

# Input/output properties of neuron-type population model of V1

Sean Bittner

May 12, 2020

## 1 Introduction

In section 3.3 of our bioRxiv preprint [1], we analyze the input/output properties of a 4 neuron-type model of V1. We identify degeneracies and sensitivities of neuron-type responses w.r.t specific inputs by visual inspection of the EPI distributions. In my opinion, these “novel insights” are not particularly interesting. Furthermore, these insights are easily obtained using an existing method – approximate Bayesian computation (ABC).

In this write-up, I present a new line of inquiry in the same V1 circuit model as the paper. Rather than conditioning on levels of increase of each neuron-type, we condition on the stabilization of each neuron-type. EPI then presents us with a distribution of stimulation to this V1 circuit that does not change much a particular neuron-type response.

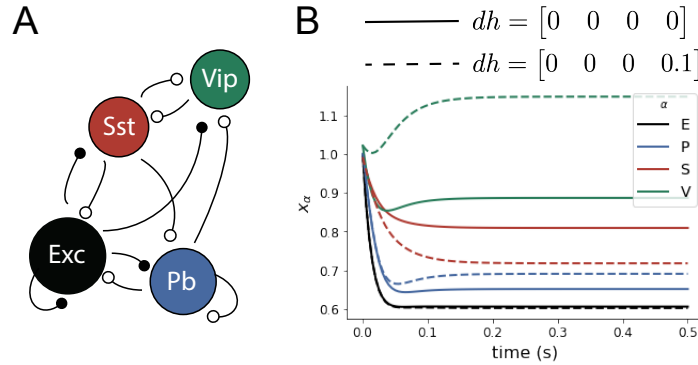
First, I show EPI converges to a distribution covering the set of samples produced by ABC. This important sanity check, along with some ABC bias visualization, will make nice appendix figures. Then, I show how EPI readily identifies the most efficacious patterns of stimulation. Finally, we confirm two distinct mechanisms of E-population stabilization using EPI on the V1 model with inactivated populations.

## 2 V1 model (same as bioRxiv post)

We consider a four-dimensional circuit model with dynamical state given by the firing rate  $x$  of each neuron-type population  $x = [x_E, x_P, x_S, x_V]^T$  (Fig. 1A). Given a time constant of  $\tau = 20$  ms and a power  $n = 2$ , the dynamics are driven by the rectified and exponentiated sum of recurrent ( $Wx$ ) and external  $h$  inputs:

$$\tau \frac{dx}{dt} = -x + [Wx + h]_+^n. \quad (1)$$

**Figure 1: V1 model**



We considered fixed effective connectivity weights  $W$  approximated from experimental recordings of publicly available datasets of mouse V1 [2, 3]. The input  $h = b + dh$  is comprised of a baseline

input  $b = [b_E, b_P, b_S, b_V]^\top = [1, 1, 1, 1.25]^\top$  and a differential input  $dh = [dh_E, dh_P, dh_S, dh_V]^\top$  to each neuron-type population.

We want to know the differential inputs  $dh$  that maintain the steady state  $x_\alpha$  for  $\alpha \in \{E, P, S, V\}$ . We see from Figure 1B that input to a single population in the recurrent circuit elicits a variety of responses across populations: E same, P up, S down, and V up. We define the differential steady state  $dx_\alpha$  as the change in steady state  $x_\alpha$  when receiving input  $h = b + dh$  with respect to the baseline  $h = b$ . Maintaining the steady state of a neuron-type population amounts to the emergent property

$$\mathcal{B}(\alpha, \sigma) \triangleq \mathbb{E} \begin{bmatrix} dx_\alpha \\ dx_\alpha^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma^2 \end{bmatrix}. \quad (2)$$

In the following analyses, we chose  $\sigma = 0.25$ .

