

# Conceptual and technical advances define a key moment for theoretical neuroscience

Anne K Churchland & L F Abbott

**Theoretical approaches have long shaped neuroscience, but current needs for theory are elevated and prospects for advancement are bright. Advances in measuring and manipulating neurons demand new models and analyses to guide interpretation. Advances in theoretical neuroscience offer new insights into how signals evolve across areas and new approaches for connecting population activity with behavior. These advances point to a global understanding of brain function based on a hybrid of diverse approaches.**

Theoretical approaches have a long history of contributing to neuroscience research, but never before has the need for them been so high nor the prospects for advancement so great. The explosion in technologies available for measuring and manipulating neurons has created a call for analysis techniques that are scalable to extremely large data sets, that take into account the heterogeneity of neurons, and that can predict and interpret the effects of complex manipulations of activity. The growing importance of theory is also driven by developments in theoretical neuroscience itself, advances that expand the reach of theoretical approaches and extend their ability to offer insight into long-standing puzzles. In the coming years, we will obtain enormous quantities of behavioral, recording (both electrical and optical), connectomic, gene expression and other forms of data. Obtaining deep understanding from this onslaught will require, in addition to the skillful and creative application of experimental technologies, substantial advances in data analysis methods and intense application of theoretic concepts and models.

The path from data to understanding requires three stages of analysis (**Fig. 1**). First, extracting relevant signals from the raw data: for example, image stabilization and isolation of regions-of-interest for imaging. Second, reducing large and

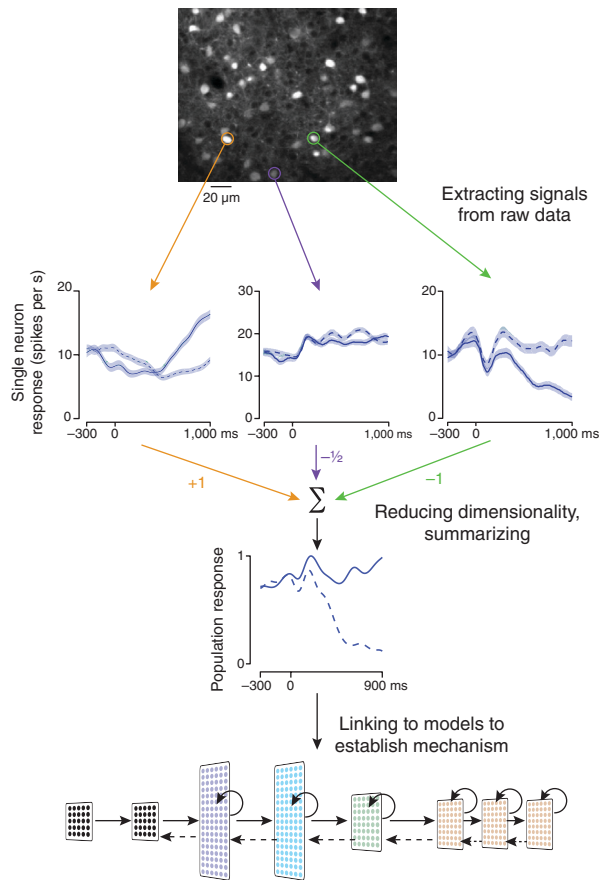
complex data into forms that summarize the results in a more compact and understandable way. Third, using modeling to link the results to underlying mechanisms and overlying principles. The perspectives and reviews in this issue primarily address the third, and some of the second, stage, surveying new developments and modeling-based insights in topics ranging from understanding and interpreting network spiking activity<sup>1–3</sup>, exploring visual processing<sup>4</sup> and memory<sup>5</sup>, and studying the representation and computation of probability<sup>6</sup>, to investigations of higher level cognition and mental illness<sup>7</sup>. In addition, major advances in the other stages of analysis have driven the entire program to evolve considerably in recent years. Two such advances and their accompanying challenges are described below.

First, there are new challenges in data analysis driven by advances in our ability to simultaneously measure responses from many neurons. Specifically, it is not clear how to reduce large and complex data sets into forms that are understandable. Simply averaging responses of many neurons could obscure important signals: neural populations often have massive diversity in cell type<sup>8</sup>, projection target<sup>9</sup> and response property<sup>10</sup>. One solution to this problem is to leverage methods that are naturally suited to multi-neuron measurements. Correlations across neurons, for example, can offer insight into the connectivity and state of a network<sup>1</sup>. Another solution to this problem is the use of machine-learning-based readouts and classifiers to interpret activity at the population level and relate it to behavior. Classifiers work by determining how best to combine neurons to extract differences in population activity across conditions (**Fig. 1**).

This approach differs critically from traditional averaging in that neurons contribute in varying degrees, or even negatively (**Fig. 1**), to the extracted signal. Simple classifiers are becoming a standard tool for interpreting neural data<sup>10,11</sup>. More complex classifiers, such as those with multiple processing layers, are rapidly being developed in both neuroscience and industry<sup>12</sup>. One example is the hierarchical convolutional neural network (HCNN), which is based on ideas developed in studies of the visual system. As discussed in this issue, HCNNs can be used to interpret activity at various stages along the primate visual pathway during object recognition<sup>4</sup>.

