On simplicity and complexity in the brave new world of large-scale neuroscience

Peiran Gao[1] and Surya Ganguli[2]

Technological advances have dramatically expanded our ability to probe multi-neuronal dynamics and connectivity in the brain. However, our ability to extract a simple conceptual understanding from complex data is increasingly hampered by the lack of theoretically principled data analytic procedures, as well as theoretical frameworks for how circuit connectivity and dynamics can conspire to generate emergent behavioral and cognitive functions. We review and outline potential avenues for progress, including new theories of high dimensional data analysis, the need to analyze complex artificial networks, and methods for analyzing entire spaces of circuit models, rather than one model at a time. Such interplay between experiments, data analysis and theory will be indispensable in catalyzing conceptual advances in the age of large-scale neuroscience.

Addresses
[1] Department of Bioengineering, Stanford University, Stanford, CA 94305, United States
[2] Department of Applied Physics, Stanford University, Stanford, CA 94305, United States

Corresponding author: Gao, Peiran (prgao@stanford.edu)

'*Things should be as simple as possible, but not simpler.*'
– Albert Einstein.

Introduction

Experimental neuroscience is entering a golden age marked by the advent of remarkable new methods enabling us to record ever increasing numbers of neurons [1–5,6•], and measure brain connectivity at various levels of resolution [7–10,11•,12–14], sometimes measuring both connectivity and dynamics in the same set of neurons [15•,16]. This recent thrust of technology development is spurred by the hope that an understanding of how the brain gives rise to sensations, actions and thoughts will lurk within the resulting brave new world of complex large-scale data sets. However, the question of how one can extract a conceptual understanding from data remains a significant challenge for our field. Major issues involve: (1) What does it even mean to conceptually understand 'how the brain works?' (2) Are we collecting the right kinds and amounts of data to derive such understanding? (3) Even if we could collect *any* kind of detailed measurements about neural structure and function, what theoretical and data analytic procedures would we use to extract conceptual understanding from such measurements? These are profound questions to which we do not have crisp, detailed answers. Here we merely present potential routes towards the beginnings of progress on these fronts.

Understanding as a journey from complexity to simplicity

First, the vague question of 'how the brain works' can be meaningfully reduced to the more precise, and proximally answerable question of how do the connectivity and dynamics of distributed neural circuits give rise to specific behaviors and computations? But what would a satisfactory answer to this question look like? A detailed, predictive circuit model down to the level of ion-channels and synaptic vesicles within individual neurons, while remarkable, may not yield conceptual understanding in any meaningful human sense. For example, if simulating this detailed circuit were the *only* way we could predict behavior, then we would be loath to say that we *understand* how behavior emerges from the brain.

Instead, a good benchmark for understanding can be drawn from the physical sciences. Feynman articulated the idea that we understand a physical theory if we can say something about the solutions to the underlying equations of the theory without actually solving those equations. For example, we understand aspects of fluid mechanics because we can say many things about specific fluid flows, without having to numerically solve the Navier–Stokes equations in every single case. Similarly, in neuroscience, understanding will be found when we have the ability to develop simple coarse-grained models, or better yet a hierarchy of models, at varying levels of biophysical detail, all capable of predicting salient aspects of behavior at varying levels of resolution. In traversing this hierarchy, we will obtain an invaluable understanding of which biophysical details matter, and more importantly, which do not, for any given behavior. Thus our goal should be to find simplicity amidst complexity, while of course keeping in mind Einstein's famous dictum quoted above.

How many neurons are enough: simplicity and complexity in multineuronal dynamics

What kinds and amounts of data are required to arrive at simple but accurate coarse grained models? In the world of large scale recordings, where we do not have access to simultaneous connectivity information, the focus has been on obtaining a state-space description of the dynamics of neural circuits through various dimensionality reduction methods (see [17] for a review). This body of work raises a key conceptual issue permeating much of systems neuroscience, namely, what precisely can we infer about neural circuit dynamics and its relation to cognition and behavior while measuring only an infinitesimal fraction of behaviorally relevant neurons? For example, given a doubling time of about 7.4 years [18] in the number of neurons we can simultaneously measure at single cell, single spike-time resolution, we would have to wait more than 100 years before we can observe $O(10^6-10^9)$ neurons typically present in full mammalian circuits controlling complex behaviors [19]. Thus, systems neuroscience will remain for the foreseeable future within the vastly undersampled measurement regime, so we need a *theory* of neuronal data analysis in this regime. Such theory is essential for firstly guiding the biological interpretation of complex multivariate data analytic techniques, secondly efficiently designing future large scale recording experiments, and finally developing theoretically principled data analysis algorithms appropriate for the degree of subsampling.

A clue to the beginnings of this theory lies in an almost universal result occurring across many experiments in which neuroscientists tightly control behavior, record many trials, and obtain trial averaged neuronal firing rate data from hundreds of neurons: in such experiments, the dimensionality (i.e. number of principal components required to explain a fixed percentage variance) of neural data turns out to be much less than the number of recorded neurons (Figure 1). Moreover, when dimensionality reduction procedures are used to extract neuronal state dynamics, the resulting low dimensional neural trajectories yield a remarkably insightful dynamical portrait of circuit computation (e.g. [20,21,22•]).

