Evolutionary constraints on visual cortex architecture from the dynamics of hallucinations

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Evolutionary constraints on visual cortex architecture from the dynamics of hallucinations

Thomas Charles Butler*,Marc Benayoun†,Edward Wallace‡‡,Wim van Drongelen*,Nigel Goldenfeld*, and Jack Cowan* 

In the cat or primate primary visual cortex (V1), normal vision corresponds to a state where neural excitation patterns are driven by external visual stimuli. A spectacular failure mode of V1 occurs when such patterns are overwhelmed by spontaneously generated spatially self-organized patterns of neural excitation. These are experienced as geometric visual hallucinations. The problem of identifying the mechanisms by which V1 avoids this failure is made acute by recent advances in the statistical mechanics of pattern formation, which suggest that the hallucinatory state should be very robust. Here, we report how incorporating physiologically realistic long-range connections between inhibitory neurons changes the behavior of a model of V1. We find that the sparsity of long-range inhibition in V1 plays a previously unrecognized but key functional role in preserving the normal vision state. Surprisingly, it also contributes to the observed regularity of geometric visual hallucinations. Our results provide an explanation for the observed sparsity of long-range inhibition in V1—this generic architectural feature is an evolutionary adaptation that tunes V1 to the normal vision state. In addition, it has been shown that exactly the same long-range connections play a key role in the development of orientation preference maps. Thus V1's most striking long-range features—patchy excitatory connections and sparse inhibitory connections—are strongly constrained by two requirements: the need for the visual state to be robust and the developmental requirements of the orientational preference map.

Author contributions: T.C.B., N.G., and J.C. designed research; T.C.B., M.B., E.W., W.v.D., N.G., and J.C. performed research; and T.C.B., E.W., N.G., and J.C. wrote the paper.

*To whom correspondence should be addressed. E-mail: nigel@illinois.edu.

the successful development of a stable orientation preference map (9).

Model
We model each local hypercolumn using the most general possible wiring, consistent with the physiological separation of excitatory and inhibitory neurons. Alternatively, the local structure can be viewed as many copies of a simplified version of the Douglas–Martin canonical microcircuit (16). This microcircuit is shown schematically in Fig. 1A.

We now introduce a model of V1 as a two-dimensional lattice of canonical microcircuits. To capture the lattice structure, we distinguish two length scales: the first length scale is local and is the one on which canonical microcircuits, each of which has a width of approximately 100 μm, interact with their neighbors. We model this local scale as comprising all the approximate 10 microcircuits in one V1 hypercolumn (1). Based on neuroanatomical data, our model includes excitatory and inhibitory connections between all microcircuits within a hypercolumn (17).

The second length scale is longer ranged. On this length scale, hypercolumns are coupled together by patchy excitatory connections. These connections have a range of approximately 4 mm with axonal arbors every 1 mm or so (18). Thus they are between hypercolumns. If we take the lattice spacing between individual microcircuits to be L mm, then the spacing between hypercolumns is $\sqrt{aL} mm$ where a is the number of microcircuits per hypercolumn. Thus the connections between differing lattice sites support a mixture of both local and nonlocal or lattice scale excitation and inhibition. These couplings are as shown in Fig. 1. For even longer-ranged connections, which would provide a third length scale, the best evidence to date (19) suggests that they are small world (20). However, in this paper, we do not consider their effects. The model presented here is a special case of that introduced to analyze the formation of geometric visual hallucinations (6). As we note later, it is also closely related to a model of the cortex introduced to study the development of stable orientation preference maps (9, 21).

