

Canonical Cortical Circuits

Rodney J. Douglas and Kevan A. C. Martin

The observation that neural circuits of the neocortex are adapted to many different tasks raises deep questions of how they are organized and operate. Most theories of cortical computation propose that the cortex processes its information in a feedforward manner through a series of hierarchically organized stages and that each of these stages is dominated by the pattern of the input to the local cortical circuit. The most influential of these models of the local circuit is Hubel and Wiesel's (1962) proposal for the circuits that underlie simple and complex cells in the cat's primary visual cortex. Felleman and Van Essen (1991) extended the notion of a processing hierarchy in their comprehensive summary wiring diagram for the primate visual system. In these models of intra- and interareal cortical circuits, sensory information from the retina is passed through successive stages of cortical processing, each of which increases the feature selectivity of visual receptive fields. Thus, from the concentric center-surround receptive fields of the retina and dorsal lateral geniculate nucleus, simple cells are created, then complex cells from simple cells, and eventually the face cells, object-specific cells, and 3-D motion-specific cells of the high levels of the cortical processing hierarchy.

This serial processing schema is conceptually simple, which makes it very attractive for theorists (e.g., Riesenhuber and Poggio, 1999). More recent experimental and theoretical considerations of the cortical circuits, however, have suggested a rather different architecture: one in which local circuits of cortical neurons are connected in a series of nested positive and negative feedback loops, called "recurrent circuits" (Fig. 2.1; Douglas et al., 1989; Douglas and Martin, 2004, 2007). Excitatory neurons outnumber the inhibitory neurons by 5 to 1, so this ratio might be expected to create an unstable positive feedback. However, because the recurrent connections also exist

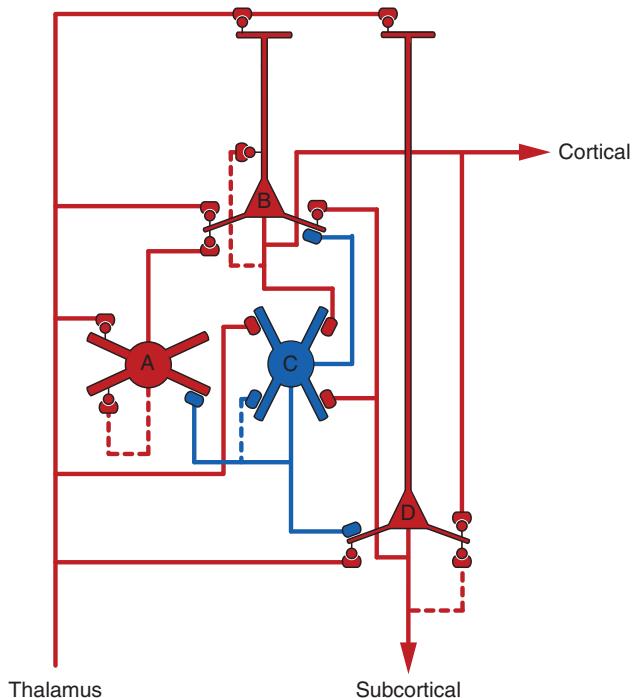


FIGURE 2–1. A canonical circuit for neocortex. Thalamic relay cells mainly form synapses in the middle layers of cortex, but they also form synapses with neurons in all six cortical layers, including the tufts of pyramidal cells in layer 1. In all layers the excitatory (red) and inhibitory (blue) neurons form recurrent connections with like cells within the same layer (dashed lines) and with other cell types (continuous lines). Layer 4 in some primary sensory cortical areas contain a specialist excitatory cell type, the spiny stellate cell (A), which projects to pyramidal cells and inhibitory cells in layer 4 and other layers. The superficial layer pyramidal cells (B) connect locally and project to other areas of cortex. Inhibitory neurons (C) are found in all layers (only one representative is shown here), and they constitute about 15% of the neurons in the neocortex. The deep layer pyramidal cells (D) also connect recurrently locally and project to subcortical nuclei in the thalamus, midbrain, and spinal cord.

between excitatory and inhibitory neurons, inhibition increases in proportion to excitation and the two opposing forces remain approximately in balance.

In the feedforward model, the thalamic input is strong, and it dominates the output of the neurons. In the recurrent model, however, input to the local circuits from the thalamus, or from other cortical areas, is thought to be relatively weak and the recurrent circuits either amplify or suppress this input (Douglas et al., 1989). The oldest and most notable example of “selective” amplification is the orientation preference of the neurons in the layer 4 of the cat’s primary visual cortex. Although these neurons receive monosynaptic input from thalamic neurons that have nonoriented receptive fields, they can amplify the excitation generated by optimally oriented stimuli and suppress

the thalamic excitation generated by nonoptimal stimuli. Thus, the goodness of “fit” of the input pattern to the “expectation” of the cortical circuits determines whether the input is amplified. These features of recurrent excitation and inhibition, amplification of weak inputs, and balanced excitation and inhibition, are fundamental attributes of the cortical circuits. To the extent they are features that appear in all cortical areas so far examined, they are defining characteristics of the proposed “canonical” circuit of neocortex (Fig. 2.1; Douglas et al., 1989, Douglas and Martin, 2004).

