Bridging large-scale neuronal recordings and large-scale network models using dimensionality reduction

Ryan C. Williamson[1,2,3], Brent Doiron[1,4], Matthew A. Smith[1,5,6], Byron M. Yu[1,7,8]

1: Center for the Neural Basis of Cognition, Pittsburgh, PA, USA.
2: Department of Machine Learning, Carnegie Mellon University, Pittsburgh, PA, USA.
3: School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.
4: Department of Mathematics, University of Pittsburgh, Pittsburgh, PA, USA.
5: Department of Ophthalmology, University of Pittsburgh, Pittsburgh, PA, USA.
6: Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA.
7: Department of Electrical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.
8: Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.

Highlights

- The interplay of neuronal recordings and network models is becoming ever stronger.

- Population activity structure provides common ground for incisive comparisons.

- Dimensionality reduction can be used to identify population activity structure.

- This approach is used to study working memory, decision making, motor control, etc.

1

Abstract ₁

A long-standing goal in neuroscience has been to bring together neuronal recordings and neural   ₂
network modeling to understand brain function. Neuronal recordings can inform the   ₃
development of network models, and network models can in turn provide predictions for   ₄
subsequent experiments. Traditionally, neuronal recordings and network models have been   ₅
related using single-neuron and pairwise spike train statistics. We review here recent studies   ₆
that have begun to relate neuronal recordings and network models based on the   ₇
multi-dimensional structure of neuronal population activity, as identified using dimensionality   ₈
reduction. This approach has been used to study working memory, decision making, motor   ₉
control, and more. Dimensionality reduction has provided common ground for incisive   ₁₀
comparisons and tight interplay between neuronal recordings and network models.   ₁₁

Introduction ₁₂

For decades, the fields of experimental neuroscience and neural network modeling proceeded   ₁₃
largely in parallel. Whereas experimental neuroscience focused on understanding how the   ₁₄
activities of individual neurons relate to sensory stimuli and behavior, the modeling community   ₁₅
sought to understand theoretically how neural networks can give rise to brain function. In   ₁₆
recent years, developments in neuronal recording technology have enabled the simultaneous   ₁₇
recording of hundreds of neurons or more [1]. Concurrently, increases in computational power   ₁₈
have enabled the simulation of large neural networks [2]. Together, these developments should   ₁₉
enable experimental data to more stringently constrain network model design and network   ₂₀
models to better predict neuronal activity for subsequent experiments [3,4].   ₂₁

A key question is how to relate large-scale neuronal recordings with large-scale network   ₂₂
models. Network models typically do not attempt to replicate the precise anatomical   ₂₃
connectivity of the biological network from which the neurons are recorded, since the   ₂₄
underlying anatomical connectivity is usually unknown (although technological developments   ₂₅
are making this possible [5]). In such settings, there is not a one-to-one correspondence of each   ₂₆
recorded neuron with a model neuron. To date, comparisons between recordings and models   ₂₇
have primarily relied on aggregate spike train statistics based on single neurons (e.g.,   ₂₈
distribution of firing rates [6], distribution of tuning preferences [7], and Fano factor [8]) and   ₂₉
pairs of neurons (e.g., spike time [9] and spike count correlations [10,11]), as well as   ₃₀
single-neuron activity time courses [12,13]. To go beyond single-neuron and pairwise statistics,   ₃₁
recent studies have examined the multi-dimensional structure of neuronal population activity to   ₃₂
uncover important insights into mechanisms underlying neuronal computation (e.g., [14–24]).   ₃₃
This has motivated the inquiry of whether network models reproduce such population activity   ₃₄
structure, in addition to single-neuron and pairwise statistics, raising the bar on what   ₃₅
constitutes an agreement between a network model and neuronal recordings [3].   ₃₆

Population activity structure can be characterized using dimensionality reduction [25–27],   ₃₇
which provides a concise summary (i.e., a low-dimensional representation) of how a population   ₃₈
of neurons covaries and how their activities unfold over time. Several dimensionality reduction   ₃₉
methods have been applied to neuronal population activity, including principal component   ₄₀
analysis (e.g., [14,15,20,28,29]), demixed principal component analysis [30], factor   ₄₁
analysis [16,19,31], Gaussian-process factor analysis [32], latent factor analysis via dynamical   ₄₂
systems [33], tensor component analysis [34], and more (see [25] for a review). The   ₄₃
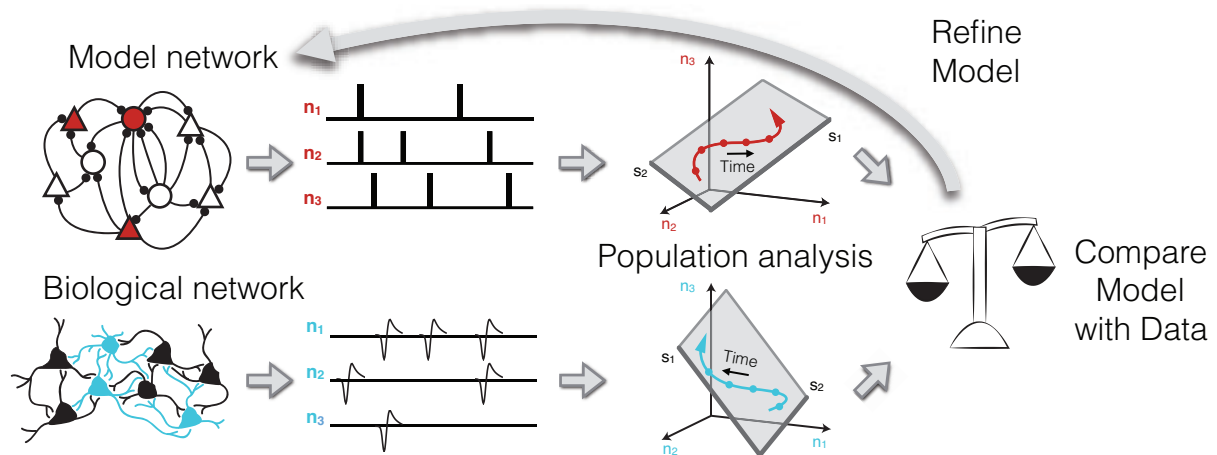
2