These assumptions yield variants of the Wilson–Cowan equations (22) for local density of neural excitation of excitatory ($\varphi$) and inhibitory ($\psi$) neurons (see SI Text). On length scales much greater than the lattice scale, the Wilson–Cowan equations reduce to the partial differential equations

$$\partial_t \varphi = -\alpha_{EE} \varphi + (1 - \varphi) f_E(\Delta \varphi) \quad \partial_t \psi = -\alpha_{II} \psi + (1 - \psi) f_I(\Delta \psi)$$

[1]

with currents given by

$$s_E = w_{EE}(1 + g_{EE}^1 \Delta \varphi) \psi - w_{EI}(1 + g_{EI}^1 \Delta \varphi) \psi + h_E$$

$$s_I = w_{IE}(1 + g_{IE}^1 \Delta \psi) \varphi - w_{II}(1 + g_{II}^1 \Delta \psi) \varphi + h_I.$$  

[2]

The functions $f_I$ and $f_E$ are sigmoidal and capture the saturating response of neurons to external stimuli. The symbol $\Delta$ represents the continuous Laplacian in two dimensions. The matrix $W$ captures the local synaptic interactions as in Fig. 1A; for example, $w_{EE}$ denotes the synaptic weight of $E$–$E$ connections. The parameters $g_{ij}^1$ are effective length scales of the indicated connections, and we take $g_{EE}^1 = g_{II}^1 = g_I^1$. For the purposes of this study of spontaneous pattern formation, external stimuli $h$ are set to zero. Although most of the key conclusions of the present work are analytical, typical simulation parameters are $w_{EE} = 1.3$ with all other $w = 1$, $\alpha_{EE} = \alpha_{II} = 0.1$. The lattice scale spacing is taken in units of $\sqrt{g_{EE}^1}$. Previously cited neuroanatomical data (13) indicates that $g_I^1 \ll g_{IE}^1$.

Results
Provided couplings that promote excitation, such as $w_{EE}$, are sufficiently large compared to relaxation and inhibitory couplings, Eq. Eq. 1 with lattice scale effects neglected ($g_{II}^1 = 0$), support a stable fixed point at nonzero excitation levels of both excitatory and inhibitory neurons. When such effects are restored, normal vision corresponds to a stable homogeneous steady state. Failure of normal vision to geometric visual hallucinations occurs when the homogeneous steady state becomes unstable to spatially inhomogeneous perturbations, leading to regular pattern formation (see Fig. 2A).

The exotic spiral structure of hallucinations reported by patients (7) and shown in Fig. 2B arises from regular pattern formation through the retinotopic map. The retinotopic map transforms coordinates of excitation on V1 into visual field coordinates through an approximate logarithmic conformal map, as demonstrated experimentally in ref. 2. When regular patterns are subjected to a logarithmic conformal map, they are transformed into logarithmic spiral patterns, implying that regular pattern formation on V1 results in the logarithmic spiral patterns observed in geometric visual hallucinations (5).

In the full model above, with $g_I^1 / g_{EE}^1 \ll 1$ as physiologically motivated, geometric hallucinations occur when

$$\frac{g_I^1}{g_{EE}^1} > \frac{w_{EE} [g_I^1 + f_I(1 - \psi)]/[w_{II}]}{w_{IE}(1 - \psi)f_E[w_{IE}]} + O\left(\frac{g_I^1}{g_{EE}^1}\right).$$

[3]

where all functions, derivatives, and concentrations of firing in inhibitory neurons are evaluated at the homogeneous fixed point (see SI Text). Such conditions may possibly be achieved through the effects of hallucinogenic drugs (5, 7).

How would V1 behave if extensive lattice scale $I$–$I$ connections were present? This extensive lattice scale inhibition means that inhibitory activity in one microcircuit suppresses inhibition in distant microcircuits, leading to more excitatory activity in the distant site: $I$–$I$ connections tend to enhance negatively correlated fluctuations in activity. We introduce such connections by relaxing the requirement that $g_I^1 \ll g_{IE}^1$. Linear stability analysis shows that a sufficient condition for the normal visual state to be unstable to spontaneous spatial order is

$$\frac{g_I^1}{g_{EE}^1} > \frac{(1 - \psi)f_E[w_{EE}]}{(1 - \psi)f_I[w_{II}]}$$

[4]

(see ref. 23). The right-hand side of the above inequality is typically less than one for reasonable parameters. Because, in the

