

# Multi-stage Cortical Recurrent Circuit Implementing Normalization

**Asit Pal (ap6603@nyu.edu)**

Department of Chemistry, New York University  
Center for Soft Matter Research, Dept. of Physics, New York University  
NYC, NY, USA

**Shivang Rawat (sr6364@nyu.edu)**

Courant Institute of Mathematical Sciences, New York University  
Center for Soft Matter Research, Dept. of Physics, New York University  
NYC, NY, USA

**David Heeger (david.heeger@nyu.edu)**

Dept. Of Psychology, New York University  
Center for Neural Science, New York University  
NYC, NY, USA

**Stefano Martiniani (stefano.martiniani@nyu.edu)**

Department of Chemistry, New York University  
Center for Soft Matter Research, Dept. of Physics, New York University  
Courant Institute of Mathematical Sciences, New York University  
Simons Center for Computational Physical Chemistry, New York University  
NYC, NY, USA

## Abstract

Communication between cortical areas is supported by long-range reciprocal connections. Given feedback connections' hypothesized role in attentional modulation, it is essential to develop a multistage network model of the brain for studying attention and inter-area communication. Here, we present a dynamically stable hierarchical recurrent neural circuit model with feedback that implements divisive normalization *exactly* at each stage of its hierarchy. We consider a two-stage model (V1 and V4), each stage receives input from the preceding area and feedback from the subsequent area and the responses in each area are normalized by local inhibitory signals. We note that an increase in feedback from V4 to V1, amplifies responses in both stages, with a more pronounced increase in higher cortical areas (Fig.1), consistent with experimental findings (Maunsell & Cook, 2002). Additionally, our model predicts that feedforward and feedback signals in the brain propagate via distinct frequency channels, gamma and alpha frequencies respectively, in line with empirical evidence (Van Kerkoerle et al., 2014; Bastos et al., 2015). Furthermore, our model admits a low-dimensional communication subspace (within and across areas) and predicts that enhancing feedback improves inter-areal communication, yet decreases within-area communication (Fig.3). In summary, our hierarchical model provides a robust and analytically tractable framework for exploring normalization, attention, and inter-areal communication.

**Keywords:** Multi-stage; Recurrent Neural Network; Attention; Communication; Coherence

## Multi-stage model

For each cortical area, the principal output response is given by input gain  $\times$  input drive + recurrent gain  $\times$  recurrent drive + feedback gain  $\times$  feedback drive, and the corresponding dynamical equation is given by,

$$\tau_{y_1} \frac{dy_1}{dt} = -y_1 + \beta [z_1] + \frac{1}{1 + [a_1]} (\mathbf{W}_{11}[y_1] + \gamma \mathbf{W}_{12}[y_2]) \quad (1)$$

where  $y_1$  and  $y_2$  denotes the membrane potentials of neurons in cortical areas V1 and V4, with  $[y_1]^2$ ,  $[y_2]^2$  representing their respective firing rates approximated by the half-wave rectified and squared membrane potentials (Carandini & Heeger, 1994).  $z_1$  denotes the input drive from the preceding cortical area, and  $a_1$  is the inhibitory neuron, implementing divisive normalization through modulation of recurrent excitation.  $\mathbf{W}_{11}$  is the recurrent weight matrix for V1, while  $\mathbf{W}_{12}$  is the inter-areal connectivity matrix from V4 to V1.  $\beta$  and  $\gamma$  modulate the relative input and feedback gains in the excitatory ( $y_1$ ) and inhibitory ( $a_1$ ) populations. Normalization in our circuit model is achieved by the two modulatory neurons  $\mathbf{a}$  and  $\mathbf{u}$  whose dynamics are prescribed by the following equations,

$$\begin{aligned} \tau_{a_1} \frac{da_1}{dt} &= -a_1 + (1 + a_1) * \sqrt{[u_1]} + \frac{g \mathbf{W}_{12}[y_2]}{[y_1]} + \alpha_1 \frac{du_1}{dt} \\ \tau_{u_1} \frac{du_1}{dt} &= -u_1 + (b \sigma_1)^2 + \mathbf{W}_{n1} ([y_1]^2 * [u_1]) \end{aligned} \quad (2)$$

where  $\mathbf{W}_{n1}$  is the normalization connectivity matrix (see Eq. 3).  $b$  and  $g$  denote baseline values for the input and feedback gating variables. Division of vectors, and multiplication