These results raise several profound and timely questions: what is the origin of the underlying simplicity implied by the low dimensionality of neuronal recordings? How can we trust the dynamical portraits that we extract from so few neurons? Would the dimensionality increase if we recorded more neurons? Would the portraits change? Without an adequate theory, it is impossible to quantitatively answer, or even precisely formulate, these important questions. We have recently started to develop such a theory [41,42]. Central to this theory is the mathematically well-defined notion of neuronal task complexity (NTC). Intuitively, the NTC measures the volume of the manifold of task parameters (see Figure 2a for

Figure 1
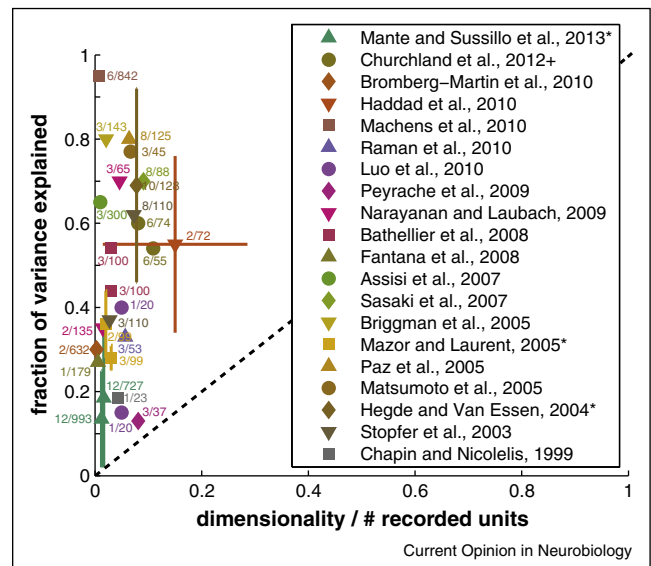
In many experiments (e.g. in insect [20,23–26] olfactory systems, mammalian olfactory [26,27], prefrontal [21,22•,28–30], motor and premotor,[31,32], somatosensory [33], visual [34,35], hippocampal [36], and brain stem [37] systems) a *much* smaller number of dimensions than the number of recorded neurons captures a large amount of variance in neural firing rates.

the special cases of simple reaches) measured in units of the neuronal population autocorrelation scale across each task parameter. Thus the NTC in essence measures how many neuronal activity patterns could possibly appear during the course of an experiment given that task parameters have a limited extent and neuronal activity patterns vary smoothly across task parameters (Figure 2b). With the mathematical definition of the NTC in hand, we derive that the dimensionality of neuronal data is upper bounded by the NTC, and if the neural data manifold is sufficiently randomly oriented, we can accurately recover dynamical portraits when the number of observed neurons is proportional to the log of the NTC (Figure 2c).

These theorems have significant implications for the interpretation and design of large-scale experiments. First, it is likely that in a wide variety of experiments, the origin of low dimensionality is due to a small NTC, a hypothesis that we have verified in recordings from the motor and premotor cortices of monkeys performing a simple 8 direction reach task [43]. In any such scenario, simply increasing the number of recorded neurons, without a concomitant increase in task complexity will not lead to richer, higher dimensional datasets — indeed data dimensionality will be independent of the number of recorded neurons. Moreover, we confirmed in motor cortical data our theoretically predicted result that the number of recorded neurons should be proportional to the

Figure 2
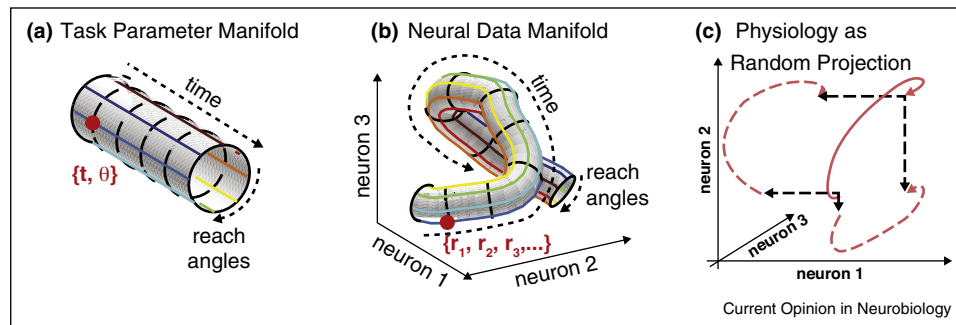
(a) For a monkey reaching to different directions, the trial averaged behavioral states visited by the arm throughout the experiment are parameterized by a cylinder with two coordinates, reach angle θ, and time into the reach t. **(b)** Trial averaged neural data is an embedding of the task manifold into firing rate space. The number of dimensions explored by the neural data manifold is limited by its volume and its curvature (but not the total number of neurons in the motor cortex), with smoother embeddings exploring fewer dimensions. The NTC is a mathematically precise upper bound on the number of dimensions of the neural data manifold given the volume of the task parameter manifold and a smoothness constraint on the embedding. **(c)** If the neural data manifold is low dimensional and randomly oriented w.r.t. single neuron axes, then its shadow onto a subset of recorded neurons will preserve its geometric structure. We have shown, using random projection theory [38,39,40•] that to preserve neural data manifold geometries with fractional error ϵ, one needs to record $M \geq (1/\epsilon)K \log(NTC)$ neurons. The figure illustrates a $K = 1$ dimensional neural manifold in $N = 3$ neurons, and we only record $M = 2$ neurons. Thus, fortunately, the intrinsic complexity of the neural data manifold (small), not the number of neurons in the circuit (large) determines how many neurons we need to record.

logarithm of the NTC to accurately recover dynamical portraits of neural state trajectories. This is excellent news: while we must make tasks more complex to obtain richer, more insightful datasets, we need not record from many more neurons within a brain region to accurately recover its internal state-space dynamics.