Fig 1. Relating biological and model networks using population analyses: Because a model network typically does not attempt to replicate the precise anatomical connectivity of a biological network, there is not a one-to-one correspondence of each biological neuron with a model neuron. Dimensionality reduction can be used to obtain a concise summary of the population activity from each network. This provides common ground for incisive comparisons between biological and model networks. Discrepancies in the population activity structure between biological and model networks can then help to refine model networks.

low-dimensional representation describes a neuronal process being carried out by the larger ⁴⁴ circuit from which the neurons were recorded [32, 35]. The same dimensionality reduction ⁴⁵ method can be applied to the recorded activity and to the network model activity, resulting in ⁴⁶ population activity structures that can be directly compared (Fig. 1). This benefit is also true ⁴⁷ of related methods for comparing neuronal recordings and network models involving neuronal ⁴⁸ decoding, population response similarity, and predicting the activity of one neuron from a ⁴⁹ population of other neurons [3]. ⁵⁰

    Dimensionality reduction has been adopted by recent studies to relate neuronal recordings ⁵¹ and network models to study working memory, decision making, motor control, and more. ⁵² Although many studies have separately employed large-scale neuronal recordings, large-scale ⁵³ network models, and dimensionality reduction, this review focuses on studies that incorporate ⁵⁴ all three components. Below we describe these studies, organized by the aspect of population ⁵⁵ activity structure used to relate neuronal recordings and network models: population activity ⁵⁶ time courses, functionally-defined neuronal subspaces, and population-wide neuronal variability. ⁵⁷ These were chosen first because they represent the key ways in which dimensionality reduction ⁵⁸ has been used in the literature to relate population recordings and network models. More ⁵⁹ importantly, these three categories represent fundamental aspects of population activity ⁶⁰ structure – how it unfolds over time, how different types of information can be encoded in ⁶¹ different subspaces, and how it varies from trial to trial. ⁶²

Population activity time courses ⁶³

Dynamical structures, such as point attractors, line attractors, and limit cycles, arising from ⁶⁴ network models have long been hypothesized to underlie the computational ability of biological ⁶⁵

networks of neurons [36–38]. Such dynamical structures have been implicated in decision   66
making [39,40], memory [41–43], oculomotor integration [44,45], motor control [46],   67
olfaction [47], and more. A fundamental question in systems neuroscience is whether these   68
dynamical structures are actually used by the brain. Although single-neuron and pairwise   69
metrics can be informative [42,45], analyzing the activity of a population of neurons together   70
has enabled deeper connections. In particular, the time course of the activity of a population of   71
neurons can be summarized by low-dimensional neuronal trajectories [25], as identified by   72
dimensionality reduction. These neuronal trajectories can provide a signature of a particular   73
dynamical structure. For example, a point attractor shows convergent trajectories. The   74
neuronal trajectories extracted from the recorded activity can then be compared with those   75
extracted from the network model activity. Such a comparison does not require a one-to-one   76
correspondence between each recorded neuron and a model neuron, but instead relies on a   77
summary of the population activity time courses.   78

This approach was recently used to study how the brain flexibly controls the timing of   79
behavior [48,49]. By applying dimensionality reduction to neuronal activity recorded from   80
medial frontal cortex, Wang et al. found that population activity time courses for different time   81
intervals followed a stereotypical path, but traversed that path at different speeds (Fig. 2A,   82
top). To understand how a network of neurons can accomplish this, the authors trained a   83
recurrent network model with 200 neurons to produce only the appropriate stimulus-behavior   84
relationships. Wang et al. then applied dimensionality reduction to the activity from the   85
network model. They surprisingly observed that the neuronal trajectories of the network model   86
also followed a stereotypical path, even though the network model was not trained to reproduce   87
the recorded activity (Fig. 2A, bottom). This population-level correspondence enabled by   88
dimensionality reduction laid the foundation for them to then dissect the network model to   89
understand the core neuronal mechanisms [50]. They found that the input to the network drove   90
the network activity from one fixed point to another, where the transition speed was   91
determined by the depth of the energy basin created by the input (Fig. 2A, bottom).   92