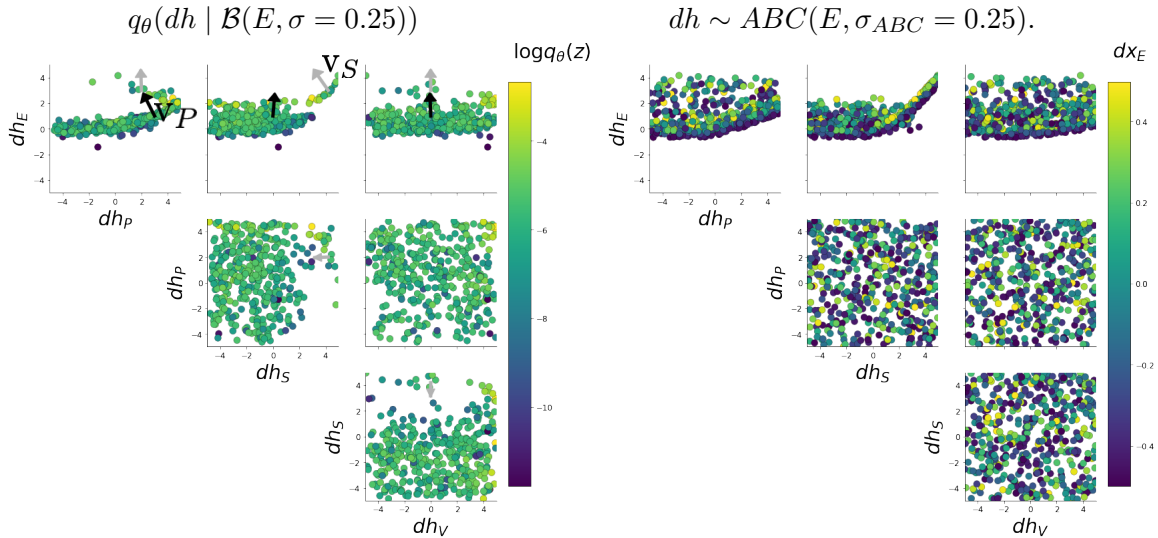
### 3 EPI agrees with ABC

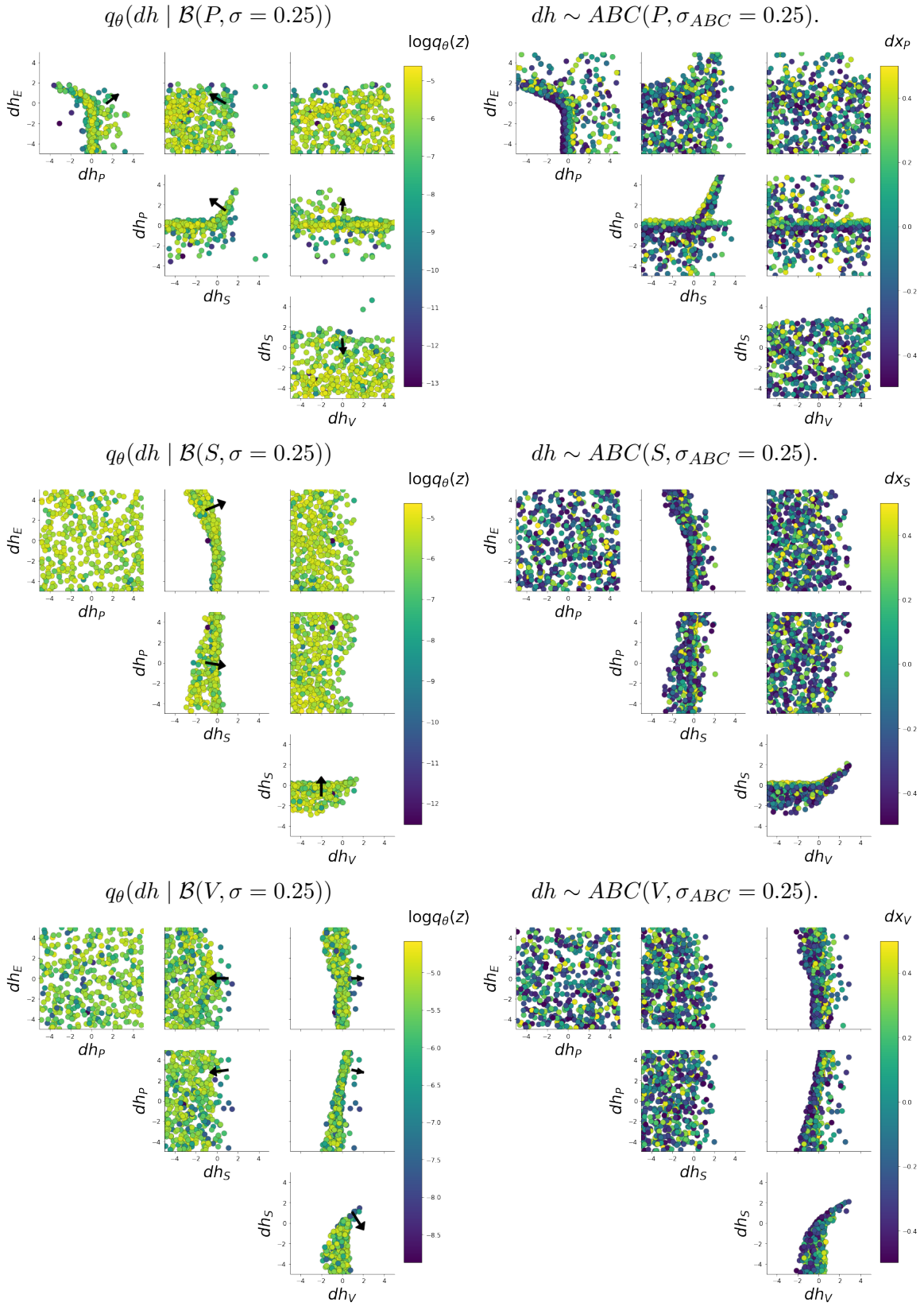
To get an idea of what distribution of parameters ( $dh$ ) we should expect from EPI, we can use ABC to obtain a set of parameters related to the emergent property. We compare EPI to ABC with a rejection heuristic defined by the standard deviation of the differential responses  $\sigma_{ABC}$