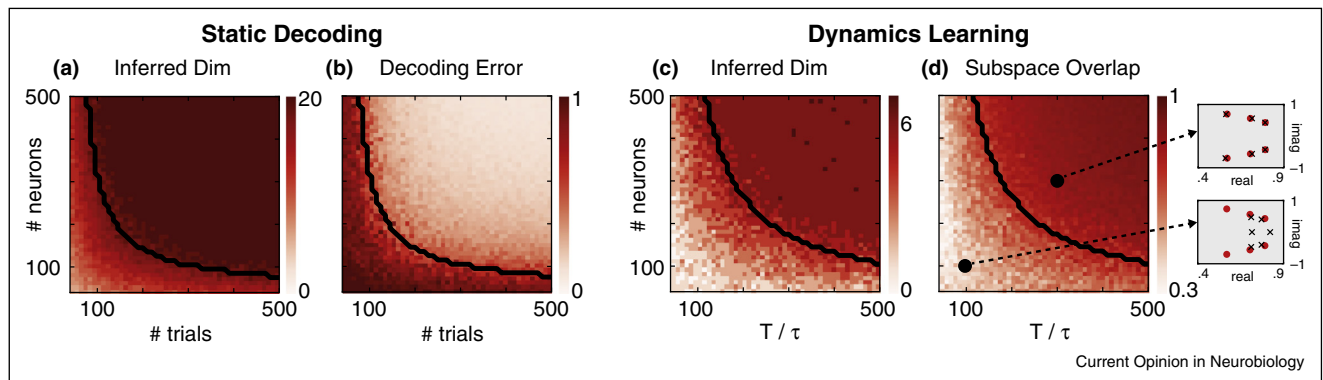## Towards a theory of single trial data analysis

The above work suggests that the proximal route for progress lies not in recording more neurons alone, but in designing more complex tasks and stimuli. However, with such increased complexity, the same behavioral state or stimulus may rarely be revisited, precluding the possibility of trial averaging as a method for data analysis. Therefore it is essential to extend our theory to the case of single trial analysis. A simple formulation of the problem is as follows: suppose we have a K dimensional manifold of behavioral states (or stimuli), where K is not necessarily known, and the animal explores P states in succession. The behavior is controlled by a circuit with N neurons but we measure only M of them. Furthermore, each neuron is noisy with a finite SNR, reflecting single trial variability. For what values of M, N, P, K, and the SNR can we accurately estimate the dimensionality K of neural data, and accurately decode behavior on single trials? We have solved this problem analytically in the case in which noisy neural activity patterns reflecting P discrete stimuli lie near a K dimensional subspace (Figure 3a,b). We find, roughly that the relations $M, P > K$ and $\text{SNR}\sqrt{MP} > K$ are sufficient [42]. Thus, it is an intrinsic measure of neural complexity K, and not the total number of neurons N, that sets a lower bound on how many neurons M and stimuli P we must observe at a given SNR for accurate single trial

analyses. Moreover, we have generalized this analysis to learning dynamical systems (Figure 3c,d).

Both our preliminary analyses reveal the existence of phase transitions in performance as a function of the number of recorded neurons, and the amount of recording time, stimuli, or behavioral states. Only on the correct side of the phase boundary are accurate dimensionality, dynamics estimation and single trial decoding possible. Such phase transitions are often found in many high dimensional data analysis problems [49•], for example in compressed sensing [50,51] and matrix recovery [52]. They reveal the beautiful fact that we can recover a lot of information about large objects (vectors or matrices) using surprisingly few measurements when we have seemingly weak prior information, like sparsity, or low-rank structure (see [53,54] for reviews in a neuroscience context). Moreover, in Figure 3, we see that we can move along the phase boundary by trading off number of recorded neurons with recording time.

