PNAS COMMENTARY

Ordering in heterogeneous connectome weights for visual information processing

Taro Toyoizumi[a,b,1]

What defines ourselves? People believe that neuronal circuits in the brain characterize who we are and how we behave. Following the triumph of the Human Genome Project, connectomics projects started as the next big biology initiative to map out the wiring diagrams between neurons in the brain. Since the completion of the first connectome of a tiny worm, *Caenorhabditis elegans*, in 1986 (1), large-scaled connectome data have been accumulated from nanoscale to macroscale across different animal species (2). However, understanding these data remains challenging (3). The reasons include the difficulty in identifying important parameters in high-dimensional observations, data variability across experimental conditions, and relating the anatomical structure to brain function. These problems also exist if we focus on the connectome data from the primary visual cortex, V1, of mice (4–6). The connection statistics differ depending on the source and target cell types across cortical layers. Even within the same classification type, some statistics deviate significantly across datasets. Thus, finding general rules is challenging. Finally, it is nontrivial to understand how the anatomical circuit is related to visual information processing. In PNAS, Kraynyukova et al. (7) conducted model-based analyses that address these problems.

Kraynyukova et al. studied the so-called stabilized supralinear network (SSN) model (8). It consists of (Fig. 1, *Left*) the excitatory, *E*, and inhibitory, *I*, neural populations in V1 that receive input from the visual thalamus (dorsal lateral geniculate nucleus, i.e., dLGN). Each population is characterized by its firing rate r_X (X is either E or I), which monotonically and supralinearly increases with its synaptic input, given in the form of excitation minus inhibition. The model has six connection weights: two thalamocortical weights, g_E and g_I, to and four cortical weights, J_{EE}, J_{EI}, J_{IE}, and J_{II}, between the two cortical populations. In addition, the model has a power exponent n that describes the nonlinearity of the two cortical populations. Kraynyukova et al. performed in vivo electrophysiological recordings of mice V1 and dLGN neurons in response to the presentation of visual grating stimuli with various contrast levels and orientations to fit this SSN model. They quantified, for each of the dLGN and V1 putative excitatory and inhibitory populations, the population response by averaging the corresponding neurons' two-dimensional receptive field, i.e., the firing rate profile, after

Author affiliations: [a]Laboratory for Neural Computation and Adaptation, RIKEN Center for Brain Science, Wako, Saitama 351-0106, Japan; and [b]Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan

Author contributions: T.T. wrote the paper.

The author declares no competing interest.

See companion article, "In vivo extracellular recordings of thalamic and cortical visual responses reveal V1 connectivity rules," 10.1073/pnas.2207032119.

[1]Email: taro.toyoizumi@riken.jp.

Published November 21, 2022.

