Transient chaotic dimensionality expansion by recurrent networks

Summary

Cortical neurons communicate with spikes, which are discrete events in time. Functional network models often employ rate units that are continuously coupled by analog signals. Is there a benefit of discrete signaling? By a unified mean-field theory for large random networks of rate and binary units, we show that both models have identical second order statistics. Their stimulus processing properties, however, are radically different: We discover a chaotic submanifold in binary networks that does not exist in rate models. Its dimensionality increases with time after stimulus onset and reaches a fixed point that depends on the synaptic coupling strength. Low dimensional stimuli are transiently expanded into higher-dimensional representations that live within the manifold. We find that classification performance peaks when stimulus dimensionality matches the submanifold dimension; typically within a single neuronal time constant. Classification shows a high resilience to noise that exceeds rate models by orders of magnitude. Our theory mechanistically explains all these observations.

These findings have several implications. 1) Optimal performance is reached with weaker synapses in discrete state networks compared to rate models; implying lower energetic costs for synaptic transmission. 2) The classification mechanism is robust to noise, compatible with fluctuations in biophysical systems. 3) Optimal performance is reached when each neuron in the network has been activated only once; this demonstrates efficient event-based computation with short latencies. 4) The presence of a chaotic sub-manifold has implications for the variability of neuronal activity; the theory predicts a transient increase of variability after stimulus onset. Our theory thus provides a new link between recurrent and chaotic dynamics of functional networks, neuronal variability, and dimensionality of neuronal responses.

Additional Detail

We illustrate the theory for two seminal network models. First, the classical network of ref. [5], their eq. (2), where neurons have continuous state variables $x_i(t) \in \mathbb{R}$. Second, the model of binary neurons from ref. [6], their eqs. (2.1,2). Neurons here communicate with discrete events and have a binary state $x_i(t) \in \{-1, 1\}$ that changes only at discrete points in time. We use $\phi = \tanh$ as the nonlinear gain function for both models. The synaptic coupling matrix has random i.i.d. Gaussian entries with zero mean and variance $N^{-1}g^2$. Extensions to more realistic settings are straight forward.

By methods from disordered systems and field theory [2–4], we derive a dynamical mean-field theory for the autocorrelation function $Q(\tau) := \langle h(t+\tau)\, h(t)\rangle$ of the synaptic input $h_i(t) = \sum_j J_{ij}\, \phi(x_j(t))$ that is exact for large networks $N \to \infty$. We find for both models the same equation of motion of Newtonian form

$$\ddot{Q}\,(\tau) = -V'_{Q_0,g^2}\,(Q\,(\tau))\,, \tag{1}$$

where the "potential" V depends on the variance of the input $Q_0 = Q(0)$, the gain function ϕ, and the coupling strength g^2. For the rate model this is the classical result of ref. [5], for the binary model this result is new. This unified theory allows us to match the parameters for the two model classes to have identical mean-field theories; we can thus compare collective information processing across model classes in the very same dynamical state.

To analyze how the binary network processes stimuli, we compute the dimensionality $d(t)$ of the state space covered by the network's response: $d(t)$ measures the maximal number of neuronal coordinates that differ across any pair of network states.

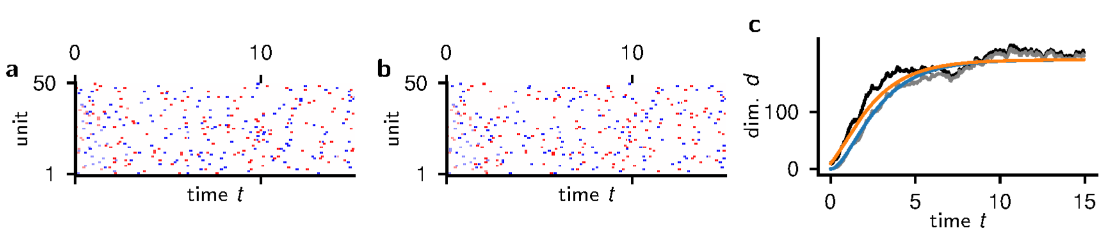

Figure 1. **Transient increase of response dimensionality in a binary network.** **a** Time evolution of first 50 neurons; up transitions in red, down in blue. Stimulus applied by initializing first $L = 10$ neurons randomly $\in \{-1, 1\}$ with equal probability. **b** Time evolution for different realization of stimulus. **c** Dimensionality of the activity submanifold $d_{\mathrm{signal}}(t)$ as a function of time, explored by the network across different stimuli (black; theory (2) in red). Dimensionality $d_{\mathrm{noise}}(t)$ due to different noise realizations: Gaussian noise with standard deviation $\sigma = 0.20$ added to each of the L entries of the initial state (gray; theory (2) in blue). Other parameters: $N = 500$ neurons, coupling strength $g = 0.8$.

