Circuitry and the classification of simple and complex cells in V1

Jim Wielaard and Paul Sajda

Laboratory for Intelligent Imaging and Neural Computing
Department of Biomedical Engineering
Columbia University, New York, USA

June 18, 2006

Abstract

Based on a large-scale neural network model of striate cortex (V1), we present a simulation study of extra- and intra-cellular response modulations for drifting and contrast reversal grating stimuli. Specifically, we study the dependence of these modulations on the neural circuitry. We find that the frequently used ratio of the first harmonic to the mean response to classify simple and complex cells is highly insensitive to circuitry. Limited experimental sample size for the distribution of this measure makes it unsuitable for distinguishing whether or not the dichotomy of simple and complex cells origi- nate from distinct LGN axon connectivity and/or local circuitry in V1. We show that a possible useful measure in this respect is the ratio of the intracellular second to first harmonic response for contrast reversal gratings. This measure is highly sensitive to neural circuitry and its distribution can be sampled with sufficient accuracy from a limited amount of experimental data. Further, the distribution of this measure is qualitatively similar to that of the subfield correlation coefficient. It is however more robust and easier to obtain experimentally.

keywords: visual cortex, simple cells, complex cells, spatial summation, model, simulation

Hubel and Wiesel (Hubel and Wiesel, 1962; Hubel and Wiesel, 1968) postulated the existence of two distinct cell classes in the primary visual cortex (V1): cells that receive input from the lateral geniculate nucleus (LGN), designated as “simple” cells, and cells that do not receive input from the LGN, designated as “complex” cells. There seems at present little doubt that in terms of their extra- cellular responses (spikes), V1 cells can be sensibly divided into two categories. However, there is an ongoing debate as to whether the circuitry part of the Hubel & Wiesel hy- pothesis is true, i.e. whether in fact there exist two V1 cell classes in terms of LGN axon connectivity and/or distinct cortical circuitry within V1 (Gilbert, 1977; Gilbert, 1983; Ferster and Lindstrom, 1983; Spitzer and Hochstein, 1988; Hirsch et al., 1998; Alonso and Martinez, 1998; Chung and Ferster, 1998; Chance et al., 1999; Alonso et al., 2001; Martinez and Alonso, 2001; Hirsch et al., 2002; Meichler and Ringach, 2002; Abbott and Chance, 2002; Priebe et al., 2004; Martinez et al., 2005).

Hubel and Wiesel based their hypothesis on their obser- vation that apparently two distinct classes of ON/OFF sub- fields makeup the receptive fields of V1 cells. A different, but reasonably consistent (Mata and Ringach, 2005) measure for simple and complex cells can be given in terms of extra- cellular response modulations in response to drifting grating stimuli. The distribution of $S1/S0$ (the “modulation index”) for spike responses (where $S1$ is the first harmonic and $S0$ the mean of the spike train) resulting from a drifting grat- ing stimulus is profoundly bimodal for V1 cells (Movshon et al., 1978a; Movshon et al., 1978b; DeValois et al., 1982; Lennie et al., 1990; Skottun et al., 1991; Casanova et al., 1992; Chino et al., 1994; DeAngelis et al., 1994; Ohzawa et al., 1996; Smith et al., 1997; Ohzawa et al., 1997; Sceniak et al., 1999; Cumming et al., 1999). Simple cells are conventionally defined as having $S1/S0 > 1$, and complex cells $S1/S0 < 1$. There seems to be somewhat less consistency between the $S1/S0$ and subfield measures in alert monkeys (Kagan et al., 2003). Also, the bimodality of the $S1/S0$ distribution was observed to be much less profound in alert monkeys (Kagan et al., 2003), but is unclear whether this is due to the much larger number of complex cells than simple cells in the sample, which by itself would have a diminishing effect on a bimodal- ity. However, in terms of subfield organization of the receptive field for spike responses, also in alert monkeys two classes are clearly distinguishable (Kagan et al., 2003).

From a more general signal processing perspective, simple and complex cells in V1, defined either via subfield or the $S1/S0$ criterion, also tend to have distinct properties. Simple cell responses are approximately linear and show a linear dependence on spatial phase, whereas complex cells display a strong nonlinear behavior and are insensitive to spatial phase. Basic neural computations regarding visual perception seem to require the linearity provided by simple cells (Gra- ham, 1989; Wandell, 1995). Further, neurons in other sensory
cortices have linear signal processing properties that resemble those seen in V1. Examples are the primary auditory cortex (Kowalski et al., 1996) and the primary somatosensory cortex (DiCarlo and Johnson, 2000). Finally, simple and complex cells have been found in the primary visual cortex of many other species of mammals, such as owl monkeys (O’Keefe et al., 1998), baboons (Kennedy et al., 1985), tree shrews (Kaufman and Somjen, 1979), rats (Burne et al., 1984), mice (Drager, 1975), rabbits (Glanzman, 1983) and sheep (Kennedy and Martin, 1983). Clearly, understanding the neural mechanisms and circuitry that lead to the creation of simple and complex cells is of fundamental importance.

It has been recently suggested (Priebe et al., 2004) that evidence for the existence of two distinct circuitry classes related to the simple/complex dichotomy in V1 seems to be lacking in terms of intracellular recordings in cat. Distributions of intracellular response properties of V1 cells in response to drifting gratings, as well as of the subfield correlation coefficient based on intracellular responses, were found to be unimodal and were interpreted as not to provide any indication of two distinct classes of circuitry.