What is the experimental evidence that the thalamic input, which provides the cortex with its major input from the peripheral sense organs and from the basal ganglia, is relatively weak? The best evidence is from cat area 17, where anatomical and physiological studies indicate that the thalamus provides only a fraction (10%) of the total excitatory input to their main target neurons (Douglas et al., 1989; Binzegger et al., 2004; Da Costa and Martin 2009). The remaining excitatory synapses in layer 4 are contributed by other cortical neurons. Electrophysiological studies in slices of cat area 17 showed that while the amplitudes of the excitatory postsynaptic potentials (EPSPs) generated by putative thalamic axons were two-fold larger than those from local cortical neurons when stimulated at 1 Hz, they depressed with repeated stimulation (Stratford et al., 1996; Bannister et al., 2002). Thus, *in vivo*, where the spontaneous activity of thalamic afferents is relatively high, the amplitude of thalamocortical EPSPs may be considerably reduced by synaptic depression even before a stimulus arrives. In the rodent sensory cortices, the thalamic synapses are also outnumbered by the synapses arising from neighboring cortical neurons (White, 1989).

The evidence for recurrent connections between cortical neurons comes from a consideration of the distributions of cortical synapses. The most comprehensive analysis of the cortical circuit (Binzegger et al., 2004) indicates how much recurrent excitatory connections dominate within and between cortical layers. The intralaminar excitatory connections are most prominent in layers 2 and 3, where the pyramidal cells form most of their local excitatory synapses with each other, so much so that their recurrent connections involve one-fifth of all the excitatory synapses in area 17 (Binzegger et al., 2004). The consequence of this is that the recurrent connections between layer 2 and 3 pyramidal cells may predominate, whereas for other layers the interlaminar recurrent connections may have a greater role. For example, the spiny stellate neurons in layer 4 of cat visual cortex receive 40% of their excitatory synapses from pyramidal cells in layer 6 and only about 20% from their neighboring spiny stellate cells in layer 4.

The concept of serial processing within a cortical “column,” introduced by Hubel and Wiesel in 1962, brought to attention the importance of the interlaminar connections. However, neurons live in a 3-D space and they can have extensive projections not just within a column, but laterally (Fig. 2.2). One of the most impressive examples of this is that of the superficial layer

