The Asynchronous State in the Cerebral Cortex

Alfonso Renart,[1*] Jaime de la Rocha,[1,2*] Peter Bartho,[1,3] Liad Hollender,[1]
Néstor Parga,[4] Alex Reyes,[2] Kenneth D. Harris[1,5]

[1] Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102, USA.
[2] Center for Neural Science, New York University, New York, NY 10003, USA.
[3] Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest 1083, Hungary.
[4] Departamento de Física Teórica. Universidad Autónoma de Madrid, Madrid 28049, Spain.
[5] Similow Research Center. New York University Medical School, New York 10016, USA.

[*] These authors contributed equally to this work.

The magnitude of spiking correlations within a local cortical circuit is thought to be set by the fraction of presynaptic inputs shared by neurons in the network. Here we show that this is not the way activity is organized in balanced recurrent networks of excitatory and inhibitory neurons. Instead, the network settles into an asynchronous state where spiking correlations are marginal despite large amounts of shared input. Spontaneous fluctuations in the activity of the excitatory and inhibitory populations in this state accurately track each other, generating negative correlations in the synaptic currents which precisely cancel the effect of shared input. In agreement with the theory, we validate experimentally that correlations are indeed marginal in the anesthetized rodent cortex. Our results provide a theoretical foundation to a large body of previous work and suggest a re-examination of the constraints that circuit architecture impose on information processing.

Pairs of nearby cortical neurons share a significant fraction of their pre-synaptic inputs, in part due to the relatively large local connection probability in the cortex (*1–3*). Temporal fluctuations in the spiking activity of these shared inputs induces correlated fluctuations in the synaptic currents to their post-synaptic targets, leading to correlated spiking (*4, 5*). Since the typical magnitude of spiking correlations measured *in vivo* (*6–9*) is similar the amount of correlation produced by realistic amounts of shared input in feed-forward networks (*5, 10, 11*), spiking correlations have been assumed to reflect shared input, a hardwired feature of the cortical anatomy. Although measured correlations are typically not too strong (*6–9*), they have a large impact on the way information in the neural activity can be decoded by downstream targets (*12, 13*). This has lead to the suggestion that the anatomy of a cortical micro-circuit, through its effect on spiking correlations, could severely limit the efficiency of the whole organism for performing sensory discriminations (*5, 7, 14*). It remains unknown, however, whether the intuitions gained in simple feed-forward networks extend to more realistic recurrent architectures.

In a randomly connected recurrent network, pairs of neurons can become correlated by *(i)* a direct connection (orange pair in Fig.1A), *(ii)* shared inputs (pink pair) and *(iii)* non-shared correlated inputs (blue pair). We first examine the effect of these different sources of correlation for a pair of neurons receiving N feed-forward excitatory (E) inputs. The relationship between the fraction of shared inputs p and the output correlation $r_{\rm out}$ is roughly linear with slope smaller than one (*5, 10, 11*) (Fig. 1B). In contrast, the relationship between the correlations between inputs $r_{\rm in}$, and $r_{\rm out}$ has initially a very high gain (Fig. 1C). This amplification occurs because the correlation c of the two input currents is approximately equal to

$$c \sim p + N r_{\rm in} \tag{1}$$

when p and $r_{\rm in}$ are small (*15*). Although this phenomenon can lead to a synchrony explosion (Fig. 1D), as observed in multi-layered feed-forward networks (*16*), it can also have a potentially strong desynchronizing effect when inhibitory (I) pre-synaptic inputs are taken into account, because positive correlations between E and I cells contribute negatively to the current correlation c (Fig. 1E-F). In fact, we now show that in recurrent networks, neural activity self-organizes in a way in which the net effect of spiking correlations on c is *negative*, precisely canceling the hardwired positive component of c due to shared inputs.

We studied the simplest model of a recurrent, randomly-connected network consistent with the low firing rate and high temporal irregularity exhibited by cortical neurons (Fig. 2A).