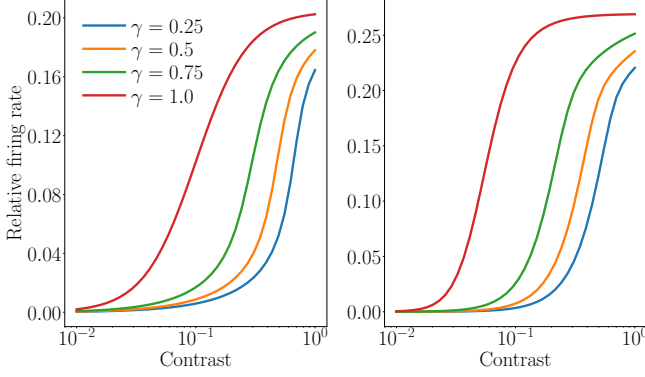


Figure 1: **Effect of feedback-mediated attention on principal neuron responses.** **Left:** Lower cortical area (V1). **Right:** Higher cortical area (V4). Increasing attention enhances cortical responses, with greater impact in higher cortical areas.

denoted by ‘\*’, are element-wise. When  $\gamma$  and  $g$  are set to unity, so that  $\mathbf{y}_1$  and  $\mathbf{a}_1$  receive identical feedback input, and  $\mathbf{W}_{11}$  has spectral radius  $\rho < 1$ , the steady-state response of the principal neurons in V1 follows the normalization equation exactly (Heeger & Mackey, 2019).

$$|\mathbf{y}_1|^2 = \frac{|\beta \mathbf{z}_1|^2}{(b \sigma_1)^2 + \mathbf{W}_{n1} |\beta \mathbf{z}_1|^2} \quad (3)$$

Note that we’ve only presented the dynamical equations for V1; V4 follows similar dynamics. We observe that to maintain dynamics with a stable fixed point (vs. a limit cycle corresponding to sustained oscillations), the feedback to the inhibitory neuron population must be equal to or higher than that to the excitatory neuron population, in line with studies of attention in Macaques (Mitchell, Sundberg, & Reynolds, 2007; Snyder, Morais, & Smith, 2016).

### Coherence and Inter-Areal Communication

Given that we have the fixed-point solution for our system of equations (Eq. 3), linearizing around the fixed point and adding to each equation an additive noise drive  $\mathbf{L}d\mathbf{W}$ , where  $d\mathbf{W}$  is a vector of independent Gaussian increments with correlation matrix  $\mathbf{D}$ , leads to a wide-sense stationary Gaussian process,  $\mathbf{x}(t)$ , whose power spectral density matrix is given by

$$\mathcal{S}(\omega) = (i\omega\mathbf{I} + \mathbf{J})^{-1} \mathbf{L} \mathbf{D} \mathbf{L}^T (-i\omega\mathbf{I} + \mathbf{J})^{-T} \quad (4)$$

where  $\mathbf{J}$  is the Jacobian at the fixed point. From  $\mathcal{S}(\omega)$  we can directly compute the coherence,  $\kappa_{ij} = |S_{ij}|^2 / (S_{ii}S_{jj})$ , between the maximally firing neurons in the two areas as a function of feedback strength (Fig. 2). We observe two distinct frequency channels for feedforward and feedback signals, consistent with experimental findings (Bastos et al., 2015; Van Kerkoerle et al., 2014). Moreover, enhancing the relative feedback (by increasing  $\gamma$ ) causes a redshift, whereas boosting the input gain (by increasing  $\beta$ ) results in a blueshift in the peak coherence frequency (Fig. 2).