Thus, to guide future large-scale experimental design, it will be exceedingly important to determine the position of these phase boundaries under increasingly realistic biological assumptions, for example exploring the roles of spiking variability, noise correlations, sparsity, cell types, and network connectivity constraints, and how they impact our ability to uncover true network dynamics and single trial decodes in the face of subsampling. In essence, we need to develop a Rosetta stone connecting biophysical network dynamics to statistics. This dictionary will teach us when and how the learned parameters of statistical models fit to a subset of recorded neurons

Figure 3

Static Decoding   **Dynamics Learning**

(a) Inferred Dim   (b) Decoding Error   (c) Inferred Dim   (d) Subspace Overlap

Current Opinion in Neurobiology

(a,b) The inferred dimensionality and held-out single trial decoding error as a function of P simulated training examples (single trials) and M recorded neurons in a situation where stimuli are encoded in a $K = 20$ dimensional subspace in a network of $N = 5000$ neurons, with SNR=5. Inference was performed using low rank matrix denoising [44], and our new analysis of this algorithm reveals a sufficient condition for accurate inference, $\mathrm{SNR}\sqrt{\frac{M}{P}}(\sqrt{P} - \sqrt{K})^2\left(\sqrt{\frac{N-K}{N}} - \sqrt{\frac{K}{M}}\sqrt{\frac{N-M}{N}}\right)^2 \geq K$ [42]. The black curve in (a,b) reflects the theoretically predicted phase boundary in the P, M plane separating accurate from inaccurate inference. This expression simplifies in the experimentally relevant regime $K, M \ll N$ and $K \ll M, P$ to $\mathrm{SNR}\sqrt{MP} > K$. **(c,d)** Learning the dimensionality and dynamics, via subspace identification [45] of a linear neural network of size $N = 5000$ from spontaneous noise driven activity. The low-rank connectivity of the network forces the system to lie in a $K = 6$ dimensional subspace. Performance is measured as a function of the number of recorded neurons M and recording time T. By combining and extending time series random matrix theory [46•], low-rank perturbation theory [47], and noncommutative probability theory, [48], we have derived a theoretically predicted phase boundary (black curve in (c,d)), that matches simulations. In (d), left, the subspace overlap is the correlation between the inferred subspace and the true subspace, with 1 being perfect correlation, or overlap. In (d), on the right, dynamics (eigenvalues) are correctly predicted only on the right side of the boundary (red dots are true eigenvalues, blue crosses are estimated eigenvalues).

ultimately encode the collective dynamics of the much larger, unobserved neural circuit containing them — an absolutely fundamental question in neuroscience.

Understanding complex networks with complete information

As we increasingly obtain information about both the connectivity and dynamics of neural circuits, we have to ask ourselves how should we use this information? As a way to sharpen our ideas, it can be useful to engage in a thought experiment in which experimental neuroscience eventually achieves complete success, in enabling us to measure detailed connectivity, dynamics and plasticity in full neural sub-circuits during behavior. How then would we extract understanding from such rich data? Moreover, could we arrive at this same understanding without collecting all the data, perhaps even only collecting data within reach in the near future? To address this thought experiment, it is useful to turn to advances in computer science, where deep or recurrent neural networks, consisting of multiple layers of cascaded nonlinearities, have made a resurgence as the method of choice for solving a range of difficult computational problems. Indeed, deep learning (see [55–57] for reviews) has led to advances in object detection [58•,59], face recognition [60,61], speech recognition [62], language translation [63], genomics [64], microscopy [65], and even modeling biological neural responses [66,67•,68,69]. Each of these networks can solve a complex computational problem. Moreover, we

know the full network connectivity, the dynamics of every single neuron, the plasticity rule used to train the network, and indeed the entire developmental experience of the network, in terms of its exposure to training stimuli. Virtually any experiment we wish to do on these networks, we can do. Yet a meaningful understanding of how these networks work still eludes us, as well has what a suitable benchmark for such understanding would be. Following Feynman's guideline for understanding physical theories, can we say something about the behavior of deep or recurrent artificial neural networks *without* actually simulating them in detail? More importantly, could we arriving at these statements of understanding without measuring every detail of the network, and what are the minimal set of measurements we would need? We do not believe that understanding these networks will directly inform us about how much more complex biological neural networks operate. However, even in an artificial setting, directly confronting the question of what it means to understand how complex distributed circuits compute, and what kinds of experiments and data analytic procedures we would need to arrive at this understanding, could have a dramatic impact on the questions we ask, experiments we design, and the data analysis we do, in the pursuit of this same understanding in biological neural circuits.

The *theory* of deep learning is still in its infancy, but some examples of general statements one can make about deep

circuits without actually simulating them include how their functional complexity scales with depth [70,71], how their synaptic weights, over time, acquire statistical structures in inputs [72,73], and how their learning dynamics is dominated by saddle points, not local minima [72,74]. However, much more work at the intersection of experimental and theoretical neuroscience and machine learning will be required before we can address the intriguing thought experiment of what we should do if we could measure anything we wanted.