Fig. 1. (A) Simplified Douglas–Martin microcircuit. The blue circle corresponds to inhibitory neurons, and the red circle to excitatory neurons. Similarly, excitatory connections are shown as red arrows from their source and inhibitory connections as blue arrows. Inputs to the microcircuit vary and are not shown. (B) Simplified representation of patchy connections between hypercolumns in V1. Each hypercolumn is represented by a circle, with both $E$–$E$ and $E$–$I$ connections between hypercolumns indicated by red arrows.
which a single characteristic length scale emerges from a pat-visual hallucinations and contrast with the usual scenario by

Butler et al. PNAS Early Edition

absence of fine tuning, the existence of extensive lattice scale \( I-I \) connections requires the left-hand side of inequality \( 4 \) to be \( O(1) \) or greater (to avoid dangling axons between hypercolumns), we can conclude that the requirement that \( V1 \) represent visual stimuli through patterns of excitation is incompatible with such \( I-I \) connections. Only with fine tuning could such \( I-I \) connections be constructed so as to not generate spontaneous spatial order. Because changing conditions in the brain make such fine tuning impossible, the organization of \( V1 \) for robust visual function requires that lattice scale \( I-I \) connections be sparse. Experimental support for this statement is provided by data on the connections made by a special class of inhibitory cells called large basket cells (LBC), which have long axons and so can provide lattice scale inhibition (14): A recent count (13) of the number of lattice scale inhibitory synapses found on \( V1 \) LBC indicates that such connections are sparse compared with the number of lattice scale excitatory synapses found on such neurons and on \( V1 \) pyramidal neurons.

The spatial structures that occur in the presence of extensive lattice scale \( I-I \) coupling differ from those observed in geometric visual hallucinations and contrast with the usual scenario by which a single characteristic length scale emerges from a pat-tern-forming process. In such a case, the homogeneous steady state is typically unstable for some bounded region of wave vectors greater than zero. Surprisingly, in the case of lattice scale \( I-I \) connections, the instability occurs for all spatial frequencies greater than some threshold, so that the long wavelength approximation breaks down (see SI Text). These effects originate in the suppressive nature of inhibition, captured in the negative sign of the Laplacian for inhibitory connections in Eq. 2. Laplacian operators with positive signs are associated with signals from excitatory neurons. When the long-range connections are primarily from excitatory neurons, the overall sign of the Laplacian is positive, and the dynamics are those associated with normal diffusion: spatial smoothing if primarily excitatory neurons are excited at long ranges, and Turing patterns if primarily inhibitory neurons are excited at long ranges. When the long-range connections primarily are associated with inhibitory neurons, the overall sign of the Laplacian is negative, leading to reverse diffusion. Reverse, or backward diffusion, has exactly the opposite effect of normal diffusion. Where normal diffusion leads to smoothing of excitations, reverse diffusion leads to clumping of excitations at ever shorter length scales (see Fig. 3). These dynamics rely on the unusual spatial distribution of \( I-I \) connections. When an inhibitory neuron fires, connected inhibitory neurons at distant lattice sites become less active, allowing excitatory neurons at that site to become more active. Meanwhile, the level of excitation at the original lattice site may be maintained by local feedbacks with excitatory neurons. Close to each of these sites, excitatory activity is suppressed via short range \( I-E \) or \( E-I \) connections, unless \( E-E \) connectivity is strong enough to counteract the overall effect. This local feedback leads to increasingly incoherent local excitations, with activity patterns nearby lattice sites less strongly correlated. If the long-range connections in the network are dominated by \( I-I \) connections, then increasingly localized clumping of excitations results.

Recent theoretical studies of Turing patterns in reaction-diffusion systems have shown that intrinsic and extrinsic noise enhances the stability of Turing patterns (24–26) through an extension of the quasi-cycle mechanism of McKane and Newman (27, 28). If such results were to hold for pattern formation in \( V1 \), achieving robust visual function would be very difficult. We have investigated the effects of noise on the dynamics of our model of \( V1 \) and found that, when lattice scale \( I-I \) connections are sparse, noise does not enlarge the set of parameters that support pat-

Fig. 2. (A) Turing pattern of neural excitation in visual cortex coordinates. (B) The same pattern represented in visual field coordinates (i.e., in the coordinates that a patient undergoing geometric visual hallucinations would see). The logarithmic map between visual field and \( V1 \) is responsible for the dramatic logarithmic spiral structure of the hallucination. Although the image shown here is from computation, its qualitative features are very similar to those reported by patients (7). Figure generated with the parameters indicated in the text plus \( g_{II} = 6.5, g_{EI} = 0.1 \) with all other \( g_{ij} = 1 \).

