# NEWS AND VIEWS

Whether this suppressive function of M1 is a special feature of the rodent whisker system or a general principle conserved across mammals is another topic for further investigation. The current results do provide a new perspective on the function of this critical area: that promotion and suppression of behavior can be driven by the same cortical structure.

**COMPETING FINANCIAL INTERESTS**
The authors declare no competing financial interests.

1. Sasaki, K., Gemba, H. & Tsujimoto, T. *Brain Res.* **495**, 100–107 (1989).
2. Chambers, C.D., Garavan, H. & Bellgrove, M.A. *Neurosci. Biobehav. Rev.* **33**, 631–646 (2009).
3. Coxon, J.P., Stinear, C.M. & Byblow, W.D. *J. Neurophysiol.* **95**, 3371–3383 (2006).
4. Ebbesen, C.L., Doron, G., Lenschow, C. & Brecht, M. *Nat. Neurosci.* **20**, 82–89 (2017).
5. Hill, D.N., Curtis, J.C., Moore, J.D. & Kleinfeld, D. *Neuron* **72**, 344–356 (2011).
6. Brecht, M., Schneider, M., Sakmann, B. & Margrie, T.W. *Nature* **427**, 704–710 (2004).
7. Haiss, F. & Schwarz, C. *J. Neurosci.* **25**, 1579–1587 (2005).
8. Matyas, F. *et al. Science* **330**, 1240–1243 (2010).
9. Mao, T. *et al. Neuron* **72**, 111–123 (2011).
10. Guo, J.Z. *et al. eLife* **4**, e10774 (2015).
11. Georgopoulos, A.P., Schwartz, A.B. & Kettner, R.E. *Science* **233**, 1416–1419 (1986).
12. Graziano, M.S., Taylor, C.S. & Moore, T. *Neuron* **34**, 841–851 (2002).
13. Zagha, E., Ge, X. & McCormick, D.A. *Neuron* **88**, 565–577 (2015).
14. Hill, D.N., Bermejo, R., Zeigler, H.P. & Kleinfeld, D. *J. Neurosci.* **28**, 3438–3455 (2008).
15. Chakrabarti, S. & Schwarz, C. in *Sensorimotor Integration in the Whisker System* (ed. P. Krieger & A. Groh) 129–148 (Springer New York, 2015).

# Correlations demystified

Peter E Latham

**An elegant study answers a long-standing question: how do correlations arise in large, highly interconnected networks of neurons? The answer represents a major step forward in our understanding of spiking networks in the brain.**

One of the most notable features of single neurons in the mammalian brain is the highly variable rates at which they spike. A neuron might emit 8 spikes on one trial, 5 on another and 10 on a third, even when conditions on each trial are virtually identical. At the population level, this variability tends to be correlated: upward fluctuations in the activity of one neuron are often mirrored by upward fluctuations in other, nearby neurons, and similarly for downward fluctuations. For example, on the same three trials, a second neuron might emit 7, 3 and 12 spikes.

These correlated fluctuations are termed noise correlations, and there are two reasons to care about them. First, as their name implies, they are mainly a nuisance: noise correlations can greatly reduce the amount of information in a population, in many cases by orders of magnitude[1]. Second, and related, computations must be efficient in the face of these noise correlations. So to understand computations in the brain, it is essential to understand how noise correlations arise. In this issue of *Nature Neuroscience*, Rosenbaum *et al.*[2] show, using arguments that are both elegant and simple, that correlations must arise when external input to a network varies over a length scale that is small compared to that of its lateral connectivity.

To put this work in context, flash back to 1998, when van Vreeswijk and Sompolinsky published what has become the de facto standard model of large networks of spiking neurons[3]. An underappreciated assumption in that model was that connectivity was so sparse that correlations were eliminated

altogether. This allowed the rigorous development of an elegant theory describing large networks of spiking neurons. Of course, because of the ultrasparse assumption, this may seem like a classic case of looking where the light is. However, in a rare stroke of luck for theoreticians, the analysis gave very accurate predictions even when the sparseness assumption was violated. But why the theory worked so well remained a mystery that was not solved for another 14 years, when Renart and colleagues showed that it is the interplay of excitation and inhibition that causes correlations to dynamically cancel, making them near zero on average[4].