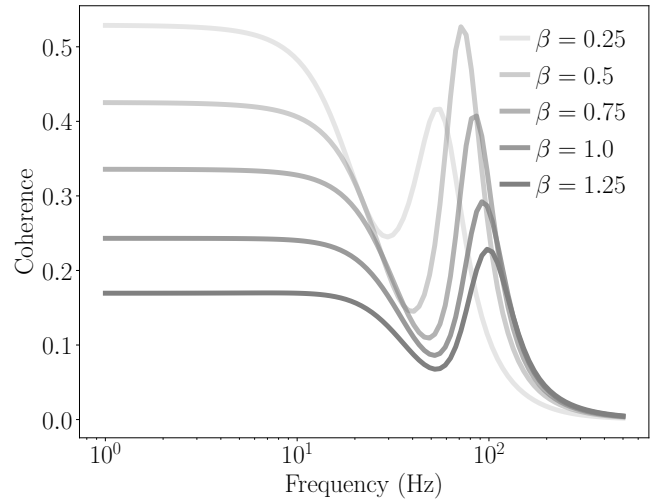
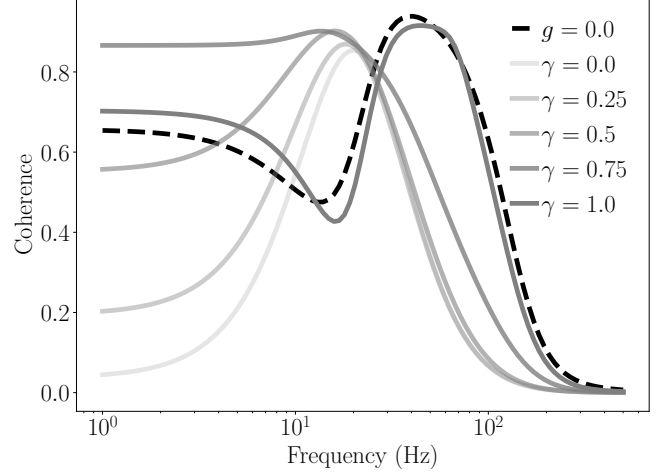


Figure 2: **Effect of attention (mediated by feedback and input gain) on coherence between V1 and V4.** **Top:** The relative feedback to the excitatory neuron ( $\mathbf{y}_1$ ) w.r.t the inhibitory neuron ( $\mathbf{a}_1$ ) can be modulated by varying the parameter  $\gamma$ . As we increase feedback to  $\mathbf{y}_1$  by increasing  $\gamma$ , the peak coherence frequency is redshifted towards the alpha frequency band. In the absence of any feedback ( $\gamma = g = 0$ ) the peak shifts to the gamma frequency band, analogous to the scenario when there is equal feedback to  $\mathbf{y}_1$  and  $\mathbf{a}_1$ . **Bottom:** The relative input to  $\mathbf{y}_1$  w.r.t  $\mathbf{a}_1$  can be modulated by varying the parameter  $\beta$ . As we increase  $\beta$ , the peak coherence frequency is blue-shifted.

The correlation matrix for  $\mathbf{x}(t)$ ,  $\mathbf{C}(0)$ , can be obtained as the solution to the Lyapunov equation

$$\mathbf{J} \mathbf{C}(0) + \mathbf{C}(0) \mathbf{J}^T = -\mathbf{L} \mathbf{D} \mathbf{L}^T \quad (5)$$

Given  $\mathbf{C}(0)$ , we follow the method used by (Semedo, Zandvakili, Machens, Byron, & Kohn, 2019) to arrive at our analytical model of communication subspace (CS). We divide the mean-subtracted responses of the total population of neurons,  $\mathbf{x}$ , into the mean-subtracted responses of the source neurons,