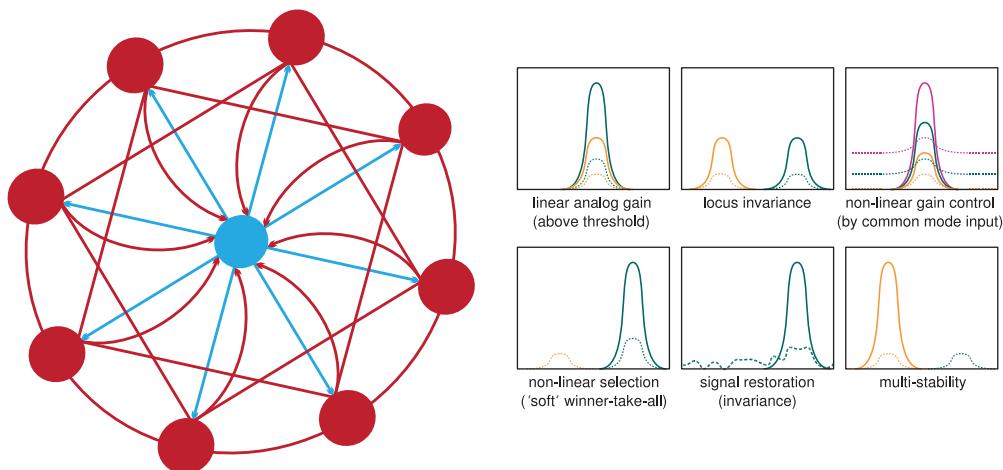


FIGURE 2–2. Recurrent circuit formed by lateral connections and some of the basic computations it could perform. (Left) Single or pools of excitatory neurons (red filled circles) are recurrently connected (red curved lines) with their neighbors and with a pool of inhibitory cells (blue filled circle). A parameter (e.g., orientation preference) is mapped around the circle of excitatory neurons, so that nearest neighbors lie closer together in the parameter space (have similar preferences) and are more strongly connected than more distant neurons in the map. Cortical Daisies are represented here by bidirectional excitatory connections that skip nearest neighbors (straight red lines) and may connect to neurons with dissimilar functional preferences. (Right) Illustration of various computations such as linear analog gain, where above threshold, the network amplifies its hill-shaped input (stippled lines) with constant gain (output, solid lines). Locus invariance occurs when the gain remains the same across the map (provided that the connection weights are homogenous across the network). In gain modulation, the network gain is modulated by an additional constant input applied to all the excitatory neurons and superimposed on the hill-shaped input. The gain is least when no constant input is applied (input, orange stippled line; output, orange solid line) and largest for a large constant input (mauve lines). When two inputs of different amplitude are applied to the network, it selects the stronger one by a nonlinear selection or “winner-take-all” operation. Signal restoration restores the hill-shaped input, even when that input is embedded in noise. When separate inputs have the same amplitudes, multistability is the operation that selects only one input: which input is selected depends upon the initial conditions of the network at the time the input is applied.

pyramidal cells, which distribute their synaptic boutons in patches or clusters. If a small cluster of neurons is viewed from the surface of the cortex, their axons form patches of terminals that have the appearance of the petals of a flower. This structure, which we refer to as the cortical “Daisy” (Douglas and Martin, 2004), is found in the cortical areas of all nonrodent species studied so far. Our view is that these horizontal axon clusters are the means by which pyramidal cells collectively participate in a selection network (Fig. 2.2). The selection mechanism is a soft winner-take-all or soft MAX mechanism,

which is an important element of many neuronal network models (Riesenhuber and Poggio, 1999; Maass, 2000; Yuille and Geiger, 2003). In this way, the superficial layer neurons would cooperate to explore all possible interpretations of input, and so select an interpretation consistent with their various subcortical inputs. However, these same pyramidal neurons not only participate in the local cortical circuit, but many of them also project outside their own cortical area to other cortical areas or subcortical structures and do this according to precise rules that govern the numbers and laminar origins of the pyramidal cells that form the interareal projections (Kennedy and Bullier, 1985; Barone et al., 2000). Thus, many of same neurons that form a Daisy in one area also provide input to Daisies in other cortical areas.

The inhibitory cells are recurrently connected with the spiny excitatory cells and with each other (Figs. 2.1 and 2.2). This arrangement is probably an early feature in the evolution of nervous systems, not just neocortex, and was originally revealed by Charles Sherrington in his studies of the spinal cord reflexes. For Sherrington, excitatory and inhibitory neurons were always in tandem and together they provided the algebra of the nervous system: "The net change which results there when the two areas are stimulated concurrently is an algebraic sum of the plus and minus effects producible separately by stimulating singly the two antagonistic nerves" (Sherrington, 1908). All cortical inhibitory neurons are GABAergic, and they conveniently divide according to which of three different calcium-binding proteins they contain (Douglas and Martin, 2004). The presence of these calcium-binding proteins correlates with the morphology of the different types of inhibitory cells and their specific connections with the spiny cells, which form 85% of their synaptic targets. The parvalbumin-containing cells, like chandelier and basket cells, which target the soma, proximal dendrites, and axon initial segment, seem well-positioned to control the output of the cell. The calbindin- or calretinin-containing cells, such as the double bouquet cells or Martinotti cells, form synapses with the more distal dendrites and thus are probably concerned with controlling the input to pyramidal cells, which are their major targets. The GABAergic cells may also colocalize polypeptides such as somatostatin, vasointestinal polypeptide, or cholecystokinin. Interestingly, although the basket cells were so named because they formed a pericellular nest of terminals around the cell body of pyramidal cells, they actually form most of their synapses with the dendrites of their target excitatory cells (Douglas and Martin, 2004). Some evidence for the effectiveness of distal inhibition has come from studies of the apical dendritic tuft of large layer 5 pyramidal neurons in the somatosensory cortex of the rat. This neuron, which projects to subcortical structures of the thalamus and midbrain, possesses the longest apical dendrite of any neuron in the cortex. Its apical tuft is the source of a calcium spike that can be gated by a distal inhibitory input, probably from Martinotti cells (Murayama et al., 2009).

Recurrent inhibition has found its most universal incarnation in the “normalization” model of visual cortex (Carandini et al., 1997). This model was developed to correct for the deficiencies of models of simple cells and complex cells, in which the inputs were summed linearly and their output passed through a spiking threshold. Standard linear models with rectification do not explain many experimental observations, such as why the responses of all cortical cells saturate, or adapt, or are suppressed by masking stimuli. The modification is to add a recurrent inhibitory pathway in which the inhibition is proportional to the pooled activity of a large number of cortical cells and acts to divide the firing rate of each cell in the pool. Quite how this might be implemented mechanistically is not at all clear, but such normalization models do at least offer one means of correcting the deficiencies of the linear models.