Other studies have also used this approach to understand how the time course of neuronal   93
activity relates to computations underlying motor control [12,51–53], decision   94
making [17,54–56], and working memory [13,57]. In each of these studies, a network model was   95
constructed without referencing the recorded activity. Dimensionality reduction was applied to   96
extract neuronal trajectories to obtain a correspondence between the neuronal recordings and   97
network models. To study the neuronal mechanisms underlying the observed time courses, the   98
network models were then dissected to reveal dynamical structures, such as fixed points or   99
point attractors [17,54,55], line attractors [17], and oscillatory modes [12,51–53]. Whether or   100
not these dynamical structures are indeed at play in real neuronal networks is still an open   101
question. Nevertheless, these studies are beginning to demonstrate that it is at least fruitful to   102
interpret neuronal activity in terms of these dynamical structures, a process facilitated by   103
dimensionality reduction.   104

Functionally-defined neuronal subspaces   105

Recent studies have investigated how distinct types of information encoded by the *same*   106
neuronal population can be parsed by downstream brain circuits [58–60]. An enticing proposal   107
is that different types of information are encoded in different subspaces within the population   108
activity space, where the subspaces are identified using dimensionality reduction. For example,   109
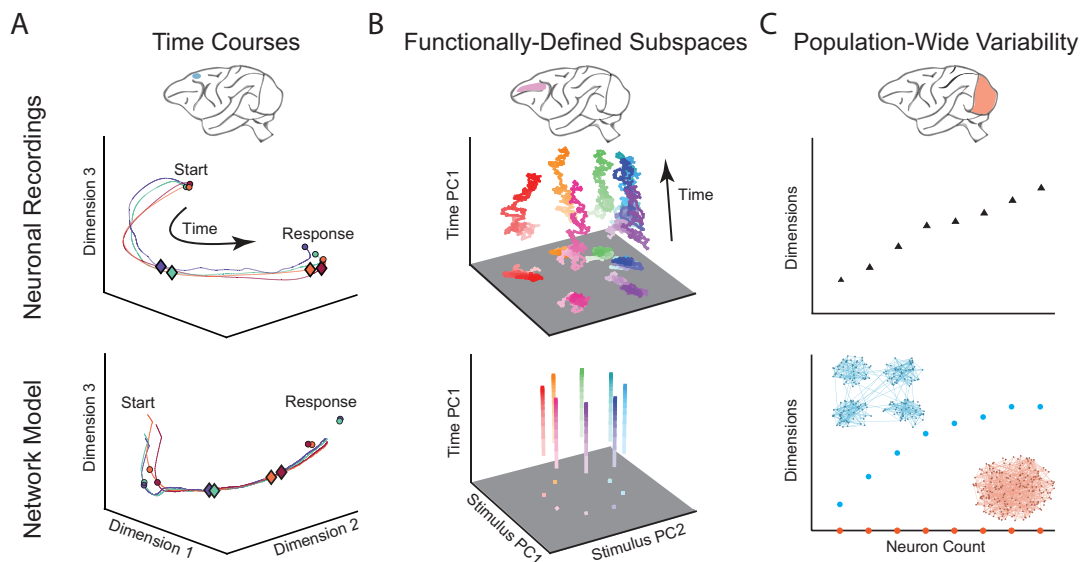
4

Fig 2. Examples of comparing neuronal recordings and network models using dimensionality reduction. A. (Top) Population activity time courses from medial frontal cortex during a time production task. Each trajectory represents a time course of neuronal activity during a different produced time interval. Circles represent the start and end of the time production interval and diamonds represent a fixed time interval after the start circle. Diamonds appear closer to the start of the trajectory on long-interval trials (blue) than short-interval trials (red), indicating that neuronal activity traverses the path at different speeds during the two intervals. (Bottom) Circles represent fixed points in the model network's dynamics and diamonds represent a fixed time interval after the start of the time production task. A similar difference in traversal speed is observed in the model network as was observed in the neural recordings. Adapted with permission from [48]. B. (Top) Delay period activity from prefrontal cortex during a delayed saccade task. Each trajectory represents a different stimulus condition. The trajectories for different stimuli remain well-separated in a stimulus subspace throughout the delay period. (Bottom) Network model activity demonstrating similar subspace stability. Adapted with permission from [20]. C. (Top) Dimensionality of population-wide neuronal variability in primary visual cortex increases with the number of neurons recorded. (Bottom) A similar dimensionality trend is observed for a spiking network model with clustered excitatory connections (blue), but not for a model with unstructured connectivity (red). Adapted with permission from [31].