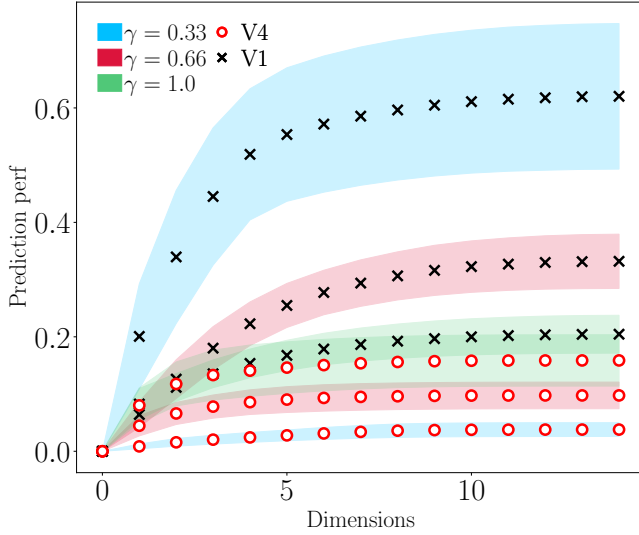


Figure 3: **Effect of feedback-mediated attention on within and inter-area communication subspaces.** **Source:** V1, **Target:** V1 (within-area), V4 (inter-area). Increasing the relative feedback by increasing  $\gamma$  improves the inter-areal prediction performance, while within-area prediction performance declines.

$\mathbf{y}$  (V1), and of the target neurons,  $\mathbf{z}$  (V4). For a linear readout  $\mathbf{z} = \mathbf{B}^T \mathbf{y}$ , minimization of the  $L_2$  norm yields  $\mathbf{B}_{\text{opt}} = \mathbf{C}_1^{-1} \mathbf{C}_3$  with residual mean squared error  $\varepsilon = \text{Tr}(\mathbf{C}_2 + \mathbf{B}^T \mathbf{C}_1 \mathbf{B} - 2\mathbf{B}^T \mathbf{C}_3)$ , where  $\mathbf{C}_1 = \mathbb{E}[\mathbf{y}\mathbf{y}^T]$ ,  $\mathbf{C}_2 = \mathbb{E}[\mathbf{z}\mathbf{z}^T]$  and  $\mathbf{C}_3 = \mathbb{E}[\mathbf{y}\mathbf{z}^T]$ . These matrices can be obtained directly from  $\mathbf{C}(0)$ . Given  $\mathbf{B}_{\text{opt}}$ , we use reduced-rank regression to find the dimensionality of the CS within and across areas. We find that an increase in feedback-mediated attention correlates with improved inter-areal prediction performance (Fig. 3) while the dimensionality of the subspace remains unchanged, whereas within-area communication diminishes significantly with increased feedback.

## Acknowledgements

The authors acknowledge valuable discussions with Adam Kohn, Ruben Coen-Cagli, John Maunsell, and Jason MacLean. This work was supported by the National Institute of Health under award number R01EY035242. S.M. acknowledges the Simons Center for Computational Physical Chemistry for financial support. This work was supported in part through the NYU IT High-Performance Computing resources, services, and staff expertise.

## References

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., ... Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401.

- Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, *264*(5163), 1333–1336.
- Heeger, D. J., & Mackey, W. E. (2019). Oscillatory recurrent gated neural integrator circuits (organics), a unifying theoretical framework for neural dynamics. *Proceedings of the National Academy of Sciences*, *116*(45), 22783–22794.
- Maunsell, J. H., & Cook, E. P. (2002). The role of attention in visual processing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1063–1072.
- Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area v4. *Neuron*, *55*(1), 131–141.
- Semedo, J. D., Zandvakili, A., Machens, C. K., Byron, M. Y., & Kohn, A. (2019). Cortical areas interact through a communication subspace. *Neuron*, *102*(1), 249–259.
- Snyder, A. C., Morais, M. J., & Smith, M. A. (2016). Dynamics of excitatory and inhibitory networks are differentially altered by selective attention. *Journal of neurophysiology*, *116*(4), 1807–1820.
- Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., Van Der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, *111*(40), 14332–14341.