In contrast, recent intracellular measurements in cat (Martinez et al., 2005) from another group were interpreted as evidence in favor of two distinct classes of circuitry. In that study, distributions of different intracellular quantities such as the subfield correlation coefficient, subfield overlap and the intracellular push-pull index were found to be clearly bimodal as the subfield correlation coefficient. The simulations are based on a large-scale neural network model of macaque V1 (Wieland and Sajda, 2005), which allows us to investigate the dependency of the distributions of response modulations on circuitry. It enables us to clarify if and how such measures could provide insight into the question of whether there are two distinct classes of V1 cells in terms of LGN input (Hubel and Wiesel, 1962; Hubel and Wiesel, 1968; Martinez et al., 2005) in line with the original Hubel & Wiesel hypothesis, or if LGN input arrives in V1 in a more egalitarian manner (Chance et al., 1999; Mechler and Ringach, 2002; Abbott and Chance, 2002; Tao et al., 2004; Priebe et al., 2004).

**Methods**

**Model summary.** We provide here only a very brief description of the model used. For a complete description and further details of the model we refer to Wieland and Sajda, 2005. Some additional background information can also be found in previous work (McLaughlin et al., 2000; Wieland et al., 2001) by one of the authors (JW).

The model consists of 8 ocular dominance columns and 64 orientation hypercolumns (i.e. pinwheels), representing a 16 mm² area of a macaque V1 input layer 4Coα or 4Cβ. The model contains approximately 65,000 cortical cells and the corresponding appropriate number of LGN cells. Our cortical cells are modeled as conductance-based integrate-and-fire point neurons, 75% are excitatory cells and 25% are inhibitory cells. Based on circuitry, the model contains two classes of cortical cells: cells that receive LGN input and cells that do not. Our LGN cells are rectified spatiotemporal linear filters. The model is constructed with isotropic short-range cortical connections (< 500 μm), realistic LGN receptive field sizes and densities, and realistic sizes of LGN axons in V1.

Dynamic variables of a cortical model-cell $i$ are its membrane potential $v_i(t)$ and its spike train $S_i(t) = \sum_k \delta(t - t_{i,k})$, where $t$ is time and $t_{i,k}$ is its $k$th spike time. Membrane potential and spike train of each cell obey a set of $N$ equations of the form

$$\frac{dv_i}{dt} = -g_{L,i}(v_i - v_L) - g_{E,i}(t, [S]_E, \eta_E)(v_i - v_E)$$

$$-g_{I,i}(t, [S]_I, \eta_I)(v_i - v_I), \quad i = 1, \ldots, N .$$

These equations are integrated numerically using a second order Runge-Kutta method with time step $0.1$ ms. Whenever the membrane potential reaches a fixed threshold level $v_T$ it is reset to a fixed reset level $v_R$ and a spike is registered. The equation can be rescaled so that $v_i(t)$ is dimensionless and $C_i = 1, v_L = 0, v_E = 14/3, v_I = -2/3, v_T = 1, v_R = 0$, and conductances and currents have dimension of inverse time.

The quantities $g_{E,i}(t, [S]_E, \eta_E)$ and $g_{I,i}(t, [S]_I, \eta_I)$ are the excitatory and inhibitory conductances of neuron $i$. They are defined by interactions with the other cells in the network, external noise $\eta_E(I)$, and, in the case of $g_{E,i}$ possibly by LGN input. The notation $[S]_{E(I)}$ stands for the spike trains of all excitatory (inhibitory) cells connected to cell $i$. Both, the excitatory and inhibitory populations consist of two subpopulations $P_k(E)$ and $P_k(I), \; k = 0, 1$. The $k = 1$ populations receive LGN input and the $k = 0$ populations do not. In the model presented here, 30% of both the excitatory and inhibitory cell populations receive LGN input. We assume noise, cortical interactions and LGN input act additively in contributing to the total conductance of a cell,

$$g_{E,i}(t, [S]_E, \eta_E) = \eta_E(t) + g_{E,i}^{cor}(t, [S]_E) + \delta_i^{LGN}(t)$$

$$g_{I,i}(t, [S]_I, \eta_I) = \eta_I(t) + g_{I,i}^{cor}(t, [S]_I) ,$$

where $\delta_i = \ell$ for $i \in \{P_k(E), P_k(I)\}, \quad \ell = 0, 1$. The terms $g_{E,i}^{cor}(t, [S]_E)$ and $g_{I,i}^{cor}(t, [S]_I)$ are the contributions from the cortical excitatory ($\mu = E$) and inhibitory ($\mu = I$) neurons and include only isotropic connections,

$$g_{\mu,i}^{cor}(t, [S]_\mu) = \frac{v_i - v_{\mu}}{\eta_{\mu}(t)}, \quad \mu = E, I.$$
where $i \in P_{\mu'}(\mu)$ Here $\vec{x}_i$ is the spatial position (in cortex) of neuron $i$, the functions $G_{\mu,j}(\tau)$ describe the synaptic dynamics of cortical synapses and the functions $C_{\mu',k}(r)$ describe the cortical spatial couplings (cortical connections). The length scale or excitatory and inhibitory connections is about 200µm and 100µm respectively.

In agreement with experimental findings, the LGN neurons are modeled as rectified center-surround linear spatiotemporal filters. A cortical cell, $j \in P_{\ell}(\mu)$ is connected to a set $N_{Q,j}^{\mu}$ of left eye LGN cells, or to a set $N_{R,j}^{\mu}$ of right eye LGN cells, $g_j^{\mu}(t) = \sum_{\ell \in N_{Q,j}^\mu} [g^0_\ell + g^V_\ell]$

\[ \int_{-\infty}^{+\infty} ds \int_{-\infty}^{+\infty} dy \ G_{\ell}^{\mu}(t-s) \ L_{\ell} \left( ||\vec{y} - \vec{y}_i|| \right) I(\vec{y}, s)_{+}, \]

where $Q = L$ or $R$. Here $[x]_+ = x$ if $x \geq 0$ and $[x]_+ = 0$ if $x \leq 0$. $L_{\ell}(r)$ and $G_{\ell}^{\mu}(r)$ are the spatial and temporal LGN kernels respectively. $\vec{y}_i$ is the receptive field center of the $i$th left or right eye LGN cell, which is connected to the $j$th cortical cell, $I(\vec{y}, s)$ is the visual stimulus. The parameters $g^0_\ell$ represent the maintained activity of LGN cells and the parameters $g^V_\ell$ measure their responsiveness to visual stimuli. The LGN kernels are of the form