Whereas a similar model had been previously analyzed in the sparse connectivity limit in which neurons share no inputs (*17*), we assume that connectivity in our network is *dense*, i.e., the probability of connection p (equal to the mean fraction of shared input) takes a fixed realistic value (e.g. $p = 0.2$ as in Fig. 2). The network consists of two populations of N excitatory and inhibitory binary neurons. Both types of neurons receive excitatory projections (with the same probability p) from an external (X) population of N cells assumed, for simplicity, to fire independently (*18*). Synaptic connections are *strong*, i.e., although each neuron receives an average number of inputs proportional to N, the number of excitatory inputs needed to make the cell fire is only proportional to $\sqrt{N}$ (*17*). We developed an analytical theory that describes the distribution of firing rates and correlations in the network in stationary conditions (*18*). The theory shows that, if inhibition is powerful and not too slow, the network settles into a stationary state of low irregular firing resulting from a dynamic balance between the large net excitatory and inhibitory drives to the cells (Fig. 2B), as described previously in sparse networks (*17*, *19*). Furthermore, the average correlation in the neuronal firing across the population $\bar{r}$ in this state is very weak. In fact, if one considers networks of different sizes, $\bar{r}$ decreases in a way inversely proportional to N. (Fig. 2C white dots). In this type of network states, formally called asynchronous (*20*), correlations are marginal in the sense that they do not fundamentally constrain how well the average firing rate of the neuronal population can be estimated (Fig. S1).

Since the synaptic current to each cell consists of an excitatory and an inhibitory component, the average current correlation across cell pairs, c, can be decomposed in three terms c_{EE}, c_{II} and c_{EI}. In the asynchronous state in our network, the magnitude of these three terms is large and does not decrease with N (Fig. 2C, colored dots and Fig. 2F *i*), due to the amplification of weak firing correlations (equation 1). How is it possible that asynchronous firing co-exists with strongly correlated input currents? The answer lies in the way the instantaneous activities of the excitatory and inhibitory populations ($m_E(t)$ and $m_I(t)$) self-organize in the asynchronous state. As shown in Fig. 2D, $m_I(t)$ tracks $m_E(t)$ with a small lag ($EI\text{-}Lag$), and they both closely follow the external instantaneous activity $m_X(t)$. In larger networks, tracking becomes more accurate, and it becomes perfect in the large N limit,

$$m_E(t) = A_E\, m_X(t) \qquad\qquad m_I(t) = A_I\, m_X(t) \tag{2}$$

where A_E and A_I are constants which depend on the network architecture (*18*). Tracking occurs because, when the connectivity is strong and dense, even the small spontaneous fluctuations in instantaneous activity, of order $1/\sqrt{N}$, are large enough to recruit inhibitory feedback. Thus,

tracking reflects the constant suppression of small fluctuations in the excitatory population by inhibition. Tracking yields a nearly instantaneous balance between the E and I current components, a phenomenon that has been observed experimentally (*21–23*). We propose that it should be present even in conditions of tonic firing, and that its functional role is to prevent epileptic-like activity arising from a synchrony explosion in the local circuit.

It is simple to show mathematically that tracking (equation 2), is equivalent to the precise cancellation of the three components of the average current correlation (*24*)

$$c_{EE} + c_{II} + 2c_{EI} \sim 1/\sqrt{N} \tag{3}$$

The term c_{EE} is positive and contains both the effect of shared inputs and the correlations between E cells (equation 1), and similarly for c_{II}. Correlations between E and I cells generated by tracking (Fig. 2E) lead to a large and negative c_{EI} which results in the cancellation shown in equation 3 (Fig. 2E and 2F *ii*).

Even after the cancellation, the instantaneous current correlation c is still larger than the correlation in firing $\bar{r}$ (Fig. 2C). This is possible because the transformation between synaptic current and firing activity is not instantaneous: Since neurons effectively integrate their synaptic input, the instantaneous correlation r is related to the area under the current cross-correlogram over a window of the order of the neuron's time-constant (*18*). Since synaptic interactions are strong, changes in the activity of only a fraction $1/\sqrt{N}$ of the cells are enough to produce a noticeable change in synaptic input. This sets the $EI\text{-}Lag$ and the magnitude and width of the total current cross-correlogram to be of that same magnitude (Fig. 2E). The area under the current cross-correlogram is thus $1/N$, as required for asynchronous firing (Fig. 2F *iii*). This mechanism, summarized in Fig. 2F, is a robust dynamical phenomenon and does not require fine-tuning of any network parameter. It constitutes the first self-consistent description of neural activity of an anatomically plausible recurrent cortical circuit. .