A B

forward diffusion

backward diffusion

Fig. 3. Schematic representation of how long-range inhibition leads to instability in strips of visual cortex. An excess of excitatory over inhibitory activity is indicated in red and the converse in blue. A illustrates dynamics with normal visual cortex architecture, where long-range inhibition is forbid-den. These dynamics are analogous to normal diffusion. The upper row shows an initial distribution of activity and the lower row shows the evolution of this distribution at a later time. The spatial distribution of activity is smoothed. B illustrates dynamics with added long-range inhibition, whose dynamics are analogous to backward diffusion. The upper row shows the same initial distribution of activity, and the lower row shows the evolution of this distribution at a later time. Under reverse diffusion, the spatial distribution of activity becomes less smooth, leading to short length scale spatial structures.
terns. However, when extensive lattice scale $I-I$ coupling is introduced, fluctuation-induced “quasi-patterns” incompatible with normal visual function are generated (see SI Text). A further functional role of forbidding such connections may then be avoidance of quasi-pattern generation. We also note that the Turing patterns that do occur in our model of the visual cortex with realistic connectivity are deterministic and highly regular. This behavior is in contrast to most Turing systems, where quasi-patterns dominate and substantial fluctuations in the patterning are expected (24–26). However, it was shown in ref. 29 that Turing patterns generated by noise can be pinned to an underlying lattice provided by the lattice scale patchy $E-E$ connections described earlier. The lack of extensive lattice scale inhibitory connections contributes to the stability of such a pinning, and helps to explain, for example, why subjects report seeing geometric visual hallucinations that are highly regular (7).

It should be noted that, although the model we present in Eqs. 1 and 2 is highly simplified, the results are based only on the elementary features of the bifurcation structure of the model. It is well known that the bifurcation structure of models in statistical mechanics and dynamical systems is sensitive to only primitive, detail-independent considerations such as symmetry, fluctuations, range of interaction, and spatial dimension (30, 31). Thus it can be expected that the results will be largely unchanged in more detailed models of V1, which include other standard formulations, such as integrodifferential equations (5) and versions in more detailed models of V1, which include other standard formulations, such as integrodifferential equations (5) and versions

Discussion

Results on the development of the orientational preference map (9, 21) can be combined with our work to constrain the evolution of several of the basic features of the network anatomy of V1. In refs. 9 and 21, it was shown that, for the orientational preference map of V1 to develop correctly, a lattice scale Mexican hat interaction is required—i.e., short-range amplification and long-range suppression. To achieve this in a two-population excitatory-inhibitory model, either long-range inhibition or long-range $E-I$ connections must dominate long-range activation of excitatory neurons. Because our work suggests that extensive lattice scale inhibitory connections are detrimental for normal vision, the only network structure that is consistent with both of these results must have only sparse lattice scale inhibitory connections, and the lattice scale $E-I$ connections must have greater effective range than the lattice scale $E-E$ connections. However, to avoid hallucinations, the lattice scale $E-I$ range must not greatly exceed the local $E-E$ range. These V1 circuit properties apply to both the avoidance of hallucinations in normal vision and to the development of orientation preference maps. In fact, both connectivity properties deal with exactly the same problem: breaking the symmetry of translation and orientation preference.

In summary, in V1 the lattice scale network’s most elementary features—patchy excitatory connections and sparse inhibitory connections—are completely constrained by two considerations: the need for the visual state to be robust and the developmental requirements of the orientational preference map.

ACKNOWLEDGMENTS. This work was partially supported by National Science Foundation Grant NSF-EF-0526747. Additional funding for T.C.B. was provided through a Dickramer fellowship from the University of Illinois Department of Physics.