While the result of Renart *et al.*[4] was extremely important, it brought a new mystery. In some areas of the brain correlations are indeed near zero on average[5,6]. However, that's the exception, not the rule: in most areas correlation coefficients hover around 0.1–0.2 (refs. 5,6), appreciably larger than the prediction of Renart *et al.*[4]. This mystery has now been solved as well, by Rosenbaum *et al.*[2]. They showed that for networks with spatially inhomogeneous connectivity (connectivity that falls off with distance, as is found in the brain), relatively large correlations should emerge if the input is spatially localized. The resulting spatial profile of correlations, large for nearby neurons and small for more distant neurons, is qualitatively similar to what is found in the brain[7]. Quantitatively, however, it differs: the model of Rosenbaum *et al.*[2] predicts that correlations should average to zero, something that is not typically seen. We'll return to that point shortly, but first we'll explain what they did.

As is typical in the analysis of neuronal networks in mammalian cortex, Rosenbaum *et al.*[2] took advantage of the fact that each

neuron makes a large number of connections, on the order of 1,000. They then used one of the favorite tricks of physicists: when they see a moderately large number, they pretend that it is very, very large. When the external input is fixed, networks of this type can fire at the kinds of low rates seen in the brain (a few hertz), with near-Poisson variability[8]. But under realistic conditions input is never fixed. Instead, it varies from trial to trial, by enough to make any single neuron, in isolation, change its firing rate by at least a few hertz. When the variability is correlated across the population, one might think that the firing rates of all the neurons in the network would change together, leading to large correlations in their firing rates. If the neurons were disconnected, that would indeed be the case. However, for highly connected networks of excitatory and inhibitory neurons, it is not. That's because for such networks to remain stable (that is, to not exhibit runaway excitation) they must be inhibition-dominated, in the sense that small increases in excitatory firing rate cause a larger increase in inhibitory firing rate[9]. This has an interesting corollary: if one were to increase the firing rate of all excitatory neurons in a network, that would cause a sufficient increase in inhibitory firing rate that the change in synaptic drive to every neuron in the network would be negative.

To make this explicit, we plot the firing rate of a test neuron, defined as a neuron with typical connectivity, as a function of the average excitatory firing rate in the network (**Fig. 1a**). There are multiple curves in this plot; focusing for now on the thick one, such a curve could be produced experimentally by controlling the firing rate of every excitatory neuron (except for the test neuron) and monitoring the firing