Kaufman et al. [18] asked how it is possible for neurons in the motor cortex to be active during motor preparation, yet not generate an arm movement. They found that motor cortical activity during motor preparation resided outside of the activity subspace most related to muscle

contractions. This allows the motor cortex to prepare arm movements without driving   113
downstream circuits, a characteristic which can be implemented by a linear readout mechanism.   114
This concept of functionally-defined neuronal subspaces has also been used in other studies of   115
motor control [23, 61–63], decision making [30, 64], short-term memory [30, 65], learning [19],   116
and visual processing [24].   117

  To understand how a neuronal circuit can implement and exploit such functionally-defined   118
neuronal subspaces, one can construct a network model to see whether it reproduces the   119
empirical observations. If so, one can then dissect the network to study the underlying   120
mechanisms. Mante et al. [17] applied dimensionality reduction to recordings in prefrontal   121
cortex to find that motion and color of the visual stimulus were encoded in distinct subspaces.   122
They then trained a recurrent network model with 100 neurons to produce only the appropriate   123
stimulus-behavior relationships. When they applied the same dimensionality reduction method   124
to the network model activity, they surprisingly found that the motion and color of the visual   125
stimulus were also encoded in distinct subspaces, even though the network model was not   126
trained to reproduce the recorded activity. This population-level correspondence between the   127
network model and recordings was enabled by dimensionality reduction and went beyond   128
comparisons based on individual neurons or pairs of neurons. Mante et al. then dissected the   129
network model to uncover how the two types of information encoded in distinct subspaces can   130
be selectively used to form a decision.   131

  Dimensionality reduction has also revealed that, in some cases, standard network models do   132
not reproduce the functionally-defined subspaces identified from neuronal recordings. For   133
example, Murray et al. [20] applied dimensionality reduction to recordings in prefrontal cortex   134
during a working memory task to find that, even though firing rates of individual neurons   135
changed over time, there was a subspace in which the activity stably encoded the memorized   136
target location (Fig. 2B, top). They then applied the same analyses to activity from several   137
prominent network models and found that none of them reproduced both the time-varying   138
activity of individual neurons *and* the subspace in which the memory was stably encoded. This   139
provided the impetus to develop a new network model that did reproduce these features of the   140
recorded activity (Fig. 2B, bottom) (see also [66]). As another example, Elsayed et al. [67]   141
found that standard network models do not reproduce the empirical observation described   142
above that neuronal activity during movement preparation and movement execution lie in   143
orthogonal subspaces. Such insights obtained using dimensionality reduction can guide the   144
development of more sophisticated network models.   145

Population-wide neuronal variability   146

The previous sections focus largely on neuronal activity that is averaged across trials and on   147
firing rate-based network models. This naturally obscures the trial-to-trial variability that is a   148
fundamental feature of neuronal responses across the cortex [68], both at the level of single   149
neuron responses [69] as well as variability shared by the population [11, 70]. Theoretical and   150
experimental studies have focused on how the structure of that variability places limits on   151
information coding [71–74], and in turn influences our behavior. At the same time, a growing   152
body of work has demonstrated that variability can be thought of not only as noise to be   153
removed, but also a signature of ongoing decision processes and cognitive variables   154
(e.g., [75–77]). To move beyond single-neuron and pairwise measurements of neuronal   155
variability, recent studies have begun to consider population-wide measures of neuronal   156

6

variability [78–82], as enabled by dimensionality reduction. Such measures allow one to i) assess   157
whether the large number of single-neuron and pairwise variability measurements can be   158
succinctly summarized by a small number of variables (e.g., the entire population increasing   159
and decreasing its activity together can be described by a single scalar variable), and ii) relate   160
the population activity on individual experimental trials to behavior [22, 32–34, 83–85].   161

In parallel with the growing interest in neuronal variability, there have been attempts to   162
create network models that exhibit variability matching recorded neurons. In particular, a class   163
of models has used the balance between excitation and inhibition as a way to generate   164
variability as an emergent property of network structure, rather than via an external variable   165
source [71, 86, 87]. In these models, the particular structure of the network has a large impact   166
on the population-wide variability that emerges. Using the lens of factor analysis, Williamson et   167
al. [31] found that the dimensionality of spontaneous activity fluctuations in V1 neurons   168
increases with the number of recorded neurons (Fig. 2C, top). This was more consistent with   169
activity generated by networks with clustered excitatory connections [8] than networks with   170
unstructured connectivity [86] (Fig. 2C, bottom). The combination of population-wide   171
measures of variability (in this case, dimensionality) and the ability to manipulate model   172
network structures facilitated an understanding of how features of variability observed in   173
biological networks relate to network structure.   174

The approach of using dimensionality reduction to compare the population-wide variability   175
of neuronal recordings and network models has also been applied to study spontaneous versus   176
evoked activity [88, 89], the activity of different classes of neurons [90], and the activity during   177
different behavioral conditions, such as attention [81, 82]. Dimensionality reduction has also   178
been used to analyze population activity from balanced network models to help identify the   179
crucial network architecture and synaptic timescales required to produce the low-dimensional   180
shared variability that is widely reported in neuronal recordings [82, 87]. Together these studies   181
demonstrate the power of combining dimensionality reduction and network models to   182
understand the mechanisms and effects of neuronal variability.   183