$$f_{ABC}(dx_\alpha; \sigma_{ABC}) = |dx_\alpha| > 2\sigma_{ABC}.$$

In other words, we ran ABC accepting parameters that generate differential responses within two standard deviations  $\sigma_{ABC} = 0.25$  of  $dx_\alpha = 0$ . In Figure 2, we see that the distributions obtained via EPI (colored by  $\log q_\theta(z)$ ) are visually similar to those obtained via ABC (colored by  $dx_\alpha$ ).

**Figure 2:** EPI (left) vs ABC (right). Arrows in EPI distribution indicate dimensions of maximal sensitivity at selected parameters  $dh$ . The importance of  $v_S$  and  $v_P$  are explained in section 4.2.





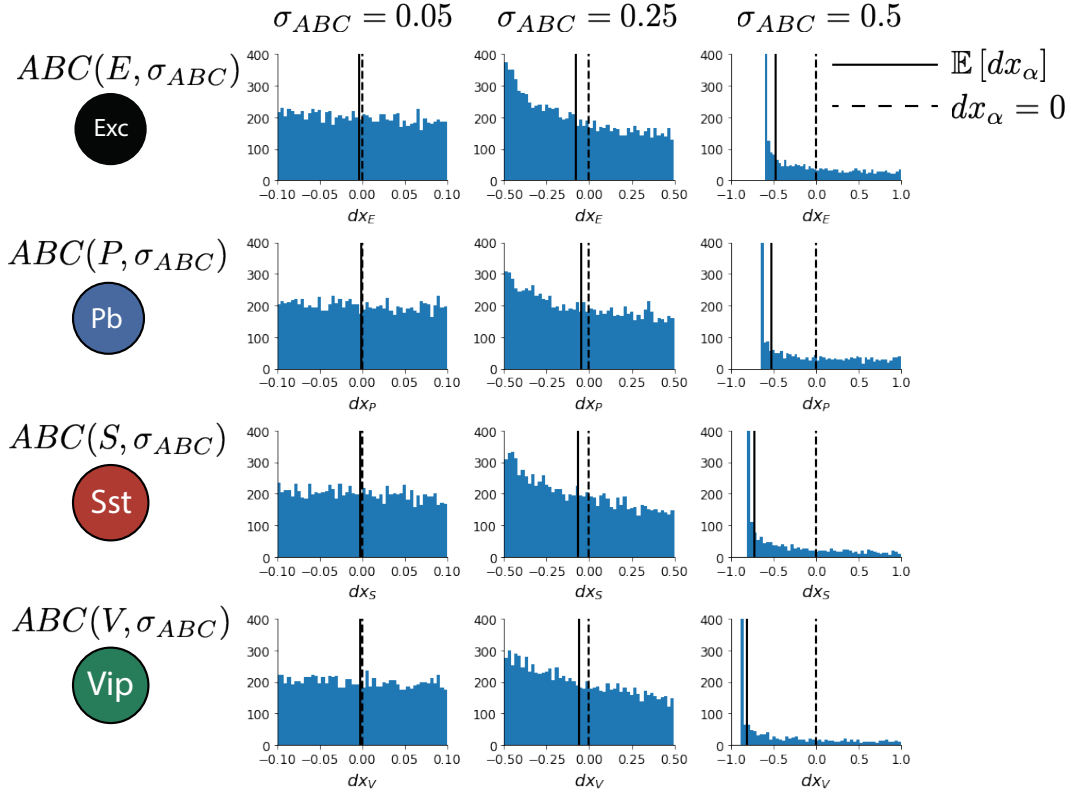
There are some subtle differences between the EPI and ABC distributions, but some difference is to be expected. Rather than simply attributing these differences to imperfection of the EPI