A signature of the asynchronous state described here is that the distribution of firing correlations r across pairs is *wide*, with a standard deviation σ_r much larger than its mean $\bar{r}$ for large networks (σ_r decays only as $1/\sqrt{N}$) resulting in similar numbers of positively and negatively correlated pairs in the network (Fig. 2G). This is due to the fact that the hard-wired sources of correlation (i.e. the presence or absence of a direct connection between two cells and variability in the fraction of shared input) have a strong impact on r (of order $1/\sqrt{N}$) and therefore generate large heterogeneity from pair to pair.

We found that a similar mechanism to the one just described in the binary network is also at

work in more biologically plausible network of spiking neurons. We simulated large networks of randomly connected conductance-based *integrate-and-fire* neurons, with parameters chosen so as to produce a balanced state where neurons fired irregularly (shared fraction $p = 0.2$; Fig. 3A; (*18*)). We quantified spiking correlations using the correlation coefficient of the spike count r in windows of 50 ms (*25*). As in the binary network (Fig. 2G), the distribution of r is wide (Fig. 3B), with an extremely low average $\bar{r} = 0.001$. To determine whether a cancellation between the components of the current correlation (equation 3) underlied the small value of $\bar{r}$ observed, we injected different levels of DC current to cell pairs in which we had disabled the spiking mechanism. The range of current levels was adjusted to isolate the EPSP- and the IPSP-components near their respective reversal potentials (*21, 22*), or combinations of EPSPs and IPSPs at intermediate potentials. The correlation between isolated EPSPs (green, Fig. 3C) and between isolated IPSPs (red) was much larger than the correlation measured at rest (i.e. no injected current), due to a cancellation with the large and negative correlation between EPSPs and IPSPs (gold). Across the whole range of holding potentials, the minimum correlation was achieved for the pairs measured at the restintg potential (Fig. 3D, solid black dot).

To investigate whether a similar phenomenon can occur in cortical circuits *in vivo*, we analyzed neuronal population recordings collected with silicon microelectrodes in somatosensory and auditory cortices of urethane-anesthetized rats (*18*). Results from both areas were pooled together, as the analysis did not reveal significant differences between them. Under urethane, cortical activity displays spontaneous changes in state homologous to those seen during sleep (*26, 27*). Alternations take place between an 'activated' state of tonic activity resembling REM (ACT, blue, Fig. 4A-B), and an 'inactivated' state characterized by global fluctuations in population activity (Up-Down transitions), resembling slow-wave-sleep (InACT, red, Fig. 4B-C). During the ACT period, correlations were, on average, remarkably small, and the correlation histogram was wide, in agreement with the theory ($\bar{r}$ = 0.0075; 46.6% of negatively correlated ($r < 0$) pairs (Fig. 4B). These values were typical of ACT state correlations across different animals (across n =11 recording sessions in 9 rats, median of $\bar{r}$ = 0.0053; [0.0024 : 0.0094] interquartile range; Fig. 4E). Across all 30,772 pairs recorded the ACT state, $\bar{r}$ = 0.0052, and 47.05% had $r < 0$. This behavior did not depend strongly on the time-scale at which correlations were measured (Fig. S2). Although the mean correlation $\bar{r}$ in the ACT state was systematically very low, in all experiments $\bar{r}$ was positive and significantly different from zero ($p < 0.005$ against a null hypothesis given by histograms of jittered surrogates; Fig. 4E). A considerable minority of both positive and negative correlations were statistically significant