Conclusion   184

Dimensionality reduction has enabled incisive comparisons between biological and model   185
networks in terms of population activity time courses, functionally-defined neuronal subspaces,   186
and population-wide neuronal variability. Such comparisons result in either i) a correspondence   187
between the neuronal recordings and the network model, in which case the model can be   188
dissected to understand underlying network mechanisms, or ii) discrepancies between the   189
neuronal recordings and standard network models, leading to the development of improved   190
models. This approach (cf. Fig. 1) has already provided insight into the neuronal mechanisms   191
underlying brain functions such as working memory, decision making, and motor control, and is   192
likely to become even more important as the scale of neuronal recordings and network models   193
grows.   194

A key consideration in network modeling is what aspects of neuronal recordings the model   195
should reproduce. We posit that the population activity structure (including population   196
activity time courses, functionally-defined neuronal subspaces, and population-wide neuronal   197
variability) will provide key signatures of how neurons work together to give rise to brain   198
function. Thus, if a network model is to provide a systems-level account of brain function, we   199
should require it to reproduce the population activity structure of neuronal recordings, in   200

7

addition to existing population metrics [3] and standard spike train statistics based on individual and pairs of neurons.

Most studies described here have used neuronal recordings to inform network models via dimensionality reduction. An important future direction is to use network models and dimensionality reduction to design new experiments and form predictions. For example, if one day we can experimentally perturb neuronal activity in specified directions in the population activity space [91], we can test whether driving the population activity in particular directions leads to particular decisions or movements predicted by the network model. The hope is to establish a virtuous cycle, where neuronal recordings and network models closely inform each other through the common ground provided by dimensionality reduction.

Acknowledgements

References

1. Stevenson, I.H. and Kording, K.P. (2011). How advances in neural recording affect data analysis. Nature Neuroscience *14*, 139–142.

2. Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J.M., Diesmann, M., Morrison, A., Goodman, P.H., Harris, F.C., et al. (2007). Simulation of networks of spiking neurons: a review of tools and strategies. Journal of Computational Neuroscience *23*, 349–398.

3. Yamins, D.L.K. and DiCarlo, J.J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature Neuroscience *19*, 356–365.

4. Barak, O. (2017). Recurrent neural networks as versatile tools of neuroscience research. Current Opinion in Neurobiology *46*, 1–6.

5. Lee, W.C.A., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., and Reid, R.C. (2016). Anatomy and function of an excitatory network in the visual cortex. Nature *532*, 370 EP –.

6. Roxin, A., Brunel, N., Hansel, D., Mongillo, G., and van Vreeswijk, C. (2011). On the distribution of firing rates in networks of cortical neurons. Journal of Neuroscience *31*, 16217–16226.

7. Chariker, L., Shapley, R., and Young, L.S. (2016). Orientation selectivity from very sparse LGN inputs in a comprehensive model of macaque V1 cortex. Journal of Neuroscience *36*, 12368–12384.

8

8. Litwin-Kumar, A. and Doiron, B. (2012). Slow dynamics and high variability in balanced   240
   cortical networks with clustered connections. Nature Neuroscience *15*, 1498–1505.   241

9. Trousdale, J., Hu, Y., Shea-Brown, E., and Josić, K. (2012). Impact of network structure   242
   and cellular response on spike time correlations. PLoS computational biology *8*, e1002408.   243

10. Stringer, C., Pachitariu, M., Steinmetz, N.A., Okun, M., Bartho, P., Harris, K.D., Sahani,   244
    M., and Lesica, N.A. (2016). Inhibitory control of correlated intrinsic variability in   245
    cortical networks. eLife *5*, e19695.   246

11. Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K., and Josić, K. (2016). The   247
    mechanics of state-dependent neural correlations. Nature neuroscience *19*, 383–393.   248

12. Sussillo, D., Churchland, M.M., Kaufman, M.T., and Shenoy, K.V. (2015). A neural   249
    network that finds a naturalistic solution for the production of muscle activity. Nature   250
    Neuroscience *18*, 1025–1033.   251

13. Rajan, K., Harvey, C.D., and Tank, D.W. (2016). Recurrent network models of sequence   252
    generation and memory. Neuron *90*, 128–142.   253

14. Mazor, O. and Laurent, G. (2005). Transient dynamics versus fixed points in odor   254
    representations by locust antennal lobe projection neurons. Neuron *48*, 661–673.   255

15. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P.,   256
    Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature   257
    487, 51–56.   258

16. Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal   259
    cortex during a virtual-navigation decision task. Nature *484*, 62–68.   260

17. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent   261
    computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.   262

18. Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2014). Cortical activity   263
    in the null space: permitting preparation without movement. Nature Neuroscience *17*,   264
    440–448.   265

19. Sadtler, P.T., Quick, K.M., Golub, M.D., Chase, S.M., Ryu, S.I., Tyler-Kabara, E.C., Yu,   266
    B.M., and Batista, A.P. (2014). Neural constraints on learning. Nature *512*, 423–426.   267

20. Murray, J.D., Bernacchia, A., Roy, N.A., Constantinidis, C., Romo, R., and Wang, X.J.   268
    (2017). Stable population coding for working memory coexists with heterogeneous neural   269
    dynamics in prefrontal cortex. Proceedings of the National Academy of Sciences *114*,   270
    394–399.   271

21. Remington, E.D., Egger, S.W., Narain, D., Wang, J., and Jazayeri, M. (2018). A   272
    dynamical systems perspective on flexible motor timing. Trends in Cognitive Sciences *22*,   273
    938–952.   274

22. Ruff, D.A., Ni, A.M., and Cohen, M.R. (2018). Cognition as a window into neuronal   275
    population space. Annual Review of Neuroscience *41*, 77–97.   276

23. Perich, M.G., Gallego, J.A., and Miller, L.E. (2018). A neural population mechanism for rapid learning. Neuron *100*, 964–976.e7.