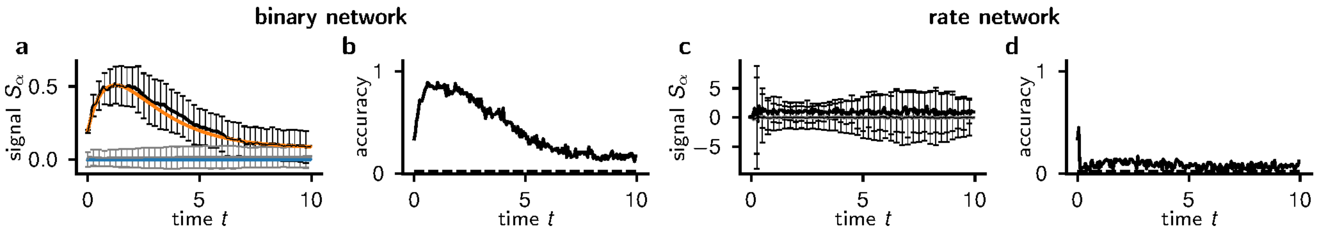

Figure 2. **Classification by binary network (a, b) and by rate network (c, d).** Both networks are in dynamically identical states, matched by the mean-field theory (1). **a**, **c** Linear readout $S_\alpha(t) = w^*_\alpha(t)^{\mathrm{T}}(x_\beta(t) + \xi)$ trained for each t to detect stimulus identity by minimizing $(S_\alpha - \delta_{\alpha\beta})^2$. Readout S_α for matching stimulus $\alpha = \beta$ (black) and for non-matching stimuli $\alpha \neq \beta$ (gray). Theoretical prediction (3) in red and blue. **b**, **d** Ratio of correct classification. Dashed horizontal line indicates chance level P^{-1}. Same network as in Figure 1. Standard deviation of readout noise ξ is 10^{-4}.

A replica calculation yields the closed form differential equation

$$\frac{d}{dt}\,d(t) = -d(t) + g\,\frac{2}{\sqrt{\pi}}\,\langle\phi'(h)\rangle_{h\sim\mathcal{N}(0,Q_0)}\,\sqrt{N\,d(t)}, \tag{2}$$

which is compared to simulation in Figure 1**c**. There are two terms: a linear term $\propto -d(t)$ and a square root term $\propto \sqrt{d(t)}$. For small d, the square root term wins, so that dimensionality increases; the dynamics is chaotic, the network progressively explores a wider space. For large d, the linear term dominates. In between there is a stable fixed point $\frac{d(\infty)}{N\,g^2} = \left(\frac{2}{\sqrt{\pi}}\langle\phi'(h)\rangle\right)^2 \leq \frac{4}{\pi}$.

What does this fixed point imply? It shows that the network does not explore its entire space of 2^Nstates. Instead, two states that differ initially, for example due to applied stimuli, always stay closer than the average distance $\sim d(\infty)/2$. This is visible in the raster plots in Figure 1**a** and **b**: Short after stimulus onset, network states are quite similar across stimuli. As time progresses, the two states become more and more distinct. An initially low-dimensional set of stimuli is thus expanded into a higher dimensional representation. For long times, however, the two states differ in only typically $d(\infty)/2$ of their neurons. This limiting dimensionality grows linearly in Ng^2. Variability across different stimuli [1] is thus intimately related to the dimensionality of the network state.

What are the functional consequences of this transient expansion of dimensionality? To find out, we train the network to classify P random binary patterns, each of length L. Each training sample is a noisy version of its prototype pattern. One linear readout w_α per pattern is trained by linear regression. The readout signal peaks at around $t \simeq 1$, a time span within which every neuron in the network has been activated only once (Figure 2**a**); classification accuracy is close to perfect (**b**). This shows that the transient chaotic dimensionality expansion is extremely fast. To explain this finding, we derive an approximation for the signal

$$S_\alpha(t) \simeq \frac{d_{\mathrm{signal}}(t)}{P}\left(1 - \frac{d_{\mathrm{noise}}(t)}{d_{\mathrm{signal}}(t)}\right), \tag{3}$$

expressed in terms of the dimensionalities $d_{\mathrm{signal}}(t)$ and $d_{\mathrm{noise}}(t)$, shown in Figure 2**a**. The first factor $d_{\mathrm{signal}}(t)/P$ dominates the initial increase of the signal. It relates the dimensionality of the signal manifold to the dimensionality of the stimulus, P: Performance increases as stimuli are expanded in higher dimensions. The latter factor depends on $d_{\mathrm{signal}}(t)/d_{\mathrm{noise}}(t)$ and determines how likely noise corrupts the signal, causing the gradual decline of the signal for long times; the maximum is a compromise between these two processes.

High resilience to noise of the binary network results from the average large distance $d(\infty)$ between any pair of states: even considerable noise (compare d_{signal} vs d_{noise} in Figure 1**c**) is tolerable. A network with continuous rate dynamics in the identical dynamical state and setting, in contrast, shows vanishing performance; a tiny amount of Gaussian readout noise ξ is sufficient to ruin classification accuracy (Figure 2**d**). What causes this drastic difference between model classes? In the rate network, states show little separation and ultimately collapse, $d(\infty) \to 0$, so that weak noise causes misclassification (Figure 2**d**). Discrete-state networks thus require weaker synaptic coupling, few neuronal activations, and therefore little resources, to implement efficient computation by transient chaotic dimensionality expansion.

REFERENCES

[1] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, 273(5283):1868–1871, 1996.

[2] A. Crisanti and H. Sompolinsky. Path integral approach to random neural networks. *Phys Rev E*, 98:062120, Dec 2018.

[3] D. Martí, N. Brunel, and S. Ostojic. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. *Phys. Rev. E*, 97:062314, Jun 2018.

[4] J. Schuecker, S. Goedeke, and M. Helias. Optimal sequence memory in driven random networks. *Phys Rev X*, 8:041029, 2018.

[5] H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Phys. Rev. Lett.*, 61:259–262, Jul 1988.

[6] C. van Vreeswijk and H. Sompolinsky. Chaotic balanced state in a model of cortical circuits. *Neural Comput.*, 10:1321–1371, 1998.