\[ G_{\ell}^{\mu}(\tau) = \begin{cases} \ 0 \ & \tau \leq \tau^0_{\ell} \\ \ k \ \tau^5 \ \left( e^{-\tau/\tau_1} - e^{-\tau/\tau_2} \right) \ & \tau > \tau^0_{\ell} \end{cases} \]

and

\[ L_{\ell}(r) = \pm (1 - K_{\ell})^{-1} \left\{ \frac{1}{\pi \sigma c_{\ell} r} \ e^{- (r / \sigma_{c,\ell})^2} - \frac{K_{\ell}}{\pi \sigma s_{\ell} r} \ e^{- (r / \sigma_{s,\ell})^2} \right\}, \]

where $K$ is a normalization constant, $\sigma_{c,\ell}$ and $\sigma_{s,\ell}$ are the center and surround sizes respectively, and $K_{\ell}$ is the integrated surround-center sensitivity.

The connection structure between LGN cells and cortical cells, given by the sets $N_{Q,j}^{\mu}$, is made so as to establish ocular dominance bands and a slight orientation preference which is organized in pinwheels (Blasdel, 1992). It is further constructed under the constraint that the LGN axonal arbors sizes in V1 do not exceed the anatomically established values.

A sketch of the model’s geometry and connection structure is given in Figure 1. Note that, contrary to the Hubel & Wiesel picture, our model can only be classified as hierarchical in terms of LGN inputs and not in terms of connections between the 4 cell classes. For example, we see that excitatory cells with LGN input (orange) receive their excitation in about equal amounts from LGN, excitatory cells without LGN input and excitatory cell with LGN input. Similarly, excitatory cells without LGN input (red) receive significant excitation from other cortical cells without LGN input. A strict Hubel & Wiesel connection scheme would consist of only the LGN input and the connection (orange) from $P_{\ell}(E)$ to $P_{0}(E)$ (assuming inhibitory cells are not included in a strict Hubel & Wiesel scheme).

**Stimuli.** The luminance of the drifting grating stimulus used is given by $I(\vec{y}, t) = I_0(1 + \epsilon \cos(\omega t - \vec{k} \cdot \vec{y}))$, with average luminance $I_0$, contrast $\epsilon$, temporal frequency $\omega$, spatial wave vector $\vec{k}$. Contrast-reversal stimuli are given by $I(\vec{y}, t) = I_0(1 + \epsilon \cos(\omega t) \cos(\vec{k} \cdot \vec{y} + \phi))$. All parameters are set to preferred values, response modulations for contrast reversal stimuli are averaged over the spatial phase $\phi$. Receptive fields are mapped out using small spots (diameter $< 1/5$th receptive field) of oscillating luminance, $I(\vec{y}, t) = I_0(1 + \epsilon \cos(\omega t) \cos(\vec{k} \cdot \vec{y} + \phi))$. On and OFF responses are collected during the bright $I(\vec{y}, t) > I_0$ and dark $I(\vec{y}, t) < I_0$ parts of the cycles. We used the spatial correlation coefficient of the ON and OFF subfields,

\[ r = \frac{\langle (r_{ON} - \langle r_{ON} \rangle)_s \rangle \langle (r_{OFF} - \langle r_{OFF} \rangle)_s \rangle_s}{\langle (r_{ON} - \langle r_{ON} \rangle)_s \rangle_s^2 \langle (r_{OFF} - \langle r_{OFF} \rangle)_s \rangle_s^2}^{1/2}, \]

where $\langle \cdot \rangle_s$ denotes the spatial average and $r_{ON}$ ($r_{OFF}$) represents the amplitude of the ON (OFF) response at a particular spatial location. All stimuli are presented monocularly and at high contrast ($\epsilon = 1$).
Results

Drifting gratings

Examples of our model’s extracellular and intracellular response modulations for a drifting grating stimulus are provided in Figure 2. These are responses of both a simple and a complex cell for several orientations of the grating, at the cells’ preferred spatial and temporal frequencies. Notice the dominance of the first harmonic in the response of the simple cell.

The distribution of the spike train modulation index $S_1/S_0$ over our cell population is shown in Figure 3A. Strength parameters (see Methods) have been set so that the distribution of these modulations in the spike responses is in agreement with experimental data for macaque (Ringach et al., 2002). Note that the distribution is profoundly bimodal and that our model cortex (as does macaque V1) contains about an equal number of simple and complex cells (as defined by the $S_1/S_0$ criterion).

It is easy to understand how the diversity in response modulations occurs in our model. The modulations enter our model cortex via the LGN input received by 30% of the cortical cells. The phases of these LGN inputs into the different cortical cells vary randomly on $[0,2\pi]$. This is so because of the receptive field off-sets of the clusters of LGN cells connected to different cortical cells, the difference in spatial organization (e.g. symmetry) of the clusters themselves, and the diversity in temporal delays in the LGN kernels (Wiesel and Sajda, 2005). A cell receives input from many other cells, thus a cell’s excitatory and inhibitory inputs will show stronger or weaker modulations depending on its specific environment in the network and whether or not it receives LGN input. Interplay between the strengths and phases of the modulations in these inputs and cell specific parameters ultimately determine the modulation in the cell’s spike and membrane potential response.

Our model, by construction, has two distinct classes of cells: cells that do and cells that do not receive LGN input (Methods). We see that, despite this fact, the classification into simple and complex cells is not “sharp”. That is, the distribution in Figure 3A is far from a binary distribution and is in fact a smooth distribution with a continuous support (distribution in Figure 3A is far from a binary distribution and is in fact a smooth distribution with a continuous support). As illustrated in Figure 3A, classification into simple and complex cells in terms of $S_1/S_0$ (i.e. $S_1/S_0 = 1$ is the class boundary) is a poor measure for the difference in underlying circuitry (LGN input vs. no LGN input). Unless specified, in what follows when referring to simple/complex cells we will mean the $S_1/S_0 = 1$ classification scheme. From Figure 3A, lower panel, we find that 24% of the cells that do not receive LGN input are simple cells ($S_1/S_0 > 1$), while 10% of the cells that do receive LGN input are complex cells ($S_1/S_0 < 1$).