Understanding not a single model, but the space of all possible models

An even higher level of understanding is achieved when we develop not just a single model that explains a data set, but rather understand the space of all possible models consistent with the data. Such an understanding can place existing biological systems within their evolutionary context, leading to insights about why they are structured the way they are, and can reveal general principles that transcend any particular model. Inspiring examples for neuroscientists can be found not only within neuroscience, but also in allied fields. For example [75] derived a single Boolean network model of the yeast cell-cycle control network, while [76] developed methods to count and sample from the space of all networks that realize the yeast cell-cycle. This revealed an astronomical number of possible networks consistent with the data, but only 3% of these networks were more robust than the one chosen by nature, revealing potential evolutionary pressure towards robustness. In protein folding, theoretical work [77•] analyzed, in toy models, the space of all possible amino acid sequences that give rise to a given fold; the number of such sequences is the designability of the fold. Theory revealed that typical folds with shapes similar to those occurring in nature are highly designable, and therefore more easily found by evolution. Moreover, designable folds are thermodynamically stable [78] and atypical in shape [79], revealing general principles relating sequence to structure. In the realm of short-term sequence memory, the idea of liquid state machines [80,81] posited that generic neural circuits could convert temporal information into instantaneous spatial patterns of activity, but theoretical work [82–84] revealed general principles relating circuit connectivity to memory, highlighting the role of non-normal and orthogonal network connectivities in achieving robust sequence memory. In the realm of long-term memory, seminal work revealed that it is essential to treat synapses as entire dynamical systems in their own right, exhibiting a particular synaptic model [85], while further theoretical work [86] analyzed the space of all possible synaptic dynamical systems, revealing general principles relating synaptic structure to function. Furthermore conductance based models of central pattern generators revealed that highly disparate conductance levels can yield similar behavior [87], suggesting that observed correlations in conductance levels across

animals [88] could reflect a signature of homeostatic design [89].

These examples all show that studying the space of models consistent with a given data set or behavior can greatly expand our conceptual understanding. Further work along these lines within the context of neuronal networks is likely to yield important insights. For example, suppose we could understand the space of all possible deep or recurrent neural networks that solve a given computational task. Which observable aspects of the connectivity and dynamics are universal across this space, and which are highly variable across individual networks? Are the former observables the ones we should focus on measuring in real biological circuits solving the same task? Are the latter observables less relevant and more indicative of historical accidents over the time course of learning?

In summary, there are great challenges and opportunities for generating advances in high dimensional data analysis and neuronal circuit theory that can aid in not only responding to the need to interpret existing complex data, but also in driving the questions we ask, and the design of large-scale experiments we do to answer these questions. Such advances in theory and data analysis will be required to transport us from the 'brave new world, that has such [complex technology] in't' [90•] and deliver us to the promised land of conceptual understanding.

Conflict of interest statement
Nothing declared.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest

1. Stevenson IH, Kording KP: **How advances in neural recording affect data analysis**. *Nat Neurosci* 2011, **14**:139-142.

2. Robinson JT, Jorgolli M, Shalek AK, Yoon MH, Gertner RS, Park H: **Vertical nanowire electrode arrays as a scalable platform for intracellular interfacing to neuronal circuits**. *Nat Nano* 2012, **7**:180-184.

3. Ahrens MB, Li JM, Orger MB, Robson DN, Schier AF, Engert F, Portugues R: **Brain-wide neuronal dynamics during motor adaptation in zebrafish**. *Nature* 2012, **485**:471-477.

4. Schrodel T, Prevedel R, Aumayr K, Zimmer M, Vaziri A: **Brain-wide 3D imaging of neuronal activity in *Caenorhabditis elegans* with sculpted light**. *Nat Methods* 2013, **10**:1013-1020.

5. Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, El Gamal A, Schnitzer MJ: **Long-term dynamics of CA1 hippocampal place codes**. *Nat Neurosci* 2013, **16**:264-266.

6. Prevedel R, Yoon Y, Hoffmann M, Pak N: **Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy**. *Nat Methods* 2014, **11**:727-730.
● Achieves high speed 3D volumetric calcium imaging in *C. elegans* and larval zebrafish by measuring the entire light field, consisting of both the location and angle of incident photons.

7. Micheva K, Smith S: **Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits**. *Neuron* 2007, **55**:25-36.

8. Wickersham IR, Lyon DC, Barnard RJ, Mori T, Finke S, Conzelmann KK, Young JA, Callaway EM: **Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons**. *Neuron* 2007, **53**:639-647.

9. Li A, Gong H, Zhang B, Wang Q, Yan C, Wu J, Liu Q, Zeng S, Luo Q: **Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain**. *Science* 2010, **330**:1404-1408.

10. Ragan T, Kadiri LR, Venkataraju KU, Bahlmann K, Sutin J, Taranda J, Arganda-Carreras I, Kim Y, Seung HS, Osten P: **Serial two-photon tomography for automated ex vivo mouse brain imaging**. *Nat Methods* 2012, **9**:255-258.

11. Chung K, Deisseroth K: **Clarity for mapping the nervous system**. *Nat Methods* 2013, **10**:508-513.
● Obtains global brain connectivity information through optical microscopy in brains that have been made transparent through the removal of lipids.

12. Takemura S, Bharioke y, Lu A, Nern Z, Vitaladevuni A, Rivlin S: **A visual motion detection circuit suggested by *Drosophila* connectomics**. *Nature* 2013, **500**:175-181.

13. Pestilli F, Yeatman JD, Rokem A, Kay KN, Wandell BA: **Evaluation and statistical inference for human connectomes**. *Nat Methods* 2014, **11**:1058-1063.

14. Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S *et al.*: **A mesoscale connectome of the mouse brain**. *Nature* 2014, **508**:207-214.