optimization routine, we consider the the effect of bias in  $dx_\alpha$  from ABC (and lack thereof from EPI).

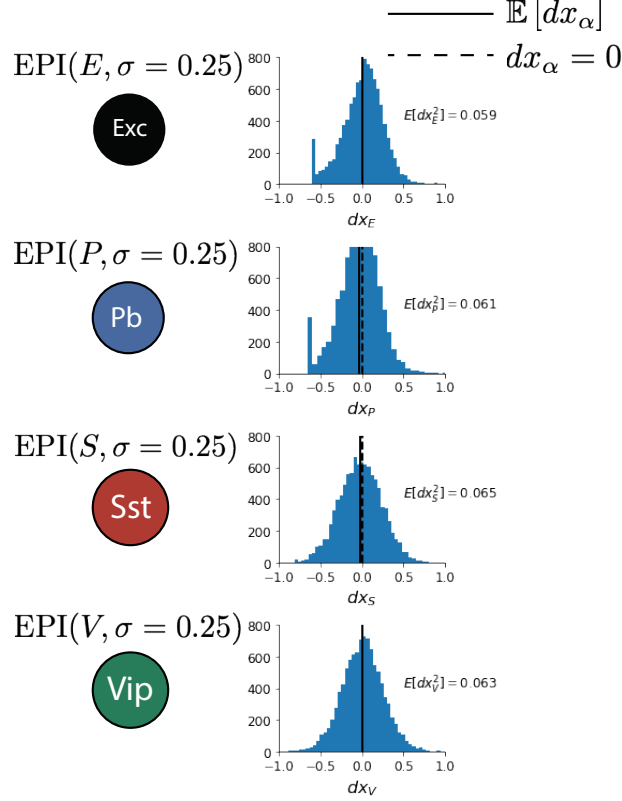
When  $\epsilon > 0$  in ABC, the samples are *not* from the posterior distribution. For  $\epsilon > 0$ , the “posterior” predictive means of ABC  $\mathbb{E}_{ABC(\alpha, \sigma_{ABC})}[dx_\alpha]$  may be far from zero. In Figure 3, we see that with ABC, there is an increasingly negative bias in  $\mathbb{E}_{ABC(\alpha, \sigma_{ABC})}[dx_\alpha]$  for greater error tolerances across all neuron-types. Additionally with ABC, there is no precise control of the variance  $\mathbb{E}_{ABC(\alpha, \sigma_{ABC})}[dx_\alpha^2]$ , which may be undesirable.

In contrast, the first and second moments of  $dx_\alpha$  are controlled to a specified degree of accuracy (Figure 4). The variances of each distribution of  $dx_\alpha$  are close to their target value  $0.25^2 = 0.0625$  (see variances next to histograms of Figure 4). The regions of the ABC pairplots which have  $|dx_\alpha| \approx 0.5$  are far less probable in the EPI distributions (by design), and can account for most mismatch observed. As a note, there is a discontinuity in Figure 2,  $EPI(E, \sigma = 0.25)$  in the  $dh_E$ - $dh_S$  plane. I’d like to see if I can run the optimization for more random seeds so that we obtain a more connected distribution.

