Draft of new V1 section

Sean Bittner, Agostina Palmigiano

October 6, 2020

1 Exploratory analysis of V1 with EPI produces a novel theory

Dynamical models of excitatory (E) and inhibitory (I) populations with supralinear input-output function have succeeded in explaining a host of experimentally documented phenomena. In a regime characterized by inhibitory stabilization of strong recurrent excitation, these models give rise to paradoxical responses [1], selective amplification [2, 3], surround suppression [4] and normalization [5]. Despite their strong predictive power, E-I circuit models rely on the assumption that inhibition can be studied as an indivisible unit. However, experimental evidence shows that inhibition is composed of distinct elements – parvalbumin (P), somatostatin (S), VIP (V) – composing 80% of GABAergic interneurons in V1 [6, 7, 8], and that these inhibitory cell types follow specific connectivity patterns (Fig. 1A) [9]. Recent theoretical advances [10, 11, 12], have only started to address the consequences of this multiplicity in the dynamics of V1, strongly relying on linear theoretical tools. Here, we use EPI to gain a comprehensive understanding of how external circuit inputs govern S-V flipping and modulate variability.

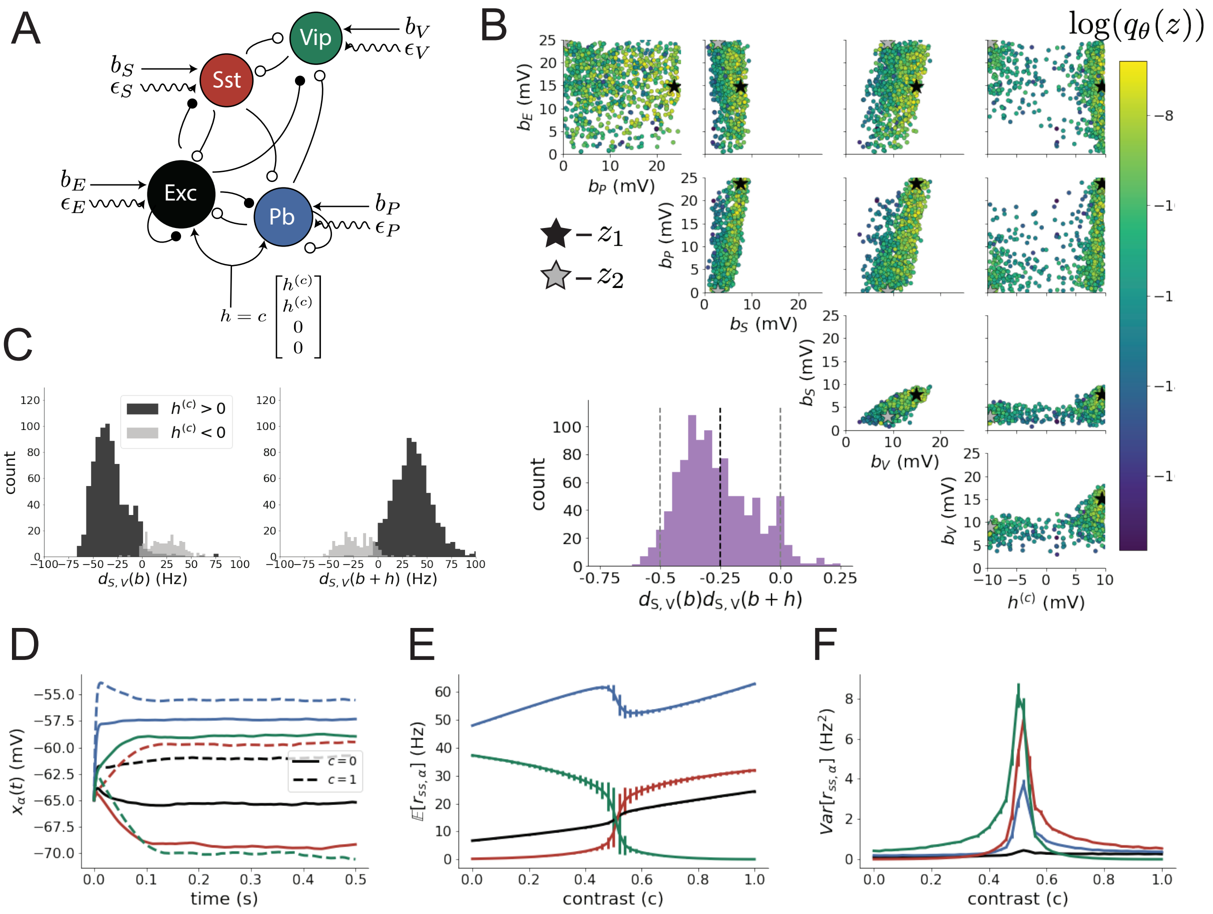
We consider contrast responses of a nonlinear dynamical V1 circuit model (Fig. 1A) with a state comprised of each neuron-type population's average membrane potential $x = [x_E, x_P, x_S, x_V]^\top$ (mV). Specifically, we examine the stochastic stabilized supralinear network [13], generalized to have inhibitory multiplicity (see Section 2). Each population of neuron-type α receives recurrent input $(Wf_r(x))_\alpha$ from synaptic projections, where W is the connectivity matrix and $f_r = [\cdot]_+^2$ is the rate nonlinearity. W is calculated from recent experiments measuring post-synaptic potentials and connectivity rates in mice [14, 15]. Additionally, each population experiences a positive baseline input of b_α, a contrast (c)-dependent input h_α, and a stochastic input ϵ_α. Little is known about the scales of these baseline and contrast-dependent inputs, so we treat these as free parameters in our application of EPI:

$$z = \begin{bmatrix} b_E & b_P & b_S & b_V & h^{(c)} \end{bmatrix}^\top . \tag{1}$$

From first principles, we may attempt to fit or infer z based on recorded V1 population responses at varying contrasts (**TODO** AP to add citation). However, such a class of constrained models rarely produce good fits that generalize well. Instead, we can use EPI to condition z on one of the most salient properties of these data-sets called "S-V flipping" – a winner-take-all relationship between the S- and V-populations. At low contrast values, the V-population dominates the S-population, and at greater contrast values, the S-population dominates the V-population. We formulate the emergent property of S-V flipping with a normalized statistic measuring the difference in steady state (x_{ss}) between the S- and V-population at a given contrast:

$$d_{S,V}(c; z) = \frac{f_r(x_{ss,S}(c; z)) - f_r(x_{ss,V}(c; z))}{|f_r(x_{ss}(c; z))|_2} \tag{2}$$

Figure 1: **A**. Four-population model of primary visual cortex with excitatory (black), parvalbumin (blue), somatostatin (red), and VIP (green) neurons. Some neuron-types largely do not form synaptic projections to others (excitatory and inhibitory projections filled and unfilled, respectively). **B**. EPI posterior $q_\theta(z \mid \mathcal{B}_{S-V})$ for S-V flipping. The obtained posterior is visualized as 500 samples from the inferred distribution colored by $\log(q_\theta(z))$. This posterior is bimodal and concentrated in planes $h^{(c)} > 0$ and $h^{(c)} < 0$ at distinct modes z_1 (black star) and z_2 (gray, star), respectively. Bottom-left: Posterior predictive distribution of the emergent property statistics with respect to the constrained means (black, dashed line) and variances (gray, dashed lines at two standard deviations). **C**. Posterior predictive distribution of $d_{S,V}$ for each mode. The z_1-mode produces V-to-S flipping with increasing contrast. **D**. Model simulations at the mode z_1 at $c = 0$ (solid) and $c = 1$ (dashed). **E**. Mean system response at varying contrasts for z_1. Error bars show noise in rate. **F**. Plot of noise with contrast from E. Error bars are standard error measured across simulations.