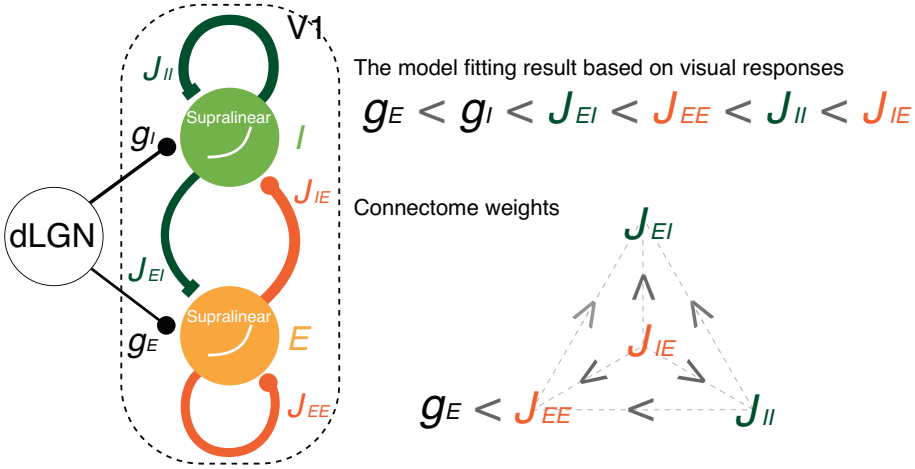

Fig. 1. (*Left*) A diagram of the SSN model of V1 (Kraynyukova et al., 2022). The cortical excitatory population (*E*: orange circle) and inhibitory population (*I*: light green circle) receive input from dLGN with weights g_E and g_I, respectively. The four types of cortical weights connecting the two populations are denoted as J_{EE}, J_{EI}, J_{IE}, and J_{II}, where J_{EI} describes the weights from *I* to *E*, for example. The line width of the thalamocortical and corticocortical connections is proportional to the mean estimated weight based on the in vivo data. In the steady state, the firing rate r_X of population X (X is either E or I) is given by $r_X = [J_{XE} r_E - J_{XI} r_I + g_X T]_+^n$, where T is the dLGN firing rate, the operation $[\cdot]_+$ transforms a negative argument to zero so that the firing rate is positive, and n is a nonlinear exponent. (*Right*) Kraynyukova et al. fitted the SSN model to the in vivo data. They found specific ordering among them (*Top*). This ordering was verified using three in vitro connectome datasets (*Bottom*). The transparency level of the inequality signs reflects the consistency among the datasets. In addition, they got $n \sim 1.9$, which indicates that the input–output function is supralinear.

aligning the neurons' preferred orientations. The orientation tuning in these populations was contrast invariant to good accuracy (i.e., the shape of orientation tuning did not change with the contrast), which warrants the examination of their population contrast response separately from their orientation tuning. Given the dLGN population contrast response, they fit the seven parameters of the SSN model to reproduce the V1 population contrast response. This analysis consistently gives a specific ordering of inferred weights (Fig. 1, *Right*). Remarkably, this ordering is consistent with the connectome datasets (4–6) despite the variability in individual weights.

The nonlinear exponent n of the SSN model also plays a crucial role. When a linear input–output function is used (i.e., $n = 1$) instead, the inferred weights follow dissimilar ordering from the connectome data. This result argues that the simple SSN model involving only the two major cortical populations is suitable for relating contrast response in V1 and the connectome data. Theoretically, the supralinear input–output function increases the cortical firing rate progressively greater than the firing rate of the thalamus at higher contrast. Consequently, the network is mainly external-input-driven at low contrast and cortical-input-driven at high contrast. The network is said to operate as an inhibition-stabilized network (ISN) (9–11) in the latter case if the network activity diverges (at least until another stabilization mechanism kicks in) under the blockade of inhibitory activity due to its strong recurrent excitation. ISNs have a salient feature, a.k.a., the paradoxical effect, where an increase in the input to the inhibitory population transiently increases its activity but subsequently reduces both the excitatory and inhibitory activity. This prediction was experimentally verified (12) across cortical areas. While alternative mechanisms, such as the disinhibition mechanism that involves more than two populations, are not excluded (11, 13, 14), the ISN regime is a simple and attractive mechanism that reproduces the normalization and surround suppression properties commonly observed in the cortex (15). Kraynyukova et al. showed that their fitting yields ISN at >4% contrast level, suggesting that ISN properties should contribute to a variety of visual stimuli.

> **"The finding of Kraynyukova et al. is a major step toward the understanding of connectome variability in identifying general rules underlying network functions."**

The findings of Kraynyukova et al. are attractive because it shows a robust ordering of connectome weights associated with the visual contrast response even though individual weights are highly variable, arguing for the importance of supralinear contrast response in V1. But how can we understand the variability in connectome weights further? Possible reasons for the variability include a bias and a finite number of samples in each dataset. Unbiased and infinite-size data are elusive. Moreover, the distribution of individual synaptic strength is often heavy-tailed (16, 17), hampering the convergence of the empirical mean with the sample size. The variability may also reflect the animal's experience and developmental trace (18) or transient changes in the circuit properties like representational drift (19) and intrinsic synaptic dynamics (17). Another candidate for the variability is the scaling property of the SSN model. In the SSN model, for any positive real number a, multiplying all cortical weights by factor $a^{1/n-1}$ and all thalamocortical weights by factor $a^{1/n}$ simply changes both the cortical excitatory and inhibitory firing rates by the factor a (c.f., the steady-state equation in Fig. 1 caption). This transformation corresponds to the invariant subspace of the SSN parameters. More generally, there are diverse solutions (i.e., configurations of biological elements) that can achieve a certain network function. This diversity has been suggested to be important for the resilience of network functions and recovery from failures (20). The solutions to achieving a certain network function could be clustered or scattered in the network parameter space. Understanding the distribution of solutions is also important in machine learning as it describes if a solution is easily found (21). Analyses of deep networks suggest that a learning path may arbitrarily follow one of many equally good paths involving multiple branching points or follow a highly restricted path if structural pruning is applied a priori. Interestingly, given that structural pruning is successful, learning is faster and achieves higher generalization capability than without pruning (22). Thus, the connectome variability around good enough solutions may be due to arbitrary fluctuations, highly structured individual differences, or their mixture. The finding of Kraynyukova et al. is a major step toward the understanding of connectome variability in identifying general rules underlying network functions.