24. Semedo, J.D., Zandvakili, A., Machens, C.K., Yu, B.M., and Kohn, A. (2018). Cortical areas interact through a communication subspace. Neuron , in press.

25. Cunningham, J.P. and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nature Neuroscience *17*, 1500–1509.

26. Gao, P. and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. Current Opinion in Neurobiology *32*, 148–155.

27. Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural manifolds for the control of movement. Neuron *94*, 978 – 984.

28. Cowley, B.R., Smith, M.A., Kohn, A., and Yu, B.M. (2016). Stimulus-driven population activity patterns in macaque primary visual cortex. PLoS Computational Biology *12*, e1005185.

29. Gallego, J.A., Perich, M.G., Naufel, S.N., Ethier, C., Solla, S.A., and Miller, L.E. (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. Nature Communications *9*, 4233.

30. Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., and Machens, C.K. (2016). Demixed principal component analysis of neural population data. eLife *5*, e10989.

31. Williamson, R.C., Cowley, B.R., Litwin-Kumar, A., Doiron, B., Kohn, A., Smith, M.A., and Yu, B.M. (2016). Scaling properties of dimensionality reduction for neural populations and network models. PLoS Computational Biology *12*, e1005141.

32. Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. Journal of Neurophysiology *102*, 614–635.

33. Pandarinath, C., O'Shea, D.J., Collins, J., Jozefowicz, R., Stavisky, S.D., Kao, J.C., Trautmann, E.M., Kaufman, M.T., Ryu, S.I., Hochberg, L.R., et al. (2018). Inferring single-trial neural population dynamics using sequential auto-encoders. Nature Methods *15*, 805–815.

34. Williams, A.H., Kim, T.H., Wang, F., Vyas, S., Ryu, S.I., Shenoy, K.V., Schnitzer, M., Kolda, T.G., and Ganguli, S. (2018). Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. Neuron *98*, 1099–1115.e8.

35. Buonomano, D.V. and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. Nature Reviews Neuroscience *10*, 113 EP –. Review Article.

36. Wilson, H.R. and Cowan, J.D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik *13*, 55–80.

10

37. Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective ₃₁₅ computational abilities. Proceedings of the National Academy of Sciences *79*, 2554–2558. ₃₁₆

38. Brody, C.D., Romo, R., and Kepecs, A. (2003). Basic mechanisms for graded persistent ₃₁₇ activity: discrete attractors, continuous attractors, and dynamic representations. Current ₃₁₈ Opinion in Neurobiology *13*, 204–211. ₃₁₉

39. Machens, C.K., Romo, R., and Brody, C.D. (2005). Flexible control of mutual inhibition: ₃₂₀ a neural model of two-interval discrimination. Science *307*, 1121–1124. ₃₂₁

40. Wang, X.J. (2008). Decision making in recurrent neuronal circuits. Neuron *60*, 215–234. ₃₂₂

41. Wang, X.J. (2001). Synaptic reverberation underlying mnemonic persistent activity. ₃₂₃ Trends in Neurosciences *24*, 455–463. ₃₂₄

42. Wimmer, K., Nykamp, D.Q., Constantinidis, C., and Compte, A. (2014). Bump attractor ₃₂₅ dynamics in prefrontal cortex explains behavioral precision in spatial working memory. ₃₂₆ Nature Neuroscience *17*, 431–439. ₃₂₇

43. Chaudhuri, R. and Fiete, I. (2016). Computational principles of memory. Nature ₃₂₈ Neuroscience *19*, 394–403. ₃₂₉

44. Seung, H.S. (1996). How the brain keeps the eyes still. Proceedings of the National ₃₃₀ Academy of Sciences *93*, 13339–13344. ₃₃₁

45. Miri, A., Daie, K., Arrenberg, A.B., Baier, H., Aksay, E., and Tank, D.W. (2011). Spatial ₃₃₂ gradients and multidimensional dynamics in a neural integrator circuit. Nature ₃₃₃ Neuroscience *14*, 1150–1159. ₃₃₄

46. Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm ₃₃₅ movements: a dynamical systems perspective. Annual Review of Neuroscience *36*, ₃₃₆ 337–359. ₃₃₇

47. Rabinovich, M., Huerta, R., and Laurent, G. (2008). Transient dynamics for neural ₃₃₈ processing. Science *321*, 48–50. ₃₃₉

48. Wang, J., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018). Flexible timing by ₃₄₀ temporal scaling of cortical responses. Nature Neuroscience *21*, 102–110. ₃₄₁

49. Remington, E.D., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018). Flexible ₃₄₂ sensorimotor computations through rapid reconfiguration of cortical dynamics. Neuron ₃₄₃ *98*, 1005–1019.e5. ₃₄₄