The minimal classification error for recognition of the underlying circuitry is obtained with a class boundary $S_1/S_0 = 1.4$, i.e. "no LGN input" $S_1/S_0 < 1.4$ vs. "LGN input" $S_1/S_0 > 1.4$. In that case we make classification errors of 7% and 21% respectively. The minimal average classification error possible based on extracellular responses is thus 11%.

In reality, intracellular responses are much more difficult to obtain than spike responses. One of the advantages of having a model as ours is that they are just as easy to extract as are spike responses (Methods). Our model’s distribution of modulations in the membrane potential with respect to a blank stimulus ($V_1/V_0$) is shown in Figure 3B. We see that the profound bimodality present for spike train modulations is not present in the $|V_1/V_0|$ distribution for the membrane potential. However, there is little mixing of the extracellularly defined simple and complex cell classes. In fact (not shown), the complex cells (as defined by $S_1/S_0 < 1$) reside mostly (84%) in the narrow core of the distribution ($|V_1/V_0| < 2$), while the simple cells reside mostly (89%) in its long tails, which extend both for positive and negative values of the modulation.
index. About 8% of the cells have negative $V_1/V_0$. Further, as shown in Figure 3B, our model predicts a discontinuous support for the distribution with a “gap” at small negative values, i.e. $V_1/V_0 < -\delta$ or $V_1/V_0 > 0$ with $\delta \approx 2$.

Because of the absence of the spike threshold, one may generally expect intracellular (or sub-threshold) responses to be a better reflection of the synaptic inputs of a cell than extracellular responses. Hence one may therefore expect its intracellular responses to be a better reflection of the circuitry relating to a cell than its extracellular responses. This is of course not necessarily true for all intracellular responses and all measures derived from them. Interestingly, our simulations show that this is in fact not true for the $V_1/V_0$ measure. As shown in Figure 3B, the minimal classification error is obtained for the class boundary $|V_1/V_0| = 2$, i.e. “no LGN input” $|V_1/V_0| < 2$ vs. “LGN input” $|V_1/V_0| > 2$. In that case we make classification errors of 14% and 5% respectively. The minimal average classification error possible based on intracellular responses is thus again 11%, as it was for extracellular responses. Contrary to general expectations, identification of the circuitry cannot be made more accurately from the intracellular modulations $V_1/V_0$ than from the extracellular modulations $S_1/S_0$.

The distribution of modulations in the membrane potential has not yet been observed experimentally in macaque. However, as mentioned in the Introduction, intracellular data for cat has recently been published (Priebe et al., 2004). Our model’s results show qualitative agreement with these data. In particular, as is shown in Figure 4, the distributions of $V_0$, $V_1$, and $V_t$ are also unimodal, while separation of simple and complex cell classes only occurs in the $V_1$ distribution (and not for $V_0$ and $V_t$). All potentials are measured with respect to the blank response $v_B$. Note that, although the threshold potential $v_T$ of our integrate & fire neurons is fixed and identical for each cell, it becomes cell dependent when measured with respect to the blank response, i.e. $V_t = v_T - v_B$.

Qualitatively, these findings can be explained from a simple rectification model (Mechler and Ringach, 2002; Priebe et al., 2004) where all intracellular modulation occur with the same (temporal) frequency as the stimulus. If one includes the “blank” response (zero contrast) in such a model, it is straightforward to show that generally a bimodal extracellular distribution corresponds to a unimodal intracellular distribution (like the one obtained for our model) including a gap in the support at small negative values. This gap occurs when $V_0$ becomes negative (with respect to the blank), the first harmonic $V_1$ must then exceed $V_0$ in order for the cell to fire. This requires the gap width $\delta > 1$. Here we assume that for a blank stimulus any action potentials in the network occur as a result of the fluctuations in the membrane potential (“fluctuation driven dynamics”), rather than as a result of its mean value, which is true for practically all cells as can be seen from Figure 4C. As mentioned earlier, in our model we observe a gap width of about 2 (Fig. 3B).

Experimental data for intracellular distributions are an important piece of the rather sparse body of information currently available on V1 circuitry. The absence of a bimodality in the intracellular distributions observed experimentally in cat V1, has been interpreted (Priebe et al., 2004) as lack of evidence for the existence of two distinct cell classes, and instead as evidence for the existence of a continuum of circuitry. However, this conclusion is strictly speaking not correct: two distinct classes of circuitry can of course just as well result in a unimodal as in a bimodal distribution. Our model results are merely a concrete example of this fact. Thus whether or not distinct circuitry is responsible for simple and complex cells remains very much an open question, which we will further address in the following sections.

**Contrast reversal gratings**

Examples of extracellular and intracellular response modulations in our model in response to a contrast reversal grating are shown in Figure 5. These are averaged response waveforms of spike train and membrane potential, with the grating...
at the preferred orientation and preferred spatial and temporal frequencies. Shown are the responses of a simple and a complex cell in the model for several spatial phases $\phi$ of the grating. Simple cells perform an approximately linear spatial summation, that is, their responses contain a dominant $S_1$ component and the spatial phase dependence of their response waveform is similar to the spatial phase dependence of the stimulus. Complex cells respond nonlinear, their response waveform is relatively insensitive to spatial phase and contain a dominant $S_2$ component (frequency doubling).