15. Bock DD, Lee WCA, Kerlin AM, Andermann ML, Hood G, Wetzel AW, Yurgenson S, Soucy ER, Kim HS, Reid RC: **Network anatomy and in vivo physiology of visual cortical neurons**. *Nature* 2011, **471**:177-182.
● Obtains simultaneous functional and anatomical information about the same set of neurons by combining optical imaging with EM microscopy.

16. Rancz EA, Franks KM, Schwarz MK, Pichler B, Schaefer AT, Margrie TW: **Transfection via whole-cell recording in vivo: bridging single-cell physiology, genetics and connectomics**. *Nat Neurosci* 2011, **14**:527-532.

17. Cunningham JP, Byron MY: **Dimensionality reduction for large-scale neural recordings**. *Nat Neurosci* 2014:1500-1509.

18. Stevenson IH, Kording KP: **How advances in neural recording affect data analysis**. *Nat Neurosci* 2011, **14**:139-142.

19. Shepherd GM: **The Synaptic Organization of the Brain**. New York: Oxford University Press; 2004.

20. Mazor O, Laurent G: **Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons**. *Neuron* 2005, **48**:661-673.

21. Machens CK, Romo R, Brody CD: **Functional, but not anatomical, separation of ''what'' and ''when'' in prefrontal cortex**. *J Neurosci* 2010, **30**:350-360.

22. Mante V, Sussillo D, Shenoy KV, Newsome WT: **Context-dependent computation by recurrent dynamics in prefrontal cortex**. *Nature* 2013, **503**:78-84.
● Elucidates a distributed mechanism for contextual gating in decision making by training artificial recurrent networks to solve the same task a monkey does, and finds that the artificial neurons behave like the monkeys prefrontal neurons.

23. Stopfer M, Jayaraman V, Laurent G: **Intensity versus identity coding in an olfactory system**. *Neuron* 2003, **39**:991-991004.

24. Assisi C, Stopfer M, Laurent G, Bazhenov M: **Adaptive regulation of sparseness by feedforward inhibition**. *Nat Neurosci* 2007, **10**:1176-1184.

25. Raman B, Joseph J, Tang J, Stopfer M: **Temporally diverse firing patterns in olfactory receptor neurons underlie spatiotemporal neural codes for odors**. *J Neurosci* 2010, **30**:1994-2006.

26. Haddad R, Weiss T, Khan R, Nadler B, Mandairon N, Bensafi M, Schneidman E, Sobel N: **Global features of neural activity in the olfactory system form a parallel code that predicts olfactory behavior and perception**. *J Neurosci* 2010, **30**:9017-9026.

27. Bathellier B, Buhl DL, Accolla R, Carleton A: **Dynamic ensemble odor coding in the mammalian olfactory bulb: sensory information at different timescales**. *Neuron* 2008, **57**:586-598.

28. Narayanan NS, Laubach M: **Delay activity in rodent frontal cortex during a simple reaction time task**. *J Neurophysiol* 2009, **101**:2859-2871.

29. Peyrache A, Khamassi M, Benchenane K, Wiener SI, Battaglia FP: **Replay of rule-learning related neural patterns in the prefrontal cortex during sleep**. *Nat Neurosci* 2009, **12**:919-926.

30. Warden MR, Miller EK: **Task-dependent changes in short-term memory in the prefrontal cortex**. *J Neurosci* 2010, **30**:15801-15810.

31. Paz R, Natan C, Boraud T, Bergman H, Vaadia E: **Emerging patterns of neuronal responses in supplementary and primary motor areas during sensorimotor adaptation**. *J Neurosci* 2005, **25**:10941-10951.

32. Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV: **Neural population dynamics during reaching**. *Nature* 2012, **487**:51-56.

33. Chapin J, Nicolelis M: **Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations**. *J Neurosci* 1999, **94**:121-140.

34. Hegdé J, Van Essen DC: **Temporal dynamics of shape analysis in macaque visual area V2**. *J Neurophysiol* 2004, **92**:3030-3042.

35. Matsumoto N, Okada M, Yasuko S, Yamane S, Kawano K: **Population dynamics of face-responsive neurons in the inferior temporal cortex**. *Cereb Cortex* 2005, **15**:1103-1112.

36. Sasaki T, Matsuki N, Ikegaya Y: **Metastability of active CA3 networks**. *J Neurosci* 2007, **27**:517-528.

37. Bromberg-Martin ES, Hikosaka O, Nakamura K: **Coding of task reward value in the dorsal raphe nucleus**. *J Neurosci* 2010, **30**:6262-6272.

38. Johnson W, Lindenstrauss J: **Extensions of Lipschitz mappings into a Hilbert space**. *Contemp Math* 1984, **26** 1-1.

39. Dasgupta S, Gupta A: **An elementary proof of a theorem of johnson and Lindenstrauss**. *Random Struct Algorithms* 2003, **22**:60-65.

40. Baraniuk R, Wakin M: **Random projections of smooth manifolds**. *Found Comput Math* 2009, **9**:51-77.
● Reveals that when data or signals lie on a curved low dimensional manifold embedded in a high dimensional space, remarkably few measurements are required to recover the geometry of the manifold.

41. Gao P, Trautmann E, Yu B, Santhanam G, Ryu S, Shenoy KV, Ganguli S: **A theory of neural dimensionality and measurement**. *Computational and Systems Neuroscience Conference (COSYNE)*. 2014.

