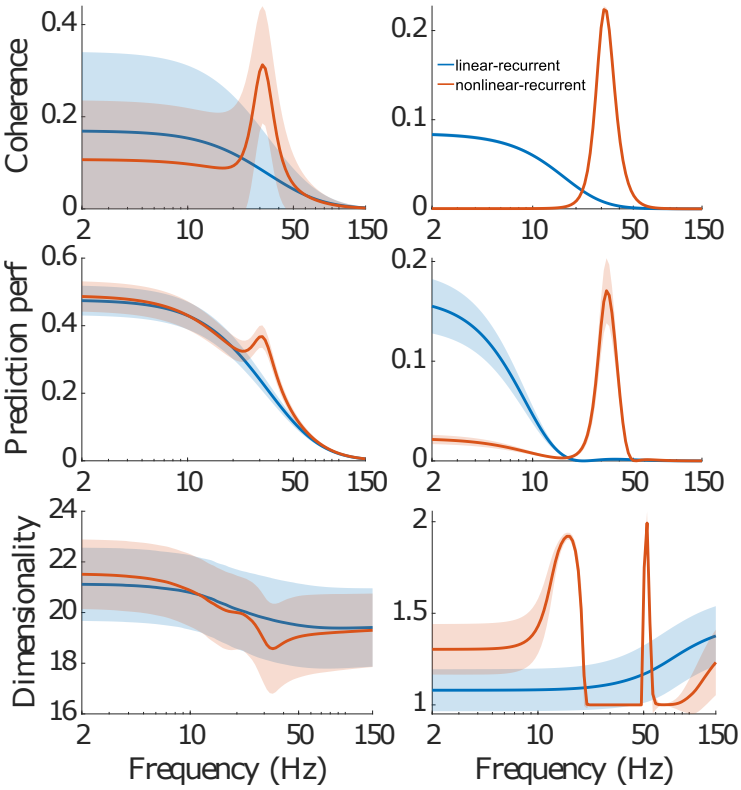

Figure 2: **Spectral decomposition of communication subspaces for linear and nonlinear recurrent models.** Source: V1. **Left:** Target: V1. **Right:** Target: V2. **A**) Coherence ($\kappa_{ij} = |\mathcal{S}_{ij}|^2/(\mathcal{S}_{ii}\mathcal{S}_{jj})$) between neurons with the largest response in the source and target populations. **B**) Optimal prediction performance: ϵ of $\mathbf{B}_{opt}$. **C**) Predictive dimensions (for >95% of ϵ_{opt}) as a function of frequency. For **B** and **C**, the envelope represents 1 SD from the mean.

Equipped with this theory, we proceed to consider a Fourier decomposition of the correlation matrix of the stochastic process $\mathbf{x}(t)$. We find that the covariance matrix of a band-pass (given by a delta function, $\delta(\omega)$) filtered signal can be expressed as $\mathcal{C}_\omega(0) = 2\mathrm{Re}[\mathcal{S}(\omega)]$. Then, from $\mathcal{C}_\omega(0)$ we compute the maximum prediction performance and the CS dimensionality as a function of frequency for the same circuit models. For the LR model we find that: 1) For V1 target, the CS dimension is a monotonic function of the prediction performance; while 2) for V2 target, it is roughly constant. Crucially, for the NR model – which exhibits bursty oscillatory activity in the γ-band – we find that for both V1 and V2 targets: 1) There is a peak in predictive performance in the γ-band; while 2) the CS dimension is lowest in the γ-band. Both features coincide with the peak in the coherence between source and target populations. This points to a frequency-selective mechanism for information transfer across cortical areas since the predictive performance is maximum, and the dimensionality lowest, in the γ-band, offering a way for communication via subspaces that is shaped by coherence.

Our theoretical analysis appears to reconcile two competing hypotheses on the mechanism of interareal communication, and consequently, we identify a possible functional role for oscillations in corticocortical communication (against our prior belief). Future work should focus on testing the generality of this result with respect to circuit models, and on attempting an experimental validation of these predictions.

References:

1. Fries, P., Trends in cognitive sciences, 9(10), 474-480 (2005).

2. Semedo, João D., et al., Neuron, 102(1), 249-259 (2019).

3. Heeger, D. J., Mackey, W. E., PNAS, 116(45), 22783-22794 (2019).