The distribution of the phase-averaged $S_2/S_1$ for a sample consisting of an equal number of cells with and without LGN input is shown in Figure 6A. The distribution displays a weak bimodality, and this behavior agrees with experimental data (Hawken and Parker, 1987). Interestingly, this property of our model follows naturally, without any parameter adjustments, after the strength parameters have been set to achieve essentially only orientation tuning and a proper distribution of response modulations in response to a drifting grating (Fig. 3A), see (Wielaard and Sajda, 2005).

The minimal classification error for extracellular classification of cells with and without LGN input occurs for a class boundary at $S_2/S_1 = 1$, as can be seen in Figure 6A. In that case we make classification errors of 20% and 5% for cells with and without LGN input respectively, so that the minimal average classification error possible based on this extracellular measure is 10%.

The distribution of the phase-averaged $S_2/S_1$ for a different sample, in this case consisting of an equal number of simple and complex cells (as defined by $S_1/S_0$) is shown in Figure 6C. The distribution looks very similar as that in Figure 6A. However, identification of simple and complex cells is notably less precise than identification of cells with and without LGN input: using the class boundary $S_2/S_1 = 1$ we make an average classification error of about 16% in classifying simple and complex cells.

It is easy to understand how the diversity in $S_2/S_1$ (and $V_2/V_1$) occurs. As explained in (Wielaard et al., 2001), for a contrast reversal grating stimulus the total LGN input into a cortical cell has, in general, a dominant $S_1$ component with a phase close to either 0 or $\pi$, determined by the positions of the ON and OFF subfields relative to the grating. The cortical excitatory and inhibitory inputs in a cell will thus have a strong $S_2$ component since they arise from many other cells. The actual strengths of $S_1$ and $S_2$ components in a cell’s excitatory and inhibitory inputs thus depends on the cell’s specific environment in the network and on whether it
receives LGN input or not. Interplay of these inputs and cell specific parameters determine the $S2/S1$ ratio in the cell’s spike response. Clearly, most cells that receive LGN input (simple) will have $S2/S1 < 1$ and most cells that do not receive LGN input (complex) will have $S2/S1 > 1$.

![Diagram of cell types](image)

Figure 7: Examples of the organization of the extracellular and intracellular ON and OFF subfields for a simple and a complex cell in the model. (A) Extracellular, complex cell. (B) Intracellular, complex cell. (C) Extracellular, simple cell. (D) Intracellular, simple cell.

No experimental data is available for the distribution of $V2/V1$ of the membrane potential waveforms. Our model’s distribution of the phase-averaged $V2/V1$ is shown in Figure 6B & D. Our model predicts that, quite contrary to the situation for the modulation index $S1/S0$ and $V1/V0$, the (weak) bimodality of the distribution of $S2/S1$ for spike waveforms is not eliminated in the $V2/V1$ distribution for membrane potential waveforms. Rather, it becomes substantially more pronounced in the $V2/V1$ distribution. Finally, we see that the intracellular measure $V2/V1$ is indeed a better reflection of the circuitry than the extracellular measure $S2/S1$. The minimal classification error for intracellular classification of cells with and without LGN input occurs for a class boundary at $V2/V1 = 1$, as can be seen in Figure 6B. We make minimal classification errors of 13% and 2% for cells with and without LGN input respectively, intracellularly the minimal average classification error possible is thus 5%, as shift by a factor of 2 with respect to the extracellular measure $S2/S1$.

The fact that the weak bimodality present in the $S2/S1$ becomes much more pronounced in the $V2/V1$ distribution, can again be intuitively understood from the simple rectification model mentioned earlier. In this model the membrane potential waveforms are subjected to a threshold to give the spike waveforms. For complex cells, both the membrane potential and spike responses will contain a strong second harmonic. Hence in this case practically all of the membrane potential waveform will be above threshold, so that evaluation of $V2/V1$ will yield about the same result as for $S2/S1$. This is apparent in Figure 6: the $S2/S1 > 1$ and $V2/V1 > 1$ sections of the distributions are very similar. For simple cells, the membrane potential and spike responses will contain a dominant first harmonic, and both will display about an equally small second harmonic component. Because of the rectification, the first harmonic present in the membrane potential waveform is substantially reduced in the spike waveform. Hence, $V2/V1$ will turn out substantially smaller than $S2/S1$ in this case. This is again apparent in Figure 6: the $V2/V1 < 1$ (simple cells) section of the distributions is shifted to the left with respect to the $S2/S1 < 1$ sections.

**Subfield correlation**

An example of the ON/OFF subfield organization for a simple and a complex cell in our model is shown in Figure 7. The subfields are obtained with a small spot of oscillating (8Hz) luminance (Methods), responses are collected during the bright (ON) and dark (OFF) parts of the cycle. In agreement with experimental observations, the simple cell (Fig. 7C,D) shows ON and OFF subfields which are separated, while for the complex cell (Fig. 7A,B) they largely overlap.

The distributions of the subfield correlation coefficients (Methods) based on spike and membrane potential responses are qualitatively similar to the $S2/S1$ and $V2/V1$ distributions respectively. This is illustrated in Figure 8. That this is so can also be understood intuitively. Cells with strongly overlapping ON and OFF subfields ($r \approx 1$) will generate a dominant second harmonic and very little first harmonic response for a contrast reversal grating, i.e. will have large $S2/S1$ and $V2/V1$ ratios. Cells with strongly non-overlapping subfields ($r \approx -1$) will generate a dominant first harmonic and little second harmonic response for a contrast reversal grating, i.e. will have small $S2/S1$ and $V2/V1$ ratios. Hence, one may naturally expect some similarity between the distributions of the subfield correlation coefficient and of the ratio of second and first harmonic in response to contrast reversal gratings.