For S and V to flip their steady states, the difference between the S- and V-population rates at no contrast ($c = 0$) must have opposite sign at full contrast ($c = 1$). Therefore, we stipulate the emergent property of S-V flipping to require the product of $d_{S,V}(c = 0)$ and $d_{S,V}(c = 1)$ to be appreciably negative. The means and variance of this emergent property statistics were chosen such that the posterior predictive distribution always produced S-V flipping.

$$\mathcal{B}_{S,V} \triangleq \begin{array}{r} \mathbb{E}\left[d_{S,V}(0)d_{S,V}(1)\right] = -0.25 \\ \text{Var}\left[d_{S,V}(0)d_{S,V}(1)\right] = 0.125^2 \end{array} \tag{3}$$

We ran EPI to obtain the parameter distribution of neuron-type population input parameters z that produce S-V flipping (Fig. 1B). We considered positive baseline inputs b_α from 0 to 25mV and contrast-dependent changes in $h^{(c)}$ from -10 to 10mV. The shape and structure of this distribution reveal some important properties of this model.

1. Two types of z result in S-V flipping. (There are two modes separated by the hyperplane $h^{(c)} = 0$.)
2. S-V flipping is sensitive to anticorrelated changes in b_S and b_V. (The pairwise marginal distribution of $q_\theta(b_S, b_V \mid \mathcal{B}_{S,V})$ reveals strong correlation between these parameters.)
3. S-V flipping is most degenerate with respect to input to the P population. (The marginal distributions of $q_\theta(h_P \mid \mathcal{B}_{S,V})$ and $q_\theta(dh_P \mid \mathcal{B}_{S,V})$ are approximately uniform along their allowed range.)

Simulation from the posterior reveals that the biologically plausible mode ($h^{(c)} > 0$) matches experiments by producing V-to-S flipping with contrast. Alternatively, the biologically implausible mode ($h^{(c)} > 0$) produces S-to-V flipping with contrast – the opposite effect from experiments (Fig. 1C).

We examined the responses of this circuit model to varying contrasts using the highest log-probability parameterization (z_1) of the plausible mode. Simulations show the stochastic responses of the S- and V-population are indeed flipped at no and full contrast (Fig. 1D). This flipping occurs gradually with graded increases in contrast (Fig. 1E). Interestingly, the responses are noisier at intermediate contrasts (Fig. 1F). This suggests a new hypothesis explaining phenomena of variability in V1: noise is greater in stabilized supralinear networks during winner-take-all competition between the S- and V- populations, but is quenched as either the S- or V-population overtakes the other.

2 V1 Supplemental section

The dynamics are driven by the rectified and exponentiated sum of recurrent inputs Wx, external inputs h, and external noise $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$:

$$T\frac{dx}{dt} = -x + x_{\text{rest}} + Wf_r(x) + b + h + \epsilon \tag{4}$$

where T is a diagonal matrix with $\tau_E = 20$ms and $\tau_P = \tau_S = \tau_V = 10$ms, $x_{\text{rest}} = -65$mV, and $f_r(x) = [x - x_{\text{rest}}]_+^2$.

The noise is an modeled as an Ornstein-Uhlenbeck process:

$$\tau_{\text{noise}} d\epsilon_\alpha = -\epsilon_\alpha dt + \sqrt{2\tau_{\text{noise}}}\sigma_\alpha dB \tag{5}$$

where τ_{noise} is slow at 50ms, $\sigma_E = 0.2\text{mV}$, $\sigma_P = \sigma_S = \sigma_V = 0.1\text{mV}$, and dB is a standard Wiener process.