**Figure 3: Biased ABC**



**Figure 4: Unbiased EPI**



## 4 EPI-enabled analyses

### 4.1 Dimensions of sensitivity

Each arrow in the EPI distributions (Fig. 2, left) points in the direction of maximal sensitivity indicated by the hessian  $\partial^2 \log q_\theta(dh) / \partial dh \partial dh^\top$  evaluated at the  $dh$  at the base of the arrow. These patterns of input elicit a greater  $dx_\alpha^2$  than (at least) 98.8% of random stimulation patterns. Each sensitivity vector is representative of a different stabilization mechanism.

### 4.2 Confirming distinct mechanisms of stabilization

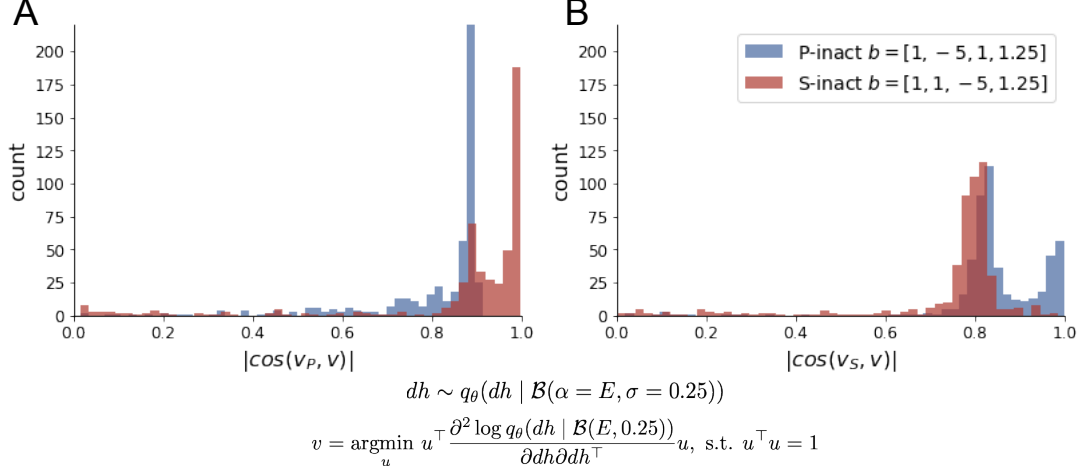
The sensitivity arrows on the EPI distribution for E-population stability (Fig. 2, top, left) indicate two distinct stability mechanisms for increased  $dh_E$ . In one case,  $dh_P$  must increase along with increased  $dh_E$  (Fig. 2, top, left black arrow  $v_P$ ). In the other case,  $dh_S$  must increase along with increased  $dh_E$  (Fig. 2, top, left gray arrow  $v_S$ ). These inferred distributions and sensitivity vectors  $v_P$  and  $v_S$  motivate the following hypothesis:

H: If the P- and S-populations are inactivated, the only E-stabilization mechanism with increasing  $dh_E$  will be that of the other population (S- or P-population, respectively) .

We can test this hypothesis with EPI, since it is optimized to identify the distribution of all inputs resulting in E-population stability. We ran EPI with the same emergent property as Eq 2 ( $\alpha = E$ ,  $\sigma = 0.25$ ) for two different inactivation scenarios. In the P-inactivated case, we changed the baseline input to  $b = b_P \triangleq [1, -5, 1, 1.25]^\top$ , and to  $b = b_S \triangleq [1, 1, -5, 1.25]^\top$  in the S-inactivated case. To