Similarly as for the intracellular $V2/V1$ measure, it can be seen from Figure 8 A&B that also the intracellular subfield correlation coefficient (Fig. 7B) better reflects the circuitry than does the extracellular subfield correlation coefficient (Fig. 7A). Extracellularly we obtain a minimal average classification error (for classifying cells with and without LGN input) of 10% while this is reduced five fold intracellularly to a minimal classification error of 2%. The classification of simple and complex cells is notably worse, see Figure 8 C&D. Using $r = 0$ as the class boundary, we make an average extracellular classification error of 22% (Fig. 8C). Intracellularly we make an average classification error of 16% (Fig. 8D). A summary...
of our results for the classification utility (LGN vs no LGN) of the different extracellular and intracellular metrics discussed are shown in Table 1.

As already mentioned in the Introduction, different results for the distribution of the intracellular subfield correlation coefficient were obtained by different experimental laboratories. A unimodal (or at best very weakly bimodal) distribution was observed by Priebe et al., 2004. A strongly bimodal distribution was observed by Martinez et al., 2005. It must be noted that the latter distribution was constructed from significantly fewer cells however. On the other hand, the conclusion made by Priebe et al., 2004, i.e. that their unimodal distribution would suggest a continuum of circuitry, is strictly speaking incorrect. What was noted earlier for the intracellular V1/V0 distribution is also true here: two distinct classes of circuitry can of course, in principle, just as well result in a unimodal as in a bimodal distribution (of the intracellular subfield correlation coefficient). In this case, however, as we will see below, our model results indeed confirm to some extent the naive intuition that a unimodal distribution of the intracellular subfield correlation coefficient suggests a continuum of circuitry.

More specifically, we have yet to address the following questions. To what extent can a strong bimodality in the intracellular V2/V1 (or subfield correlation coefficient) distribution, like the one we observe in our model, actually be interpreted as a signature of V1 circuitry? Is it correct to conclude that if intracellular experimental data would show a profoundly bimodal V2/V1 (or subfield correlation coefficient) distribution, that this is evidence for existence of two distinct classes of circuitry in V1? Could presence of a less profound bimodality, or unimodality of the observed V2/V1 (or subfield correlation coefficient) distribution be evidence for a continuum of circuitry in V1? We address these questions in the next section.

**Continuous vs. discrete classes of V1 circuitry**

Conductance and membrane potential waveforms (cycle average, see Methods) for a typical model cell are shown in Figure 9. Results shown are for a drifting grating stimulus (left column) and for a contrast reversal grating stimulus (right column). For the drifting grating (Fig. 9A) we see that the LGN input (black) is dominated by the first harmonic, while the cortical conductances have a large mean value with substantially smaller modulations. The situation is distinctly different for the contrast reversal stimulus (Fig. 9B). At the preferred phase, the LGN conductance still contains a dominant first harmonic, but it is phase sensitive and the first harmonic practically vanishes at the orthogonal phase. The cortical conductances are insensitive to the grating phase and clearly show a dominant second harmonic component.

In our model (as well as in real V1 cortex, see e.g. (Destexhe et al., 2003)) conductances are large. As a result, there is to a good approximation a simple Ohmic relation between the cycle averaged conductances and membrane potential (Wieland et al., 2001),

\[ \langle v_k(t) \rangle \approx V_k(t) = \frac{I_{D,k}(t)}{g_{T,k}(t)}, \]

(8)

where

\[ I_{D,k}(t) = g_{E,k}(t) V_E - g_{I,k}(t) |V_I|, \]

(9)

\[ g_{T,k}(t) = g_L + g_{E,k}(t) + g_{I,k}(t), \]

(10)

and \( g_{E,k}, g_{I,k} \) are the total excitatory and inhibitory conductances and \( V_E, V_I \) are the excitatory and inhibitory reversal potentials respectively. That the Ohmic approximation (8) is indeed highly accurate is shown in Figure 9E,F.

Manipulation of the relative strengths of the cortical and LGN components of the excitatory conductances changes the relative strengths of the modulations in the membrane potential (and spike train) and hence the distributions of these modulations. Together with the Ohmic approximation, this provides us with a way to predict, given results of our main simulation, the behavior of these distributions as function of
the distributions look like for an arbitrary continuum of cell (with and without LGN input) it allows us to predict what \( \zeta \) and after the transformation 

\[
(E,F) \quad \text{Membrane potential (black), Ohmic approximation (dashed)}
\]

\( g \) denotes convolution. We leave the inhibitory and leakage conductances unchanged for each cell in the population and compute the membrane potential modulation distributions from (8). Further, we take

\[
\hat{\alpha}_k(n) = \begin{cases} 
-\zeta \beta_k C_k & \text{for } n = \pm 2 \\
-\beta_k C_k & \text{else}
\end{cases},
\]

where \( C_k = g_{E,k}^\text{LGN}(0)/g_{E,k}^\text{COR}(0) \) and \( \sim \) denotes the Fourier components, e.g. \( \hat{\alpha}(n) = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} \alpha_k(t) e^{-i n \omega t} dt \). Constructed in this way, the parameter \( \alpha_k(t) \) provides a constant scaling factor with additional amplification of the second harmonic set by the parameter \( \zeta \). The construction assures cells maintain reasonable firing rates, since the mean of the total excitatory conductance is unaltered, i.e. \( \hat{G}_{E,k}(0) = \hat{g}_{E,k}(0) \) for arbitrary \( \beta_k \) and \( \zeta \). We eliminate all LGN input, and hence create a cell from the other class (without LGN input) by taking \( \beta_k = 1 \). Results of this manipulation, and \( \zeta = 2 \), for a model cell are shown in Figure 9C-F. Notice the drastic change in the membrane potential (green); the transformation in this case obviously changes the cell from a simple cell to a complex cell.

We now return to the distributions of the membrane potential modulations \( V/2V1 \). We apply the transformation to a sample of \( N_0 = 600 \) model cells with LGN input, and calibrate the transformation by fixing \( \zeta \) so that for \( \beta_k = 1, k = 1, \ldots, N_0 \) we obtain a \( V/2V1 \) distribution that approximately matches that for cells without LGN input in our model. This is shown in Figure 10A&B. The distribution of \( V/2V1 \) for the sample with LGN input is shown in Figure 10A. First we note that making the transformation (12) with \( \beta_k = 0, k = 1, \ldots, N_0 \) (arbitrary \( \zeta \)), changes little because of the high accuracy of the Ohmic approximation (compare solid and dashed curves in Fig. 10A).