Second, new challenges for data analysis arise from advances in experimentalists' ability to manipulate neural activity with temporal, pathway and cell-type precision. The results of such manipulations can be difficult to interpret. For example, the neuronal heterogeneity described above implies that a population of neurons may encode a parameter through a mixture of firing rate increases and decreases. Thus, a pan-neuronal firing rate change from optogenetic stimulation or suppression could have mixed effects. This could occur even for manipulations of an identified class of cells because recent experimental findings indicate heterogeneity far beyond the currently identified categories<sup>13,14</sup>. Furthermore, as with more traditional methods of stimulation, propagation of activity to neighboring areas can create 'off target' effects that may erroneously suggest that an area has a causal role when, in fact, it does not<sup>15</sup>. Circuit-level modeling is a basic theoretical tool for dealing with these issues, but it must be extended to the multi-area level to address them more fully<sup>16</sup>.

Anne K. Churchland is at the Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA. L.F. Abbott is at the Department of Neuroscience, and Department of Physiology and Cellular Biophysics, Columbia University College of Physicians and Surgeons, New York, New York, USA. e-mail: churchland@cshl.edu



**Figure 1** Multiple stages of analysis needed for neural data to inform our understanding of neural mechanism. Top, extracting raw signals from data. Image shows neurons in mouse parietal cortex measured using a two-photon microscope (courtesy of F. Najafi). Regions of interest corresponding to single neurons are circled. Arrows point to schematized firing rates from the neurons, grouped by behavioral condition. Middle, dimensionality reduction is used to summarize responses of a large population. In the example, a weight is assigned to each neuron, and contributions are summed to extract a population response that correlates with behavioral conditions of interest. Bottom, models then link this reduced dimensionality representation of the data to mechanism. Adapted from ref. 4.

Beyond the work described in this issue, what form might future models take? Work on the worm *C. elegans* provides insights that we expect will apply to other systems as well. In a study involving large-scale imaging of the *C. elegans* nervous system, a principal component-based description of neural activity was extracted, and patterns of activity across the entire neuronal population were correlated with behavioral states of the worm<sup>17</sup>. This analysis may ultimately lead to a model that provides an accurate, high-level description of neural activity and accounts for important aspects of worm behavior. The ‘units’ of this type of model are projections

across a large neuronal population, not neurons themselves. Such a model could be criticized for leaving unexplained the responses of individual neurons; indeed, there is a temptation to require such models to dig down and explain fine-grained phenomena as well. As an alternative, we suggest a hybrid, rather than hierarchical, approach in which high-level and fine-grained models coexist, each explaining different features of the data. For example, in *C. elegans* the activity of a small number of modulatory neurons can alter the state transitions predicted by the high level model<sup>18</sup>, but these neurons are unlikely to be described well at the population level. The best

approach may be to combine high-level and fine-grained components into a hybrid model that provides a more complete account of nervous system function than either component could by itself.

The perspectives and reviews in this issue embody this diversity of approach. Some of the articles in this issue start with basic, although simplified, biophysics and try to derive conclusions by assembling well-understood elements into less well-understood networks<sup>1–3,5</sup>. Others start with a general principle or idea of how a computation should be performed and work down from this ‘normative’ starting point to infer how functionality is achieved<sup>6,7</sup>. Yet another approach is to draw on analogies with another system, such as an HCNN<sup>4</sup>, even if it is quite different at the mechanistic level. Given the range of phenomena studied in neuroscience, demanding unity of viewpoint, methodology or depth of understanding, even in a single system, seems both unrealistic and counterproductive. Global understanding, when it comes, will likely take the form of highly diverse panels loosely stitched together into a patchwork quilt.

#### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

- Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K. & Josić, K. *Nat. Neurosci.* **19**, 383–393 (2016).
- Denève, S. & Machens, C.K. *Nat. Neurosci.* **19**, 375–382 (2016).
- Abbott, L.F., DePasquale, B. & Memmesheimer, R.-M. *Nat. Neurosci.* **19**, 350–355 (2016).
- Yamins, D.L.K. & Dicarlo, J.J. *Nat. Neurosci.* **19**, 356–365 (2016).
- Fiete, I. & Chaudhuri, R. *Nat. Neurosci.* **19**, 394–403 (2016).
- Pouget, A., Drugowitsch, J. & Kepecs, A. *Nat. Neurosci.* **19**, 366–374 (2016).
- Huys, Q.J.M., Maia, T.V. & Frank, M.J. *Nat. Neurosci.* **19**, 404–413 (2016).
- Huang, Z.J. & Zeng, H. *Annu. Rev. Neurosci.* **36**, 183–215 (2013).
- Li, N., Chen, T.W., Guo, Z.V., Gerfen, C.R. & Svoboda, K. *Nature* **519**, 51–56 (2015).
- Raposo, D., Kaufman, M.T. & Churchland, A.K. *Nat. Neurosci.* **17**, 1784–1792 (2014).
- Pagan, M., Urban, L.S., Wohl, M.P. & Rust, N.C. *Nat. Neurosci.* **16**, 1132–1139 (2013).
- LeCun, Y., Bengio, Y. & Hinton, G. *Nature* **521**, 436–444 (2015).
- Tasic, B. *et al. Nat. Neurosci.* **19**, 335–346 (2016).
- Cembrowski, M.S. *et al. Neuron* doi:10.1016/j.neuron.2015.12.013 (13 January 2016).
- Otchy, T.M. *et al. Nature* **528**, 358–363 (2015).
- Chaudhuri, R., Bernacchia, A. & Wang, X.J. *eLife* **3**, e01239 (2014).
- Kato, S. *et al. Cell* **163**, 656–669 (2015).
- Flavell, S.W. *et al. Cell* **154**, 1023–1035 (2013).