50. Sussillo, D. and Barak, O. (2013). Opening the black box: low-dimensional dynamics in ₃₄₅ high-dimensional recurrent neural networks. Neural Computation *25*, 626–649. ₃₄₆

51. Hennequin, G., Vogels, T., and Gerstner, W. (2014). Optimal control of transient ₃₄₇ dynamics in balanced networks supports generation of complex movements. Neuron *82*, ₃₄₈ 1394 – 1406. ₃₄₉

52. Michaels, J.A., Dann, B., and Scherberger, H. (2016). Neural population dynamics ₃₅₀ during reaching are better explained by a dynamical system than representational tuning. ₃₅₁ PLoS Computational Biology *12*, e1005175. ₃₅₂

53. Russo, A.A., Bittner, S.R., Perkins, S.M., Seely, J.S., London, B.M., Lara, A.H., Miri, A.,   353
Marshall, N.J., Kohn, A., Jessell, T.M., et al. (2018). Motor cortex embeds muscle-like   354
commands in an untangled population response. Neuron *97*, 953–966.e8.   355

54. Carnevale, F., de Lafuente, V., Romo, R., Barak, O., and Parga, N. (2015). Dynamic   356
control of response criterion in premotor cortex during perceptual detection under   357
temporal uncertainty. Neuron *86*, 1067–1077.   358

55. Chaisangmongkon, W., Swaminathan, S.K., Freedman, D.J., and Wang, X.J. (2017).   359
Computing by robust transience: how the fronto-parietal network performs sequential,   360
category-based decisions. Neuron *93*, 1504–1517.e4.   361

56. Mastrogiuseppe, F. and Ostojic, S. (2018). Linking connectivity, dynamics, and   362
computations in low-rank recurrent neural networks. Neuron *99*, 609–623.e29.   363

57. Barak, O., Sussillo, D., Romo, R., Tsodyks, M., and Abbott, L.F. (2013). From fixed   364
points to chaos: Three models of delayed discrimination. Progress in Neurobiology *103*,   365
214–222.   366

58. Park, I.M., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2014). Encoding and decoding   367
in parietal cortex during sensorimotor decision-making. Nature Neuroscience *17*,   368
1395–1403.   369

59. Pagan, M. and Rust, N.C. (2014). Quantifying the signals contained in heterogeneous   370
neural responses and determining their relationships with task performance. Journal of   371
Neurophysiology *112*, 1584–1598.   372

60. Fusi, S., Miller, E.K., and Rigotti, M. (2016). Why neurons mix: high dimensionality for   373
higher cognition. Current Opinion in Neurobiology *37*, 66–74.   374

61. Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in   375
premotor cortex during motor planning. Nature *532*, 459–464.   376

62. Miri, A., Warriner, C.L., Seely, J.S., Elsayed, G.F., Cunningham, J.P., Churchland,   377
M.M., and Jessell, T.M. (2017). Behaviorally selective engagement of short-latency   378
effector pathways by motor cortex. Neuron *95*, 683 – 696.e11.   379

63. Hennig, J.A., Golub, M.D., Lund, P.J., Sadtler, P.T., Oby, E.R., Quick, K.M., Ryu, S.I.,   380
Tyler-Kabara, E.C., Batista, A.P., Yu, B.M., et al. (2018). Constraints on neural   381
redundancy. eLife *7*, e36774.   382

64. Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural   383
population supports evolving demands during decision-making. Nature Neuroscience *17*,   384
1784–1792.   385

65. Daie, K., Goldman, M., and Aksay, E.F. (2015). Spatial patterns of persistent neural   386
activity vary with the behavioral context of short-term memory. Neuron *85*, 847–860.   387

66. Druckmann, S. and Chklovskii, D. (2012). Neuronal circuits underlying persistent   388
representations despite time varying activity. Current Biology *22*, 2095–2103.   389

12

67. Elsayed, G.F., Lara, A.H., Kaufman, M.T., Churchland, M.M., and Cunningham, J.P. (2016). Reorganization between preparatory and movement population responses in motor cortex. Nature Communications *7*, 13239.

68. Renart, A. and Machens, C.K. (2014). Variability in neural activity and behavior. Current Opinion in Neurobiology *25*, 211–220.

69. Faisal, A.A., Selen, L.P., and Wolpert, D.M. (2008). Noise in the nervous system. Nature Reviews Neuroscience *9*, 292–303.

70. Cohen, M.R. and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nature Neuroscience *14*, 811–819.

71. Shadlen, M.N. and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. Journal of Neuroscience *18*, 3870–3896.

72. Abbott, L.F. and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. Neural Computation *11*, 91–101.

73. Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nature Reviews Neuroscience *7*, 358–366.

74. Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. Nature Neuroscience *17*, 1410–1417.

75. Cohen, M.R. and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nature Neuroscience *12*, 1594–1600.

76. Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron *63*, 879–888.

77. Nienborg, H., R. Cohen, M., and Cumming, B.G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annual Review of Neuroscience *35*, 463–483.