From Figure 10B we see that making the transformation (12) with \( \zeta = 2, \beta_k = 1, k = 1, \ldots, N_0 \), results in a distribution of \( V/2V1 \) that approximately matches that for cells without LGN input in the model (compare solid, dotted and dashed curves).

Keeping \( \zeta = 2 \) fixed, we can now create a sample consisting of two distinct classes of cells (with and without LGN input) from the sample consisting of only cells with LGN input, and recover the results of our full simulation. We do this by making the transformation (12) with \( \beta_k = 0 \) for half of the cells and \( \beta_k = 1 \) for the other half in the sample. This is shown in Figure 10C, we see that the result (red curve) closely resembles the distribution of Figure 6B, i.e. our full simulation result for the \( V/2V1 \) distribution for a sample consisting of an equal number of cells with and without LGN input.

We can introduce a continuum of circuitry (with respect to LGN input) by making the transformation (12) with \( \beta_k, \ k = 1, \ldots, N_0 \) drawn randomly (and independently) from a uniform distribution on \([0,1]\) for each cell. The result is shown in Figure 10C (blue curve). We see that our method predicts that the profound bimodality present for two distinct classes of circuitry entirely disappears for a continuum of circuitry.

Our transformation method is not limited to a contrast reversal grating stimulus. We can apply the same transformations to the responses the stimuli that were used to obtain
the ON and OFF subfields for our model cells (Methods). The results are shown in Figure 11. In this case the transformation is somewhat less precise for cells with a higher subfield correlation, compare red and black curves. The reason is that higher subfield correlation implies higher overlap of the subfields and thus weaker responses to the spot stimuli used (Methods). Figure 11 (red and black curves) thus shows that the signal-to-noise-ratio of our data is somewhat too low for the ON and OFF subfields for our model cells (Methods). We believe our work clears up some apparent misconceptions (Priebe et al., 2004) for what concerns the interpretation of the unimodal nature of intracellular distributions measured also reported for a detailed neural network model similar in spirit to ours, but with a more continuous (egalitarian) distribution of LGN inputs (Tao et al., 2004). Its insensitivity thus makes the V1/V0 distribution a poor indicator for V1 circuitry. This is particularly true given the limited quantity of experimental data attainable even in principle with current techniques. Note for instance that to truly experimentally confirm the distribution predicted by our model, i.e. including the negative branch, the gap etc., the sample needs to be much larger than the 100 cells for which data is currently available (from Fig. 12 we may deduce that a sample size of 1200 is in fact still quite insufficient for this purpose).

Finally we apply the same transformations to the drifting grating stimulus, i.e. to the V1/V0 distribution. The results are shown in Figure 12. As reference samples we use again the samples consisting of 600 cells with and 600 cells without LGN input of Figure 10 A&B. As was the case for contrast reversal, the result for binary circuitry practically coincides with the full simulation result (black and red curves). However, unlike the V2/V1 distribution for contrast reversal, the V1/V0 distribution for drifting gratings changes little for continuous circuitry. Indeed, confirmation of the experimental data was Fig 11: Distributions of the intracellular subfield correlation coefficient for binary and continuous circuitry. Original distribution for the sample consisting of the cells with and without LGN input of Figure 10 A&B (black). Distribution for the sample of Figure 10A, after the transformation $\beta_k = 0$ for half of the cells and $\beta_k = 1$ for the other half (red). Distribution for the sample of Figure 10A, after the transformation with $\beta_k$ drawn randomly (independently) from a uniform distribution on $[0,1]$, for each cell in this sample (blue).

Fig 12: Distributions of the intracellular V1/V0 ratio for binary and continuous circuitry. (A) Original distribution of V1/V0 (drifting grating) for the cells with and without LGN input of Figure 10 A&B (black). Distribution of V1/V0 for the sample of Figure 10A, after the transformation $\beta_k = 0$ for half of the cells and $\beta_k = 1$ for the other half (red). Distribution of V1/V0 for the sample of Figure 10A, after the transformation with $\beta_k$ drawn randomly (independently) from a uniform distribution on $[0,1]$ for each cell (blue).(B) As A, but on a linear scale.

Discussion

We believe our work clears up some apparent misconceptions (Priebe et al., 2004) for what concerns the interpretation of the unimodal nature of intracellular distributions measured...
recently in cat.

In general unimodal distributions can result from one, two, a continuum or any other arrangement of classes. It is thus in principle not correct to conclude (Priebe et al., 2004) that generally unimodal distributions suggest continuous classes (circuitry). Our simulations provide illustrations of this simple fact. Our model cortex with two distinct classes (of circuitry) yields unimodal intracellular distributions for V0, V1, V1/V0 and V4 as observed experimentally. This is thus an illustration of the fact that it is incorrect to interpret the unimodal feature of these distributions as evidence for a continuum of circuitry. Our combined results for two distinct classes and for a continuum of classes, show that in the case of the intracellular subfield correlation coefficient distribution, a unimodal distribution may indeed be taken as evidence for a continuum of circuitry, in agreement with naive expectations.

Moreover, classification based on a unimodal distribution is not necessarily inferior to classification based on a bimodal distribution. Our simulations provide an example of this by means of the intracellular V1/V0 and corresponding extracellular S1/S0 distributions. As we have shown, identification of the two cell classes in the model (cells with and without LGN input) can be made just as well in terms of the core and tails of the V1/V0 distribution as in terms of the two modes of the S1/S0 distribution. Incidentally, the same seems to hold true with regard to classification of simple and complex cells on the (bimodal) S1/S0 and (unimodal) V1/V0 distributions for experimental data (Priebe et al., 2004, their Fig. 3B), as well as for a model with a continuum of LGN circuitry (Tao et al., 2004, their Fig. 6A), and for our model (not shown).