42. Gao P, Ganguli S: **Dimensionality, Coding and Dynamics of Single-Trial Neural Data**. *Computational and Systems Neuroscience Conference (COSYNE)*. 2015.

43. Byron MY, Kemere C, Santhanam G, Afshar A, Ryu SI, Meng TH, Sahani M, Shenoy KV: **Mixture of trajectory models for neural decoding of goal-directed movements**. *J Neurophysiol* 2007, **97**:3763-3780.

44. Gavish M, Donoho D: *Optimal Shrinkage of Singular Values*. 2014:. arXiv preprint arXiv:14057511.

45. Lennart L: *System Identification: Theory for the User*. 1999.

46. Yao J: **A note on a marcenko-pasteur type theorem for time-**
• **series**. *Stat Probab Lett* 2012 http://dx.doi.org/10.1016/
j.spl.2011.08.011.
Describes the structure of noise in high dimensional linear dynamical
systems; this structure partially dictates how much of the system, and for
how long, we need to measure to uncover its dynamics.

47. Benaych-Georges F, Nadakuditi R: **The singular values and
vectors of low rank perturbations of large rectangular random
matrices**. *J Multivar Anal* 2012, **111**:120135 http://dx.doi.org/
10.1016/j.jmva.2012.04.019.

48. Nica A, Speicher R: **On the multiplication of free N-tuples of
noncommutative random variables**. *Am J Math* 1996:799-837
http://dx.doi.org/10.2307/25098492 http://www.jstor.org/stable/
25098492.

49. Amelunxen D, Lotz M, McCoy MB, Tropp JA: **Living on the edge:**
• **phase transitions in convex programs with random data**. *Inf
Inference* 2014:iau005.
Reveals that a wide variety of data analysis algorithms that correspond to
convex optimization problems, have phase transitions in various mea-
sures of performance that quantify the success of the algorithm.

50. Donoho D, Maleki A, Montanari A: **Message-passing algorithms
for compressed sensing**. *Proc Natl Acad Sci U S A* 2009,
106:18914.

51. Ganguli S, Sompolinsky H: **Statistical mechanics of
compressed sensing**. *Phys Rev Lett* 2010, **104**:188701.

52. Donoho DL, Gavish M, Montanari A: **The phase transition of
matrix recovery from Gaussian measurements matches the
minimax MSE of matrix denoising**. *Proc Natl Acad Sci U S A*
2013, **110**:8405-8410.

53. Ganguli S, Sompolinsky H: **Compressed sensing, sparsity, and
dimensionality in neuronal information processing and data
analysis**. *Annu Rev Neurosci* 2012, **35**:485-508.

54. Advani M, Lahiri S, Ganguli S: **Statistical mechanics of complex
neural systems and high dimensional data**. *J Stat Mech Theory
Exp* 2013, **2013**:P03014.

55. Bengio Y, Courville A, Vincent P: **Representation learning: a
review and new perspectives**. *IEEE Trans Pattern Anal Mach
Intell* 2013, **35**:1798-1828.

56. Bengio Y: **Learning deep architectures for AI**. *Found Trends®
Mach Learn* 2009, **2**:1-127.

57. Schmidhuber J: **Deep learning in neural networks: an overview**.
Neural Netw 2015, **61**:85-117.

58. Krizhevsky A, Sutskever I, Hinton GE: **Imagenet classification**
• **with deep convolutional neural networks**. *Advances in Neural
Information Processing Systems*. 2012:1097-1105.
Seminal work that yielded substantial performance improvements in
visual object recognition through deep neural networks.

59. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D,
Erhan D, Vanhoucke V, Rabinovich A: *Going Deeper with
Convolutions*. 2014:. arXiv preprint arXiv:14094842.

60. Sun Y, Chen Y, Wang X, Tang X: **Deep learning face
representation by joint identification-verification**. Edited by
Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger
K.*Advances in Neural Information Processing Systems*. Curran
Associates, Inc.; 2014:1988-1996.

61. Taigman Y, Yang M, Ranzato M, Wolf L: **Deepface: closing the
gap to human-level performance in face verification.**.
*2014 IEEE Conference on Computer Vision and Pattern
Recognition (CVPR)*. IEEE; 2014:1701-1708.

62. Hannun AY, Case C, Casper J, Catanzaro BC, Diamos G, Elsen E
et al.: *Deep Speech: Scaling up End-to-End Speech Recognition*.
2014:. CoRR abs/1412.5567.

63. Sutskever I, Vinyals O, Le QVV: **Sequence to sequence learning
with neural networks**. Edited by Ghahramani Z, Welling M,
Cortes C, Lawrence N, Weinberger K.*Advances in Neural
Information Processing Systems*. Curran Associates, Inc.;
2014:3104-3112.

64. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D,
Yuen RK et al.: **The human splicing code reveals new insights
into the genetic determinants of disease**. *Science* 2015,
347:1254806.

65. Ciresan D, Giusti A, Gambardella LM, Schmidhuber J: **Deep
neural networks segment neuronal membranes in electron
microscopy images**. *Advances in Neural Information Processing
Systems*. 2012:2843-2851.