(compare blue and gray in Fig. 4B; Fig. 4F). Pairs with significant and negative (positive) correlations show clear troughs (peaks) in their cross-correlograms on average (Fig. 4B insets, see Fig. S3 for individual examples). Finally, the correlation histogram during the ACT state is still wide even if one only considers neurons recorded in the same shank (Fig. S4). In contrast to these results, during InACT periods, the distribution of correlations r in similar populations of neurons was consistently biased towards positive values (across n =7 sessions in 5 rats the median of $\bar{r}$ was 0.0953, [0.088 : 0.109] interquartile range; Fig. 4C-E). Across all 18,916 pairs in this condition, $\bar{r}$ = 0.096, and 8.96% had $r < 0$. However, although the whole population was co-modulated by the slow oscilation (resulting in the positive bias of r), spiking was still very weakly correlated, on average, during Up-states (*28*): removal of Down-states from the recorded spike trains (*18*) results in correlation histograms very similar to those during the ACT state (median of $\bar{r} = 0.0163$, [0.0066 : 0.023] interquartile range; Fig. 4C-E). In this condition, $\bar{r}$ = 0.0136 across all 18,916 pairs, with $r < 0$ in 44.93% of them.

Our findings establish that the dynamics of recurrent circuits results in an active decorrelation of the synaptic inputs to the neurons in the network. In agreement, theory, network simulations, and *in vivo* population recordings all show that densely connected neural circuits display extremely low spiking correlations on average. Why, then, are noise correlations such a common feature of neuronal recordings? First, cortical circuits do exhibit a variety of states with temporally structured global fluctuations under some types of anesthesia or during quiet wakefulness (*22, 26, 29–32*). Although we have shown that spiking is also very weakly correlated within periods of sustanined activity in a state with slow global fluctuations (Fig. 4D-F (*28*)), the population structure of spiking correlations in the presence of punctuated activity fluctuations (*22, 30, 32*) remains to be elucidated. More importantly, measured correlations might reflect, not the way cortical circuits operate, but rather our incomplete understanding of what cortical circuits do. Neuronal activity, even in primary sensory areas, is often affected not only by sensory stimuli but by cognitive and behavioral variables (*33–35*) which are of very difficult experimental access. If these variables have a global effect on the activity of the population being monitored, failure to condition on their precise value on every trial will lead to measurements of positive correlations, even if all neurons are conditionally independent. Nevertheless, recent data suggests that extremely weak correlations are also observed in evoked responses from the visual cortex of awake behaving monkeys (*36*). If, as we are suggesting, cortical networks are endowed with the capacity of actively decorrelating their spiking activity, then the cortical circuitry does not necessarily constitute an irreducible source of 'noise'.