Peter E. Latham is in the Gatsby Computational Unit, University College London, London, UK.
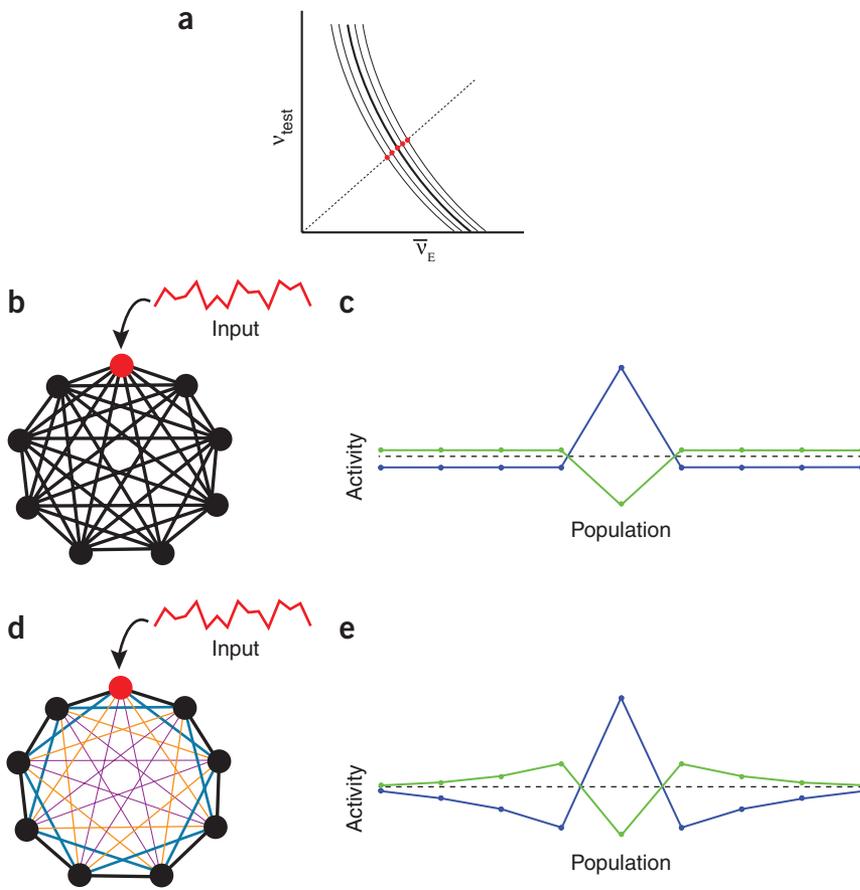e-mail: pel@gatsby.ucl.ac.uk

**Figure 1** Firing rates and correlations in large, highly connected networks of excitatory and inhibitory neurons. (**a**) Firing rate of a test neuron ($\nu_{test}$) versus the average excitatory firing rate ($\bar{\nu}_E$). The thick line is the firing rate curve when the input is fixed and static; the thin lines are firing rate curves on trials with different amounts of external input. The intersection of the firing rate curves with the dashed 45° ($\nu_{test} = \bar{\nu}_E$) line corresponds to network equilibria (as indicated by the red dots). For ease of visualization, the firing rate curve is shallow here; in high-connectivity networks it would be much steeper. This would bring the thin lines very close together (because external drive shifts the firing rate curve up and down), so there would be very little variation in firing rate as input changed. (**b**) A randomly connected network arbitrarily broken into nine populations, only one of which (red) receives external input (which is applied uniformly to all neurons in the red subpopulation). Connection strengths are negative, mirroring the fact that networks of excitatory and inhibitory neurons are effectively inhibitory, and they are uniform because the network is randomly connected. (**c**) Average population activity on two trials relative to baseline (dashed line). For the blue points, the preferred population receives positive input; for the green points, it receives negative input. The center data points correspond to activity of the red subpopulation. All neurons within each subpopulation are correlated, but the neurons in the red subpopulation are anticorrelated with the rest of the neurons. (**d**) As in **b**, except the connectivity falls off with distance, as indicated by the progressively thinner lines connecting more distant populations. Black lines, closest connections; blue lines, moderately close connections; orange lines, moderately distant connections; purple lines, distant connections (**e**) Activity on two trials, as in **c** (with the same color code). Because connectivity is short range, activity is a decreasing function of distance, leading to correlations that fall off with distance.

high-connectivity regime, even highly correlated trial-to-trial variability has virtually no effect on the average firing rate in the network, and so the average correlation coefficient is very small.

This is, essentially, the result of Renart et al.[4]: the dynamic cancellation referred to above is what produces the steep, downward sloping firing rate curve and so reduces fluctuations in firing rates. To see what Rosenbaum et al.[2] added, assume that some fraction of the excitatory neurons, the preferred population (say, 1/9), receives input that varies from trial to trial. Again, high connectivity clamps the mean firing rates. Consequently, whenever the activity of the preferred population increases, the activity of the other 8/9 of the excitatory neurons (the nonpreferred populations) will decrease, and vice versa. Thus, as a result of variability in the input, neurons in the preferred population will fluctuate together and neurons in the nonpreferred population will fluctuate together—but in the opposite direction. The network is shown in **Figure 1b**, where we have arbitrarily divided the neurons into nine groups. Because the networks are inhibition dominated (**Fig. 1a**), coupling is effectively inhibitory. It is this inhibitory coupling that causes preferred and nonpreferred populations to be anticorrelated (**Fig. 1c**).