The normalization model requires a collective computation of all the neurons in the circuit. However, each neuron forms 5000 or more synapses and the low firing rates of cortical neurons indicate that only a restricted subset of these 5000 can be active at one time. Thus, while the combinatorial possibilities of the 5000 or more inputs onto a single cortical neuron provide numbers that are more than astronomical, cortical neurons provide outputs that are highly robust and reliable in space and time, with the result that most of the time our perceptions and actions are well-matched to the environment. Here inhibition can play a key role in determining which few hundreds of neurons constitute the effective circuit at any moment, because the effective circuit is created only by those neurons that are above threshold. Although it seems likely that there is a degree of redundancy in the inputs, even then only a very restricted subset of outputs typically should occur. The number of different parameters represented in the output of a cortical neuron is likely to be tens, not hundreds or thousands. We have referred to this constraint in numbers and patterns of active neurons in a recurrently connected population as the “permitted set.”. This set is the combination of active neurons whose effective weight matrix is stable and allows the network to converge to a steady state in a given context (Hahnloser et al., 2000; Douglas and Martin, 2007). This would require that perhaps only 10% of the synapse to be active. Thus, the nature of the recurrent activity and the size of the projective fields of clusters of neurons that share common inputs indicate that the computed output will be represented by the activity of less than 1000 neurons.

The organization of the neocortical circuits and the principles of their operation are still only very partially understood. However, each technical advance over the past century has reaffirmed that repeated patterns of structure and function are seen at every level, from molecule to cell to circuit, and that many of these patterns are common across cortical areas and species. In this context, the concept of a canonical circuit, like the concept of hierarchies of processing, offers a powerful unifying principle that links structural and functional levels of analysis across species and different areas of cortex.

ACKNOWLEDGMENTS

We thank Elizabetta Chicca and Nuno Da Costa for their creative graphics. The research was supported by EU Daisy grant FP6-2005-015803.

REFERENCES

Bannister NJ, Nelson JC, Jack JJ (2002) Excitatory inputs to spiny cells in layers 4 and 6 of cat striate cortex. *Philos Trans R Soc London B* 357:1793–1808.

Barone P, Batardiere A, Knoblauch K, Kennedy H (2000) Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J Neurosci* 20:3263–3281.

Binzegger T, Douglas RJ, Martin KAC (2004) A quantitative map of the circuit of cat primary visual cortex. *J Neurosci* 24:8441–8453.

Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17:8621–8644.

Da Costa N, Martin KAC (2009) The proportion of synapses formed by the axons of the lateral geniculate nucleus in layer 4 of area 17 of the cat. *J Comp Neurol* 143:101–108.

Douglas RJ, Martin KAC (2004) Neuronal circuits of the neocortex. *Ann Rev Neurosci* 27:419–451.

Douglas RJ, Martin KAC (2007) Mapping the matrix: the ways of neocortex. *Neuron* 56:226–238.

Douglas RJ, Martin KAC, Whitteridge D (1989) A canonical microcircuit for neocortex. *Neural Comput* 1:480–488.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.

Hahnloser R, Sarpeshkar R, Mahowald M, Douglas R, Seung S (2000) Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405:947–951.

Hubel D, Wiesel T (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.

Kennedy H, Bullier J (1985) A double-labeling investigation of the afferent connectivity to cortical areas V1 and V2 of the macaque monkey. *J Neurosci* 5:2815–2830.

Maass W (2000) On the computational power of winner-take-all. *Neural Comput* 12: 2519–2535.

Murayama M, Pérez-Garcí E, Nevian T, Bock T, Senn W, Larkum ME (2009) Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. *Nature* 457: 1137–1141.

Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019–1025.

Sherrington C (1908) On the reciprocal innervation of antagonistic muscles. Thirteenth note. On the antagonism between reflex inhibition and reflex excitation, *Proc R Soc London, B* 80b:565–578 (reprinted *Folia Neuro-Biol* 1908, 1:365).

Stratford KJ, Tarczy-Hornoch K, Martin KAC, Bannister NJ, Jack JJ (1996) Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature* 382:258–261.

White EL (1989) *Cortical Circuits: Synaptic Organization of the Cerebral Cortex. Structure, Function and Theory*. Boston: Birkhauser.

Yuille AL, Geiger D (2003) Winner-take-all networks. In: Arbib M, ed. *The Handbook of Brain Theory and Neural Networks*, pp. 1228–1231. Cambridge, MA: MIT Press.