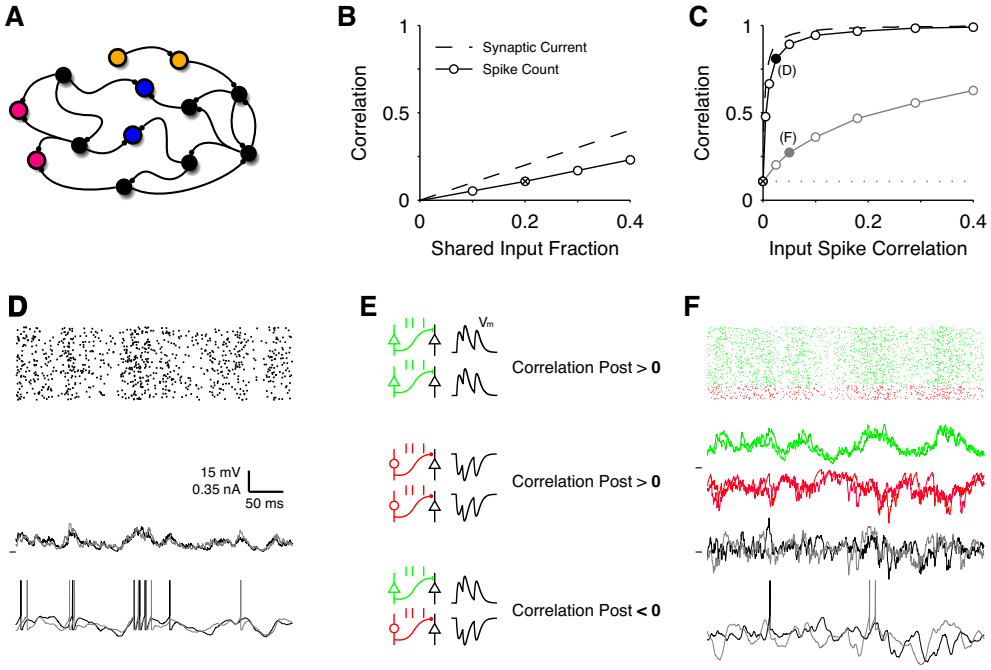

Figure 1: Sources of correlation in a neural circuit. (A) Schematic illustration of three mechanisms that lead to correlations in a recurrent network: A direct synaptic connection (orange pair), shared inputs (pink pair), and non-shared correlated inputs (blue pair). **(B)** Correlation coefficient of the synaptic current, c (dashed), and output spikes, r_{out} (dots, count window T = 50 ms) as a function of the fraction of shared inputs p. Each post-synaptic cell received 250 excitatory (E) Poisson spike trains. The correlation c is equal to p. **(C)** Correlations c (dashed) and r_{out} (dots) as a function the input spike correlation r_{in} at fixed p = 0.2. Each cell receives 250 excitatory Poisson spike trains with correlation r_{in} (black) (*18*). Point marked with a cross in (B) and (C) is identical. Adding an additional 60 inhibitory (I) input spike trains with identical statistics and correlations decreases the gain between r_{in} and r_{out} (gray). **(D)** Input raster (top), synaptic current (middle) and membrane potential (bottom) of the post-synaptic pair for the point marked (D) in panel (C). Many weakly correlated inputs ($r_{\text{in}} = 0.025$) lead to strongly correlated synaptic currents and output spikes. **(E)** Schematic illustration of the effect of positive input correlations r_{in} on the correlation of a pair of post-synaptic neurons. Correlations between E inputs (green) or between I inputs (red) lead to positive correlations. Correlations between E and I inputs leads to negative correlations. **(F)** Same as (D) but for the point marked (F) in panel (C). E and I spikes and synaptic currents are shown separately in green and red respectively. Correlations between E and I inputs ($r_{\text{in}} = 0.05$) reduces the correlation of the total synaptic input.