It is not hard to extrapolate from this scenario to one in which connectivity falls off with distance (**Fig. 1d**). The only real change is that the correlations are no longer long range, as distant neurons no longer have much effect on each other. This results in a tell-tale pattern of correlations: positive for nearby neurons, negative for intermediate neurons and zero for distant neurons (**Fig. 1e**). This tell-tale pattern is a key experimental prediction.

In both scenarios, because of the high connectivity, correlations average to zero— something that is not seen in the brain. How can positive correlations come about in the high-connectivity regime? There are at least four ways. First, networks of excitatory and inhibitory neurons are prone to oscillations and up–down states[9], collective activity that leads to large correlations. Second, synaptic strength can changes on slow timescales, from hundreds of milliseconds to seconds. Because the synaptic strength determines the equilibrium firing rate, this would cause slow fluctuations in the overall level of activity. Third, external input might fluctuate by an amount much larger than was assumed above. Fourth, neuromodulators might modify overall excitability. Which of these, if any, is responsible for nonzero average correlations is not clear, and this is an active area of research.

rate of the test neuron. While this has not been done, such a curve is consistent with every viable model of large, biologically plausible networks to date. There are two important features to this plot. First, the firing rate of the test neuron decreases as the average firing rate increases—a consequence of the fact that networks are inhibition dominated. Second, the curve is steep: in the high connectivity limit assumed by Rosenbaum et al.[2], it would be infinitely steep (here we drew it shallower

than it should be; otherwise, the thick and thin lines would be indistinguishable).

If the test neuron is sufficiently typical, then the equilibrium firing rate in the network occurs where the 45° diagonal crosses the firing rate curve. Trial-to-trial variability in the input would shift the firing rate of the test neuron up and down, as shown by the thin lines. However, because the lines are steep, the resulting shift in equilibrium firing rates is small. Consequently, in the

When analyzing experimental data in macaque primary visual cortex, Rosenbaum et al.[2] sidestepped the issue and took long-range fluctuations into account without explicitly considering their source. When subtracted from experimentally observed correlations, the resulting correlational structure (positive, then negative, then near zero as distance between neurons increased) was exactly as predicted.

The analysis by Rosenbaum et al.[2] was beautiful, elegant and, ultimately, straightforward: they simply extended results from randomly connected networks with high connectivity to networks in which connection probability falls off with distance; the rest was algebra (occupying 35 pages of supplementary information). And this was not just theory; the authors took the laudable additional step of comparing their results to experiments and, fortunately, finding agreement. Their analysis adds much-needed insight into the dynamics of large networks of spiking neurons—exactly the kind of insight we need if we are ever going to understand how the brain works.

How do these correlations affect the ability of networks to store information? The answer, as is typical in neuroscience, is that we don't know. The only correlations that reduce information are ones that make the noise look like the signal[10]. As shown recently, these correlations emerge naturally in circuits that receive very little information compared to their coding capacity[11]. Whether the internally induced correlations described by Rosenbaum et al.[2] also introduce such correlations is an open question, one that is likely to keep theorists busy for the foreseeable future.