We considered fixed effective connectivity weights W approximated from experimental recordings of publicly available datasets of mouse V1 [14, 15] (see Section 2). The input to this circuit is comprised of a nominal baseline input b and additive contrast-dependent input h

$$h = c \begin{bmatrix} h^{(c)} \\ h^{(c)} \\ 0 \\ 0 \end{bmatrix}. \tag{6}$$

We considered fixed effective connectivity weights W approximated from experimental recordings of publicly available datasets of mouse V1. Specifically, Billeh et al. [15] produce estimates of the synaptic strength (in mV)

$$M = \begin{bmatrix} 0.36 & -0.48 & -0.31 & -0.28 \\ 1.49 & -0.68 & -0.50 & -0.18 \\ 0.86 & -0.42 & -0.15 & -0.32 \\ 1.31 & -0.41 & -0.52 & -0.37 \end{bmatrix} \tag{7}$$

and connection probability

$$C = \begin{bmatrix} 0.16 & 0.411 & 0.424 & 0.087 \\ 0.395 & .451 & 0.857 & 0.02 \\ 0.182 & 0.03 & 0.082 & 0.625 \\ 0.105 & 0.22 & 0.77 & 0.028 \end{bmatrix}. \tag{8}$$

Multiplying these connection probabilities and synaptic efficacies gives us an effective connectivity matrix:

$$W = C \odot M = \begin{bmatrix} 0.0576 & -0.197 & -0.131 & -0.0244 \\ 0.589 & -0.307 & -0.429 & -0.00360 \\ 0.157 & -0.0126 & -0.0123 & -0.200 \\ 0.138 & -0.0902 & -0.400 & -0.0104 \end{bmatrix}. \tag{9}$$

References

[1] Misha V Tsodyks, William E Skaggs, Terrence J Sejnowski, and Bruce L McNaughton. Paradoxical effects of external modulation of inhibitory interneurons. *Journal of neuroscience*, 17(11):4382–4388, 1997.

[2] Mark S Goldman. Memory without feedback in a neural network. *Neuron*, 61(4):621–634, 2009.

[3] Brendan K Murphy and Kenneth D Miller. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*, 61(4):635–648, 2009.

[4] Hirofumi Ozeki, Ian M Finn, Evan S Schaffer, Kenneth D Miller, and David Ferster. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–592, 2009.

[5] Daniel B Rubin, Stephen D Van Hooser, and Kenneth D Miller. The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, 85(2):402–417, 2015.

[6] Henry Markram, Maria Toledo-Rodriguez, Yun Wang, Anirudh Gupta, Gilad Silberberg, and Caizhi Wu. Interneurons of the neocortical inhibitory system. *Nature reviews neuroscience*, 5(10):793, 2004.

[7] Bernardo Rudy, Gordon Fishell, SooHyun Lee, and Jens Hjerling-Leffler. Three groups of interneurons account for nearly 100% of neocortical gabaergic neurons. *Developmental neurobiology*, 71(1):45–61, 2011.

[8] Robin Tremblay, Soohyun Lee, and Bernardo Rudy. GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron*, 91(2):260–292, 2016.

[9] Carsten K Pfeffer, Mingshan Xue, Miao He, Z Josh Huang, and Massimo Scanziani. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature neuroscience*, 16(8):1068, 2013.

[10] Ashok Litwin-Kumar, Robert Rosenbaum, and Brent Doiron. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. *Journal of neurophysiology*, 115(3):1399–1409, 2016.

[11] Luis Carlos Garcia Del Molino, Guangyu Robert Yang, Jorge F. Mejias, and Xiao Jing Wang. Paradoxical response reversal of top- down modulation in cortical circuits with three interneuron types. *Elife*, 6:1–15, 2017.

[12] Guang Chen, Carl Van Vreeswijk, David Hansel, and David Hansel. Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation. 2019.

[13] Guillaume Hennequin, Yashar Ahmadian, Daniel B Rubin, Máté Lengyel, and Kenneth D Miller. The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability. *Neuron*, 98(4):846–860, 2018.

[14] (2018) Allen Institute for Brain Science. Layer 4 model of v1. available from: https://portal.brain-map.org/explore/models/l4-mv1.

[15] Yazan N Billeh, Binghuang Cai, Sergey L Gratiy, Kael Dai, Ramakrishnan Iyer, Nathan W Gouwens, Reza Abbasi-Asl, Xiaoxuan Jia, Joshua H Siegle, Shawn R Olsen, et al. Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *bioRxiv*, page 662189, 2019.