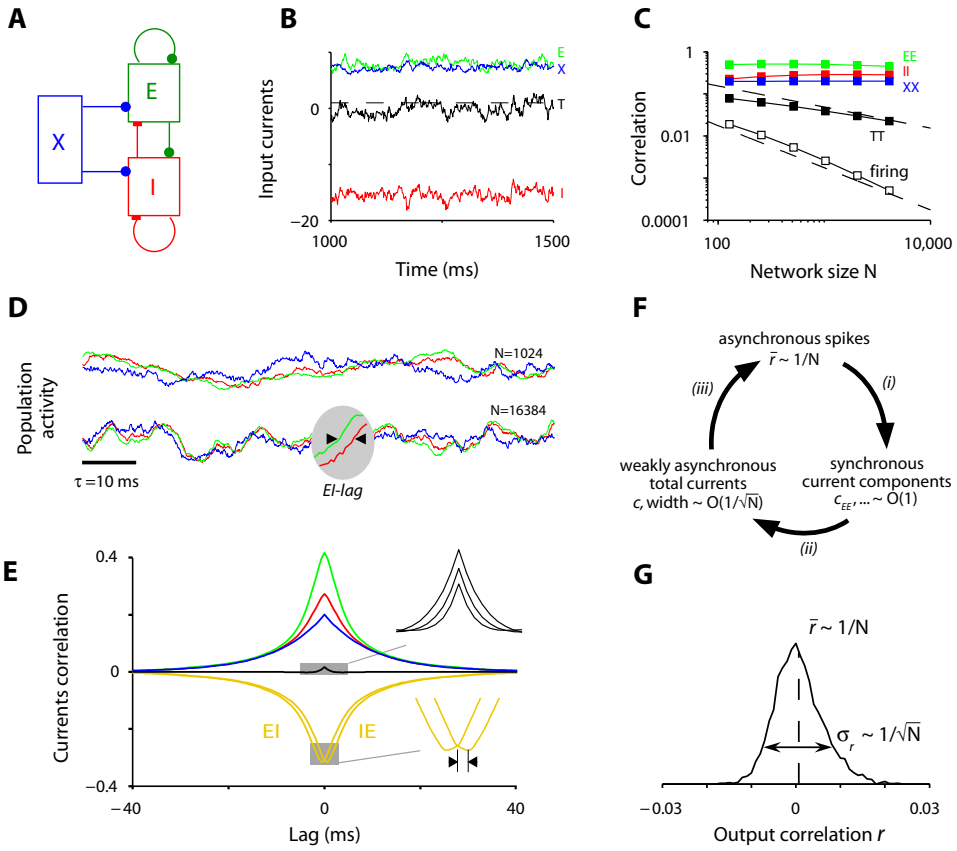

Figure 2: Asynchronous activity in a binary recurrent network. **(A)** Network scheme showing the connectivity between the excitatory (E), inhibitory (I) and external excitatory (X) populations. **(B)** Example current traces showing the balance between the large excitatory and inhibitory drives to a neuron. Dashed line represents threshold. **(C)** Population-averaged correlation coefficients of the firing activity ($\bar{r}$, white), total current (c, black) and current components for networks of different sizes N. Dashed lines show $1/\sqrt{N}$ and $1/N$ scaling for comparison. **(D)** Instantaneous population activities (transformed to z-scores) for the E, I and X populations for show that tracking becomes more accurate with increasing N. Instance of lag between E and I activities magnified in the inset. **(E)** Population-averaged correlograms of the current components. Top inset shows the decrease in magnitude and width of the cross-correlogram of the total current for N = 2048, 4096 and 8192. Bottom inset: offset of the correlation between E and I currents (arrowheads) quantifies EI-Lag. Color code as in (C). **(F)** Description of the asynchronous self-consistent solution: *(i)*, asynchronous firing yields correlated current components c_{EE} and c_{II} (equation 1) which *(ii)* are cancelled by strong negative correlations due to tracking of spontaneous fluctuations (D), resulting in total current correlations of magnitude and time-scale $\sim 1/\sqrt{N}$. *(iii)*, integration of the total currents leads to asynchronous activity. **(G)** The histogram of correlations for excitatory pairs ($N = 4096$) is wide: $\sigma_r \gg \bar{r}$.

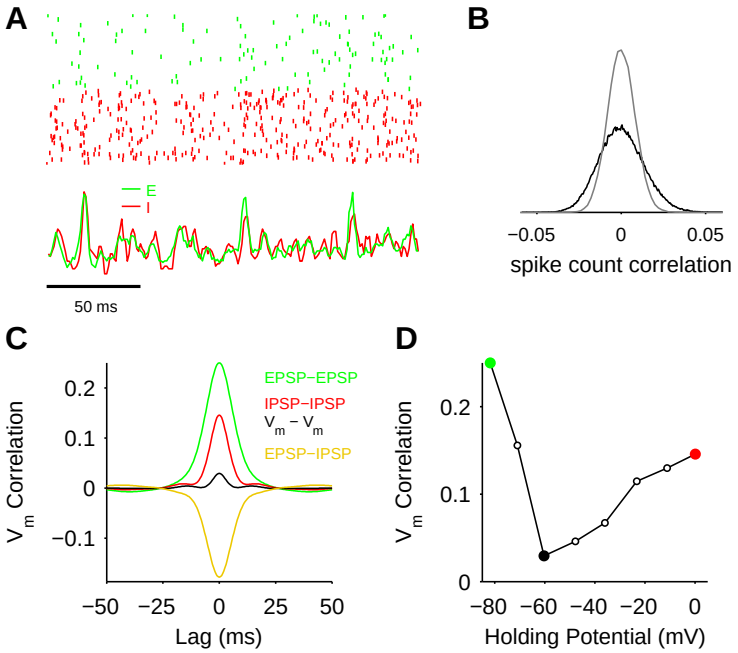

Figure 3: Cancelation of correlations in a recurrent network of spiking neurons (A) Raster (top) of 500 excitatory (green) and inhibitory (red) neurons in a conductance-based integrate-and-fire network. Bottom curves show tracking of instantaneous population activities (transformed to z-scores, bin size 3 ms). Average firing rate of E and I cells were 1 and 3.6 spike/s, respectively. **(B)** Histogram of spike count correlations in the network is wide (black, count window 50 ms). Histogram of jittered spike trains is also shown (gray). **(C)** Average membrane potential cross-correlograms of pairs in which the spiking mechanism was inactivated. DC current was injected to isolate EPSPs (green) or IPSPs (red) in both cells, or EPSPs for one cell and IPSPs for the other (gold). The black curve is from pairs at resting potential (which were not injected any current). **(D)** Peak height of the membrane potential cross-correlogram as a function of the average holding membrane potential of both neurons in the pair. Green and red filled dots correspond to pairs held at the reversal of inhibition and excitation and the black filled dot corresponds pairs at rest.