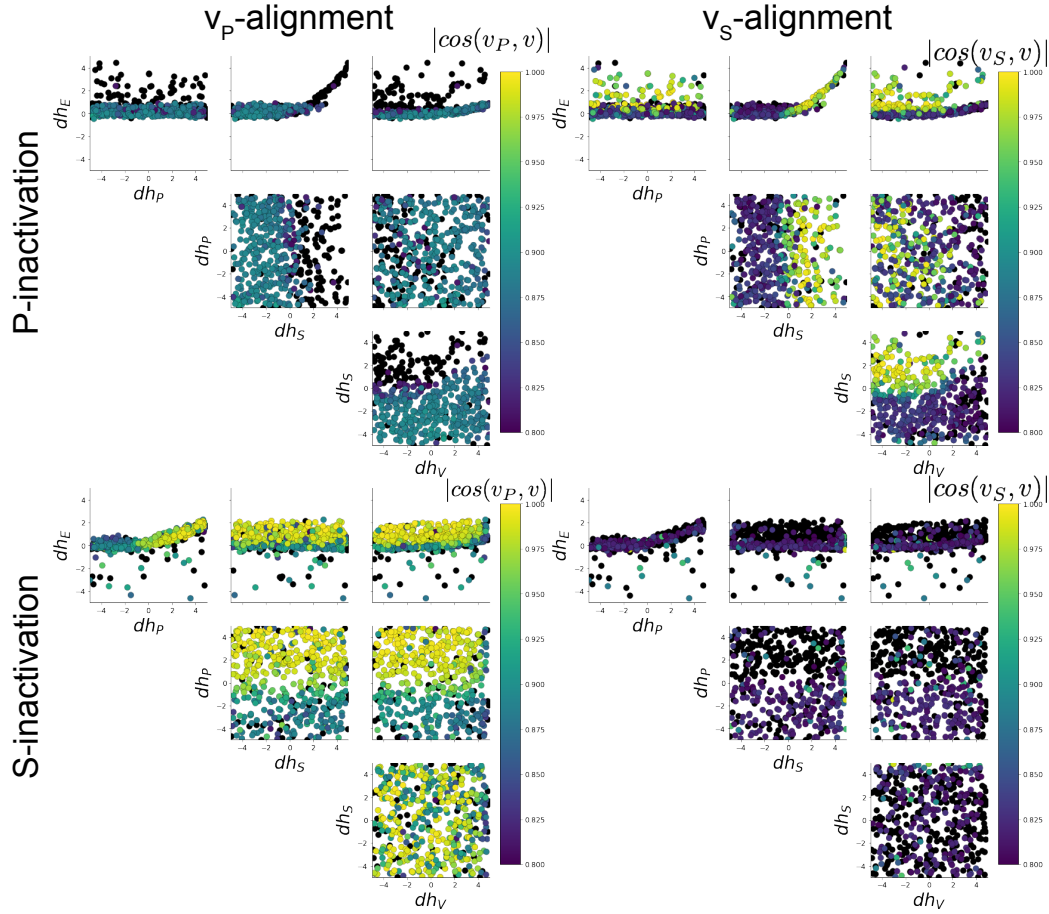
test for the existence of stabilization mechanisms corresponding to  $v_P$  and  $v_S$ , we measured the most-sensitive dimension of the Hessian at all points in the inferred distribution. We declare that the P- or S-inactivated circuit lacks a stability mechanism, if no sensitivity vectors throughout the distribution are aligned with the corresponding sensitivity vector.

**Figure 5:** Stability mechanisms of inactivated circuits.



Confirming our hypothesis, we found that only the  $v_S$  stability mechanism is present in the P-inactivated circuit (Figure 5A), and that only the  $v_P$  stability mechanism is present in the S-inactivated circuit (Figure 5B). As expected, the samples of the inferred distributions aligned with the stability mechanism correspond to the sloped regions orthogonal to  $v_S$  and  $v_P$  (Figure 6).

**Figure 6:** Inactivated circuit stability distributions.



## 5 Summary

We have improved on our previous EPI analysis of V1 in the following ways:

1. We utilize a novel ability of EPI (Hessian evaluation).
2. Our scientific conclusion is about mechanisms of stability rather than high-level characteristics of responses.
3. We have validation of structure via ABC.
4. We characterize the issue of bias in ABC (and lack thereof in EPI).
5. The emergent property of “neuron-type stabilization” follows the ethos of the definition of emergent property better than “neuron-type population rate increase of  $y$ ”.

I recommend substituting this work for section 3.3 in the bioRxiv manuscript. I’d be happy to write up a draft of such a section.

## References

- [1] Sean R Bittner, Agostina Palmigiano, Alex T Piet, Chunyu A Duan, Carlos D Brody, Kenneth D Miller, and John P Cunningham. Interrogating theoretical models of neural computation with deep inference. *bioRxiv*, page 837567, 2019.
- [2] (2018) Allen Institute for Brain Science. Layer 4 model of v1. available from: <https://portal.brain-map.org/explore/models/l4-mv1>.
- [3] 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.