Finally, we have shown that contrary to general expectations, the intracellular measure V1/V0 is not a better indicator of circuitry than the extracellular measure S1/S0 (Table 1). We have shown that the intracellular V1/V0 distribution is in fact highly insensitive to circuitry and not likely to be a very useful measure to distinguish between circuitry of one or another kind based on experimental data.

Our simulations show that the intracellular V2/V1 distribution for contrast reversal stimuli is highly sensitive to whether cells receive their LGN input in a continuous (egalitarian) or in a distinctly binary fashion. Our results suggest this measure can be utilized as an indicator of V1 circuitry. The sensitivity of this measure is caused by the fact that for contrast reversal stimuli, there is a clear qualitative difference between the phase dependence of the LGN and the collective cortical inputs into a cell. The LGN input shows a distinct phase dependence whereas the collective cortical input is practically phase insensitive and frequency doubled. If the LGN inputs arrives in V1 in a binary fashion, our simulations predict the distribution of V2/V1 for cells in the input layers is profoundly bimodal. If the LGN input is egalitarian, our simulations predict that this distribution is profoundly unimodal.

We have shown that the behavior of the intracellular subfield correlation coefficient as function of circuitry is largely similar to what we see for V2/V1. This also holds true (not shown) for the subfield overlap index, a measure which is strongly correlated with the subfield correlation coefficient (Mata and Ringach, 2005). Our results suggest that also the distribution of the intracellular subfield correlation coefficient can serve as an indicator of V1 circuitry. But the removal of the bimodal nature for continuous circuitry is somewhat less clear for this distribution than for the V2/V1 distribution. As mentioned in the Introduction, the intracellular subfield correlation coefficient distribution has, unlike the V2/V1 distribution, been observed experimentally. Naively interpreted, the experimental data so far however suggest conflicting pictures for the circuitry of simple and complex cells. On the one hand the data of Priebe et al., 2004 (i.e. their data for the intracellular subfield correlation coefficient, Fig. 7 of their paper) would suggest continuum of circuitry, while on the other hand the data of Martinez et al., 2005 (i.e. Figs. 3 of their paper and their Fig. 4 for the push-pull index distribution) would suggest two distinct classes of circuitry. Our findings presented here provide evidence that this naive interpretation is in fact correct: using numerical simulations we showed that the unimodal intracellular subfield correlation coefficient distribution of Priebe et al., 2004 is indeed likely an indication of a continuum of circuitry, while the bimodal intracellular subfield correlation coefficient distribution of Martinez et al., 2005 is likely an indication of two distinct classes of circuitry.

As for the reasons for the apparent discrepancy between the Priebe et al., 2004 and Martinez et al., 2005 data, there are several possibilities. The most obvious is the difference in sample size. The substantially smaller sized data set of Martinez et al., 2005 could simply lead to a “biased estimate” of the same phenomenon. Other reasons for the discrepancy could be related to the intracellular measurements of inhibition, as pointed out in Hirsch and Martinez, 2006. While in Martinez et al., 2005, responses to light/dark spots are frequently inverse images of each other (their Fig. 1), in Priebe et al., 2004 this is quite rare (their Fig. 6; note that pronounced negative deflections are very rare). This could be due to technical differences or also to differences in layer sampling (Priebe et al., 2004 did not identify the layers of the cells recorded; perhaps only a small proportion of their cells were recorded within layer 4). Finally there are perhaps other technical reasons, such as the fact that some of the cells from Martinez et al., 2005 were studied with QX-314.

It is important to realize that the capacity of the intracellular V2/V1 distribution, as a measure for simple/complex cell circuitry, is by no means limited to the context in which it is presented in this paper. That is, it is not limited to a situation in which the distinction between circuitry is "LGN input" and "no LGN input". Rather, it equally well applies to a situation where all simple and complex cells would be created entirely from cortical interactions (e.g. in V1 layers that do not receive LGN input). Also it applies equally well to a (purely hypothetical) situation where all simple and complex cells would be created entirely from LGN inputs. The reason
is that the behavior of V2/V1 for contrast reversal gratings is closely related to the subfield composition of a cell’s receptive field (and hence to the simple/complex cell dichotomy), and not, in principle, to whether the cell receives LGN input or not. Therefore, for what concerns the behavior of V2/V1, it is irrelevant if the subfield organization of a simple and complex cells results entirely from LGN input, entirely from cortical interactions, or from a combination of both. Thus, the phase-averaged V2/V1 distribution for contrast reversal gratings is quite in general a good indicator of whether the simple/complex dichotomy is a reflection of distinct circuitry and it is not limited to input layers. This is useful, as it means that as far as questions about "distinct" vs. "continuum" of circuitry go, experimental data from different layers could be pooled together to construct an overall V2/V1 distribution. This is not to say that layer-specific experimental observation of the V2/V1 distribution remains not of course of great interest. The distribution could very well be different for different layers, and in fact (as mentioned earlier), when so, this may be a reason for the apparent discrepancy between the Priebe et al., 2004 and Martinez et al., 2005 data.

We have shown that, unlike for the V1/V0 distribution (drifting grating), an experimental estimate of the V2/V1 distribution could be obtained for relatively small cell samples (order of hundreds of cells). The suggested V2/V1 distribution is also a robust measure: it concerns a ratio of low order Fourier components, of a generally strong response (to contrast reversal gratings), which is furthermore averaged over phases. It is for instance much less sensitive to noise and experimental errors than the subfield correlation coefficient or overlap index. As far as we know, experimental data is not yet available, or at least not in sufficient amount, to utilize this distribution as an indicator of V1 circuitry.

Acknowledgments

This work was supported by grants from ONR (MURI program, N00014-01-1-0625) and NGA (HM1582-05-C-0008).

References