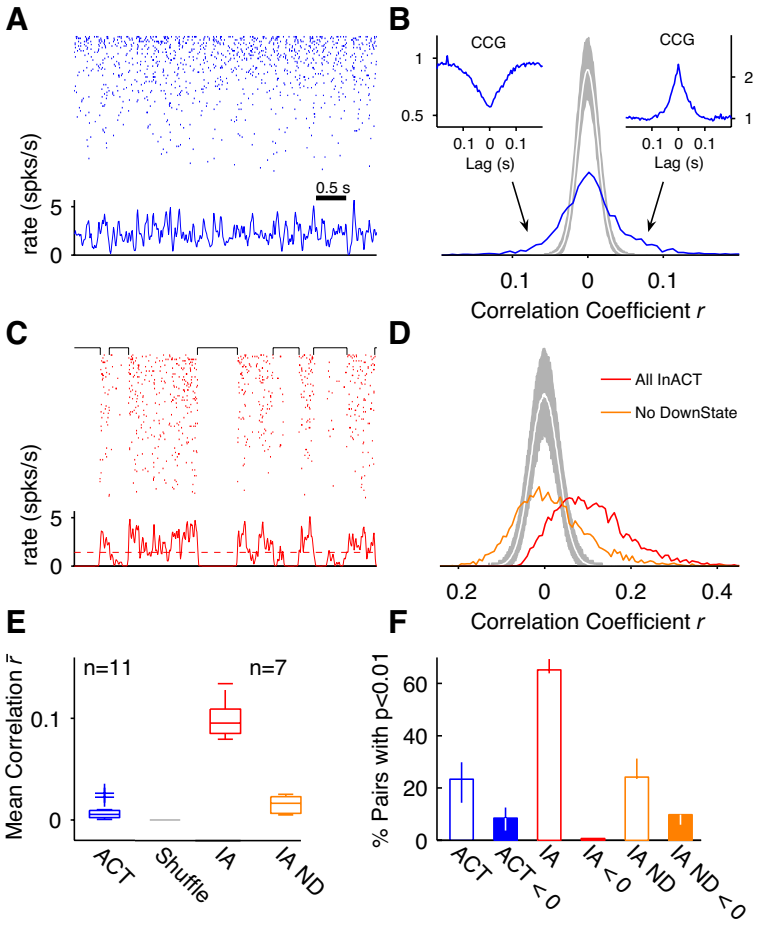

Figure 4: Marginal correlations in the activated cortex *in vivo*. (A) Raster (top) and instantaneous population activity (bottom) for a population of 100 simultaneously recorded neurons over a 5 s period of cortical activation **(B)** Histogram of correlations of the population in (A) for a single experiment (640 s of ACT) is wide. The white curve is the histogram of the spike-jittered ACT data (mean over 500 surrogate sets; gray shade 95% confidence interval (*18*)). Insets show average raw cross-correlograms of all negatively (left) and positively (right) significantly correlated ($p<0.01$) pairs. **(C-D)** Same as (A-B) for the same population of cells during 296 s of InACT. Black brackets at the top in (C) indicate Down-States. Dotted line, threshold for Down-State detection (*18*). Histogram of correlations during InACT is biased towards positive values (red). Removing Down-state periods from the data set largely eliminates the positive bias (orange). **(E)** Box-whisker plots showing the distribution of mean correlations across experiments for different conditions. **(F)** Median fraction of significantly ($p<0.01$) correlated pairs (empty) and of significantly and negatively correlated pairs (solid) across experiments. Error bars represent interquartile range.