1. Zohary, E., Shadlen, M.N. & Newsome, W.T. Nature 370, 140–143 (1994).
2. Rosenbaum, R., Smith, M.A., Kohn, A., Rubin, J.E. & Doiron, B. Nat. Neurosci. 20, 107–114 (2017).
3. van Vreeswijk, C. & Sompolinsky, H. Neural Comput. 10, 1321–1371 (1998).
4. Renart, A. et al. Science 327, 587–590 (2010).
5. Cohen, M.R & Kohn, A. Nat. Neurosci. 14, 811–819 (2011).
6. Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K & Josić, K. Nat. Neurosci. 19, 383–393 (2016).
7. Smith, M.A. & Kohn, A. J. Neurosci. 48, 12591–12603 (2008).
8. Goris, R.L., Movshon, J.A. & Simoncelli, E.P. Nat. Neurosci. 17, 858–865 (2014).
9. Latham, P.E., Richmond, B.J., Nelson, P.G. & Nirenberg, S. J. Neurophysiol. 83, 808–827 (2000).
10. Moreno-Bote, R. et al. Nat. Neurosci. 17, 1410–1417 (2014).
11. Kanitscheider, I., Coen-Cagli, R. & Pouget, A. Proc. Natl. Acad. Sci. USA 112, E6973–E6982 (2015).

# Cracking the mnemonic code

Eva Zita Patai & Hugo J Spiers

**Evidence reveals that humans share remarkably similar patterns of event-specific neural activity during spontaneous spoken recall. Posterior medial cortex appears to play a key role in transforming experience into memory.**

Like fingerprints, each person's brain has a unique pattern and organization but looks generally similar to other people's. Thus, when two people experience the same event, a similar set of brain regions will be engaged, but the exact pattern of brain activity elicited will be unique to each individual. Similarly, during subsequent recall it is thought that each person's unique pattern of activity will re-emerge[1] but will become even more differentiated by the fact that people tend to recall events in slightly different ways. Thus, it would be surprising if one person's brain activity pattern during spontaneous recall provided a better match to the activity pattern in another person's brain during recall of the same event than the match between encoding and recall within an individual. Yet this is exactly what Chen and colleagues show in this issue[2]. Their findings imply that humans share a remarkably similar neural circuit for transforming experience into memory.

Much of the basis for this new research comes from previous seminal work showing

Eva Zita Patai and Hugo J. Spiers are at the Institute of Behavioural Neuroscience, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, UK.
e-mail: e.patai@ucl.ac.uk or h.spiers@ucl.ac.uk

that neural activity during perception of stimuli is preserved across individuals[3] and is related to the content, not just the physical form, of the stimuli[4]. It has also been shown that patterns of activity during perception are reactivated during recollection in a sensory-specific manner, such that visual and auditory memories are represented in their respective sensory cortices[5]. As a final piece to the puzzle, Bird and colleagues have shown that encoding and recall activity of movie scenes overlap in the posterior medial cortex (PMC), such that higher correlation between these neural patterns predicts better recall performance[6]. This implies that the PMC is involved in consolidation and reinstatement of memories, possibly through its connections with the medial temporal lobe and other memory-related structures.

Chen and colleagues reasoned that if neural activity is preserved across mental states (perception and memory) within individuals and representations during perception are preserved across individuals, then neural activity during recall should also be preserved across individuals. To test this, they recorded functional magnetic resonance imaging (fMRI) data from participants as they viewed a 50-min movie from the BBC TV series *Sherlock*, in which the fictional detective Sherlock Holmes solves murder mysteries in a twenty-first-century London. Subsequently, during fMRI participants were asked to verbally recall the events in the movie. Using a method known as representational similarity analysis[7], the research team then examined, scene-by-scene, whether similar brain activity patterns occurred within and between individuals during both tasks: move viewing and recall.

The analysis revealed that the PMC, medial prefrontal cortex, right anterior temporal lobe, right inferior frontal gyrus and higher level visual areas had scene-specific patterns of activity when directly contrasting recall-to-recall across individuals (**Fig. 1**). This effect was robust and was not dependent on changes in acoustics of the verbal recall or differences in recall length between participants, further underscoring the notion that the neural representations of these recalled events had undergone some systematic transformation. Because everyone saw the same movie, unsurprisingly, brain activity patterns in these areas during movie viewing served to classify scenes far above chance. Impressively, the authors now reveal that classification accuracy of individual scenes during spoken recall was also substantially above chance. Thus, despite all the idiosyncrasies in how people spontaneously recall different movie scenes, there was a remarkably similar pattern of neural activity across brains for the same events.