66. Serre T, Oliva A, Poggio T: **A feedforward architecture accounts
for rapid categorization**. *Proc Natl Acad Sci U S A* 2007,
104:6424-6429.

67. Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D,
• DiCarlo JJ: **Performance-optimized hierarchical models
predict neural responses in higher visual cortex**. *Proc Natl Acad
Sci U S A* 2014:201403112.
Deep neural circuits that were optimized for performance in object
categorization contained small collections of neurons in deep layers,
whose linear combinations mimicked the responses of neurons in mon-
key inferotemporal cortex to natural images.

68. Cadieu CF, Hong H, Yamins DL, Pinto N, Ardila D, Solomon EA,
Majaj NJ, Dicarlo JJ: **Deep neural networks rival the
representation of primate IT cortex for core visual object
recognition**. *PLoS comput Biol* 2014, **10**:e1003963.

69. Agrawal P, Stansbury D, Malik J, Gallant JL: *Pixels to Voxels:
Modeling Visual Representation in the Human Brain*. 2014:. arXiv
preprint arXiv:14075104.

70. Bianchini M, Scarselli F: **On the complexity of neural network
classifiers: a comparison between shallow and deep
architectures**. *IEEE Trans Neural Netw* 2014, **25**:1553-1565.

71. Pascanu R, Montúfar G, Bengio Y: **On the number of inference
regions of deep feed forward networks with piece-wise linear
activations**. *Internal Conference on Learning Representations
(ICLR)*. 2014.

72. Saxe A, McClelland J, Ganguli S: **Exact solutions to the
nonlinear dynamics of learning in deep linear neural networks**.
International Conference on Learning Representations (ICLR).
2014.

73. Saxe AM, McClelland JL, Ganguli S: **Learning hierarchical
category structure in deep neural networks.**. *Proc. of 35th
Annual Cog. Sci. Society*. 2013.

74. Dauphin YN, Pascanu R, Gulcehre C, Cho K, Ganguli S, Bengio Y:
**Identifying and attacking the saddle point problem in high-
dimensional non-convex optimization**. *Advances in Neural
Information Processing Systems*. 2014:2933-2941.

75. Li F, Long T, Lu Y, Ouyang Q, Tang C: **The yeast cell-cycle
network is robustly designed**. *Proc Natl Acad Sci U S A* 2004,
101:4781-4786.

76. Lau K, Ganguli S, Tang C: **Function constrains network
architecture and dynamics: a case study on the yeast cell
cycle Boolean network**. *Phys Rev E* 2007, **75**:051907.

77. Li H, Helling R, Tang C, Wingreen N: **Emergence of preferred
• structures in a simple model of protein folding**. *Science* 1996,
273:666-669.
Revealed that entropic arguments alone, involving counting the number
of sequences that lead to a fold, qualitatively explains the presence of
preferred folding structures in nature.

78. Wingreen NS, Li H, Tang C: **Designability and thermal stability of
protein structures**. *Polymer* 2004, **45**:699-705.

79. Li H, Tang C, Wingreen NS: **Are protein folds atypical?** *Proc Natl
Acad Sci U S A* 1998, **95**:4987-4990.

80. Maass W, Natschlager T, Markram H: **Real-time computing
without stable states: a new framework for neural
computation based on perturbations**. *Neural Comput* 2002,
14:2531-2560.

81. Jaeger H: *Short term memory in echo state networks. GMD Report
152 German National Research Center for Information Technology*.
2001.

82. White O, Lee D, Sompolinsky H: **Short-term memory in
orthogonal neural networks**. *Phys Rev Lett* 2004, **92**:148102.

83. Ganguli S, Huh D, Sompolinsky H: **Memory traces in dynamical systems**. *Proc Natl Acad Sci U S A* 2008, **105**:18970.

84. Ganguli S, Sompolinsky H: **Short-term memory in neuronal networks through dynamical compressed sensing**. *Neural Information Processing Systems (NIPS)*. 2010.

85. Fusi S, Drew PJ, Abbott LF: **Cascade models of synaptically stored memories**. *Neuron* 2005, **45**:599-611 http://dx.doi.org/10.1016/j.neuron.2005.02.001.

86. Lahiri S, Ganguli S: **A memory frontier for complex synapses**. *Neural Information Processing Systems (NIPS)*. 2014.

87. Prinz AA, Bucher D, Marder E: **Similar network activity from disparate circuit parameters**. *Nat Neurosci* 2004, **7**:1345-1352.

88. Schulz DJ, Goaillard JM, Marder EE: **Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly variable levels of gene expression**. *Proc Natl Acad Sci U S A* 2007, **104**:13187-13191.

89. O'Leary T, Williams AH, Caplan JS, Marder E: **Correlations in ion channel expression emerge from homeostatic tuning rules**. *Proc Natl Acad Sci U S A* 2013, **110**:E2645-E2654.

90. Shakespeare W: **The Tempest**. Classic Books Company; 2001.
 •
This play is the origin of the phrase 'brave new world,' uttered by the character Miranda in praise and wonder when she saw new people for the first time. However, this phrase is mired in irony, as Miranda unknowingly overestimated the worth of the people she saw.