References and Notes

1. C. Holmgren, T. Harkany, B. Svennenfors, Y. Zilberter, *Journal of Physiology* **551**, 139 (2003).

2. S. Song, P. J. Sjostrom, M. Reigl, S. Nelson, D. B. Chklovskii, *PLoS Biol* **3**, e68 (2005).

3. A.-M. Oswald, B. Doiron, J. Rinzel, A. Reyes, *J. Neurosci.* In press.

4. G. P. Moore, J. P. Segundo, D. H. Perkel, H. Levitan, *Biophysical Journal* **10**, 876 (1970).

5. M. N. Shadlen, W. T. Newsome, *J. Neurosci.* **18**, 3870 (1998).

6. T. J. Gawne, B. J. Richmond, *J. Neurosci.* **13**, 2758 (1993).

7. E. Zohary, M. N. Shadlen, W. T. Newsome, *Nature* **370**, 140 (1994).

8. E. Vaadia, *et al.*, *Nature* **373**, 515 (1995).

9. D. Lee, N. L. Port, W. Kruse, A. P. Georgopoulos, *J. Neurosci.* **18**, 1161 (1998).

10. M. D. Binder, R. K. Powers, *J Neurophysiol* **86**, 2266 (2001).

11. J. de la Rocha, B. Doiron, E. Shea-Brown, K. Josic, A. Reyes, *Nature* **448**, 802 (2007).

12. L. Abbott, P. Dayan, *Neural Computation* **11**, 91 (1999).

13. H. Sompolinsky, H. Yoon, K. Kang, M. Shamir, *Phys. Rev. E* **64**, 051904 (2001).

14. K. H. Britten, M. N. Shadlen, W. T. Newsome, J. A. Movshon, *J. Neurosci.* **12**, 4745 (1992).

15. The correlation between two quantities, each given by sum of N variables correlated by an amount ρ_{in}, out of which Np are common is equal to $(p + r_{in}(N - p))/(1 + r_{in}(N - 1))$, which is approximately equal to $p + Nr_{in}$ when $p \sim r_{in}N \ll 1$.

16. A. Reyes, *Nature Neuroscience* **6**, 593 (2003).

17. C. van Vreeswijk, H. Sompolinsky, *Science* **274**, 1724 (1996).

18. Details of the Theory along with Materials and Methods are available as Supplementary Information.

19. D. J. Amit, N. Brunel, *Cerebral Cortex* **7**, 237 (1997).

20. I. Ginzburg, H. Sompolinsky, *Phys. Rev. E* **50**, 3171 (1994).

21. A. Hasenstaub, *et al.*, *Neuron* **47**, 423 (2005).

22. M. Okun, I. Lampl, *Nat Neurosci* **11**, 535 (2008).

23. B. V. Atallah, M. Scanziani, *Neuron* **62**, 566 (2009).

24. Strictly speaking, the cancellation takes place between the current covariances. Covariances and correlations have the same scaling with N because the variance of the synaptic currents becomes independent of N for large network sizes (*18*).

25. Although the numerical value of the spike count correlation coefficient vanishes as the counting window goes to zero, its value becomes independent of the time window (except for finite size fluctuations) as the time window becomes much longer than the time-scale of the cross-correlogram of the spike trains. In our simulations, a counting window of 50 ms is long enough for this to happen.

26. M. Steriade, *Electroencephalography* (Williams and Willkins, 1999), fourth edn.

27. E. A. Clement, *et al.*, *PLoS ONE* **3**, e2004 (2008).

28. E. A. Stern, D. Jaeger, C. Wilson, *Nature* **394**, 475 (1998).

29. A. Arieli, A. Sterkin, A. Grinvald, A. Aertsen, *Science* **273**, 1868 (1996).

30. M. DeWeese, A. Zador, *J. Neurosci.* **26**, 12206 (2006).

31. D. S. Greenberg, A. R. Houweling, J. N. D. Kerr, *Nat. Neurosci.* **11**, 749 (2008).

32. J. F. A. Poulet, C. C. H. Petersen, *Nature* **454**, 881 (2008).

33. M. Brosch, E. Selezneva, H. Scheich, *J. Neurosci.* **25**, 6797 (2005).

34. M. G. Shuler, M. F. Bear, *Science* **311**, 1606 (2006).

35. J. H. Maunsell, S. Treue, *J. Neurosci.* **29**, 317 (2006).

36. A. S. Ecker, *et al.* (2009). Submitted.