78. Ecker, A., Berens, P., Cotton, R.J., Subramaniyan, M., Denfield, G., Cadwell, C., Smirnakis, S., Bethge, M., and Tolias, A. (2014). State dependence of noise correlations in macaque primary visual cortex. Neuron *82*, 235–248.

79. Rabinowitz, N.C., Goris, R.L., Cohen, M., and Simoncelli, E.P. (2015). Attention stabilizes the shared gain of V4 populations. eLife *4*, e08998.

80. Lin, I.C., Okun, M., Carandini, M., and Harris, K.D. (2015). The nature of shared cortical variability. Neuron *87*, 644–656.

81. Kanashiro, T., Ocker, G.K., Cohen, M.R., and Doiron, B. (2017). Attentional modulation of neuronal variability in circuit models of cortex. eLife *6*, e23978.

82. Huang, C., Ruff, D., Pyle, R., Rosenbaum, R., Cohen, M., and Doiron, B.D. (2018). Circuit models of low dimensional shared variability in cortical networks. Neuron , in press.

13

83. Cohen, M.R. and Maunsell, J.H. (2010). A neuronal population measure of attention predicts behavioral performance on individual trials. Journal of Neuroscience *30*, 15241–15253. 427 428 429

84. Kiani, R., Cueva, C., Reppas, J., and Newsome, W. (2014). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. Current Biology *24*, 1542–1547. 430 431 432

85. Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2015). Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. eLife *4*, e04677. 433 434 435

86. Van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science *274*, 1724–1726. 436 437

87. Rosenbaum, R., Smith, M.A., Kohn, A., Rubin, J.E., and Doiron, B. (2017). The spatial structure of correlated neuronal variability. Nature Neuroscience *20*, 107–114. 438 439

88. Mazzucato, L., Fontanini, A., and La Camera, G. (2016). Stimuli reduce the dimensionality of cortical activity. Frontiers in Systems Neuroscience *10*. 440 441

89. Hennequin, G., Ahmadian, Y., Rubin, D.B., Lengyel, M., and Miller, K.D. (2018). The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability. Neuron *98*, 846–860.e5. 442 443 444

90. Bittner, S.R., Williamson, R.C., Snyder, A.C., Litwin-Kumar, A., Doiron, B., Chase, S.M., Smith, M.A., and Yu, B.M. (2017). Population activity structure of excitatory and inhibitory neurons. PLoS One *12*, e0181773. 445 446 447

91. Jazayeri, M. and Afraz, A. (2017). Navigating the neural space in search of the neural code. Neuron *93*, 1003–1014. 448 449

Mante et al. [17]*: In a context-dependent decision-making task, the authors found that ⁤⁤⁤⁤⁤⁤450
color and motion information were encoded in distinct subspaces of PFC population activity. A ⁤451
recurrent network model trained to perform the same task revealed a network-level mechanism ⁤452
of how the two types of information can be selectively used to form a decision. ⁤453

⁤454

Murray et al. [20]**: The authors analyzed population activity recorded in prefrontal cortex ⁤455
and identified a low-dimensional subspace in which stimulus information was reliably encoded, ⁤456
despite the fact that individual neurons showed substantial time-varying activity. They then ⁤457
found that standard models did not reproduce this empirical observation, and proceeded to ⁤458
develop a "stable subspace" model that did reproduce this observation. ⁤459

⁤460

Wang et al. [48]**: Recording from medial frontal cortex during a timing task, the authors ⁤461
found that population activity time courses followed a stereotypical path, but traversed the ⁤462
path at different speeds based on the duration of the timing interval. They found similar trends ⁤463
in a recurrent network model trained to perform the task and showed that speed of traversal ⁤464
was determined by the network inputs. ⁤465

⁤466

Sussillo et al. [12]*: The authors trained a recurrent network model to produce muscle ⁤467
activity patterns observed in an arm reaching task. The model activity surprisingly showed ⁤468
rotational dynamics that mimicked those observed empirically in M1 population recordings. ⁤469

⁤470

Chaisangmongkan et al. [55]*: The authors found that PFC and LIP neurons show mixed ⁤471
selectivity during a delayed match-to-category task, and that the neuronal trajectories ⁤472
extracted using dimensionality reduction are interpretable during each epoch of the task. They ⁤473
then constructed a recurrent network model to understand the network principles that govern ⁤474
the activity time courses during this task. ⁤475

⁤476

Williamson et al. [31]**: This study compared the population activity structure of V1 ⁤477
recordings and spiking network models while varying the number of neurons and trials analyzed. ⁤478
The scaling trends of the V1 recordings better resembled a model with clustered excitatory ⁤479
connections than one with unstructured connectivity. ⁤480

⁤481

Elsayed et al. [67]*: This study found that M1 population activity during movement ⁤482
preparation and movement execution resides in orthogonal subspaces. Standard network models ⁤483
did not reproduce this empirical observation. ⁤484

⁤485

Mazzucato et al. [88]*: Comparing gustatory cortex recordings and spiking network models, ⁤486
the authors examined how the dimensionality of population activity grows with population size ⁤487
during spontaneous and evoked activity. They then developed a theoretical upper bound on ⁤488
dimensionality based on the level of pairwise correlations. ⁤489

15