ACKNOWLEDGMENTS. I thank Ho Ka Chan and Louis Kang for their help in proofreading the manuscript. The study was supported by RIKEN Center for Brain Science, Brain/MINDS from AMED under Grant JP15dm0207001, and JSPS KAKENHI Grant JP18H05432.

1. J. G. White, E. Southgate, J. N. Thomson, S. Brenner, The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **314**, 1–340 (1986).
2. J. L. Lanciego, F. G. Wouterlood, A half century of experimental neuroanatomical tracing. *J. Chem. Neuroanat.* **42**, 157–183 (2011).
3. Y. Frégnac, Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain? *Science* **358**, 470–477 (2017).
4. Allen Institute for Brain Science, Synaptic physiology coarse matrix dataset. https://brain-map.org/explore/connectivity/synaptic-physiology. Accessed 23 September 2022.
5. X. Jiang *et al.*, Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).
6. S. B. Hofer *et al.*, Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat. Neurosci.* **14**, 1045–1052 (2011).
7. N. Kraynyukova *et al.*, In vivo extracellular recordings of thalamic and cortical visual responses reveal V1 connectivity rules. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2207032119 (2022).
8. Y. Ahmadian, D. B. Rubin, K. D. Miller, Analysis of the stabilized supralinear network. *Neural Comput.* **25**, 1994–2037 (2013).
9. H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, D. Ferster, Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62**, 578–592 (2009).
10. Y. Ahmadian, K. D. Miller, What is the dynamical regime of cerebral cortex? *Neuron* **109**, 3373–3391 (2021).
11. S. Sadeh, C. Clopath, Inhibitory stabilization and cortical computation. *Nat. Rev. Neurosci.* **22**, 21–37 (2021).
12. A. Sanzeni *et al.*, Inhibition stabilization is a widespread property of cortical networks. *Elife* **9**, e54875 (2020).
13. A. Litwin-Kumar, R. Rosenbaum, B. Doiron, Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. *J. Neurophysiol.* **115**, 1399–1409 (2016).
14. A. Mahrach, G. Chen, N. Li, C. van Vreeswijk, D. Hansel, Mechanisms underlying the response of mouse cortical networks to optogenetic manipulation. *Elife.* **9**, e49967 (2020).
15. D. B. Rubin, S. D. Van Hooser, K. D. Miller, The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron* **85**, 402–417 (2015).
16. L. Cossell *et al.*, Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399–403 (2015).
17. H. Kasai, N. E. Ziv, H. Okazaki, S. Yagishita, T. Toyoizumi, Spine dynamics in the brain, mental disorders and artificial neural networks. *Nat. Rev. Neurosci.* **22**, 407–422 (2021).

18. B. A. Linkenhoker, C. G. von der Ohe, E. I. Knudsen, Anatomical traces of juvenile learning in the auditory system of adult barn owls. *Nat. Neurosci.* **8**, 93–98 (2005).
19. M. E. Rule, T. O'Leary, C. D. Harvey, Causes and consequences of representational drift. *Curr. Opin. Neurobiol.* **58**, 141–147 (2019).
20. J. M. Goaillard, E. Marder, Ion channel degeneracy, variability, and covariation in neuron and circuit resilience. *Ann. Rev. Neurosci.* **44**, 335–57 (2021).
21. H. Li, Z. Xu, G. Taylor, C. Studer, T. Goldstein, Visualizing the loss landscape of neural nets. *Adv. Neural Inf. Proc. Sys.* **31**, (2018).
22. J. Frankle, M. Carbin, The lottery ticket hypothesis: Finding sparse, trainable neural networks. *arXiv [Preprint]* (2018). https://doi.org/10.48550/arXiv.1803.03635. Accessed 23 September 2022.