

Annotated Bibliography

Lane McIntosh

February 12, 2014

Contents

| | | |
|----------|---|----------|
| 1 | Theoretical Neuroscience | 2 |
| 1.1 | Efficient and Predictive Coding | 2 |
| 1.1.1 | Srinivasan 1982, Predictive coding: a fresh view of inhibition in the retina | 2 |
| 1.1.2 | Atick 1992, Could Information Theory be an Ecological Theory of Sensory Processing? | 2 |
| 1.1.3 | Fairhall 2001, Efficiency and ambiguity in an adaptive neural code | 3 |
| 1.1.4 | Baccus 2002, Fast and slow contrast adaptation in retinal circuitry | 3 |
| 1.1.5 | Hosoya 2005, Dynamic predictive coding by the retina | 3 |
| 1.1.6 | Sharpee 2006, Adaptive filtering enhances information transmission in visual cortex | 4 |
| 1.1.7 | Nagel 2006, Temporal Processing and Adaptation in the Songbird Auditory Forebrain | 4 |
| 1.1.8 | Wark 2007, Sensory adaptation | 4 |
| 1.1.9 | Meister 2008, Rapid Neural Coding in the Retina with Relative Spike Latencies | 5 |
| 1.1.10 | Pitkow 2012, Decorrelation and Efficient Coding in Retinal Ganglion Cells | 6 |
| 1.2 | Energy Cost and Efficiency of the Brain | 7 |
| 1.2.1 | Laughlin 1998, The metabolic cost of neural information | 7 |
| 1.3 | Methodologies | 7 |
| 1.3.1 | Strong 1998, Entropy and Information in Neural Spike Trains | 7 |
| 2 | Information Theory and Statistical Mechanics | 7 |
| 2.1 | Statistical Physics as Inference | 7 |
| 2.1.1 | Jaynes 1957, Information Theory and Statistical Mechanics I | 7 |
| 2.2 | Physics of Information | 8 |
| 2.2.1 | Still 2012, Thermodynamics of Prediction | 8 |
| 3 | Observations from Experimental Neuroscience | 8 |
| 3.1 | Observations about Early Visual Systems | 8 |
| 3.1.1 | Hubel 1988, Segregation of Form, Color, Movement, and Depth | 8 |
| 3.2 | Noise in Neural Systems | 8 |
| 3.2.1 | Stevens 1968, Synaptic noise and other sources of randomness in motoneuron interspike intervals | 8 |
| 3.3 | Activity-dependent plasticity | 9 |
| 3.3.1 | Lanahan 1998, Immediate-early genes and synaptic function | 9 |
| 3.4 | Feedback Connections | 9 |
| 3.4.1 | Sillito 2006, Always returning: feedback and sensory processing in visual cortex and thalamus | 9 |

1 Theoretical Neuroscience

1.1 Efficient and Predictive Coding

1.1.1 Srinivasan 1982, Predictive coding: a fresh view of inhibition in the retina

This paper by Srinivasan, Laughlin, and Dubs constitutes the first mention (to my knowledge) of predictive coding as applied to neural systems, although it follows in the grand tradition of efficient coding as pioneered by Barlow back in the 1950's.

The authors conceive of inhibitory interneurons as subtracting a weighted sum of global signals that constitute a prediction of the stimulus. For center-surround receptive fields for example, this amounts to inhibitory interneurons predicting what the center value might be. This prediction is then subtracted to yield predictive encoding, reducing spatial (and temporal) redundancy.

They derive the spatial and temporal filters with the minimum mean-squared error between the predicted value and the true center (space) or present (time) value. They find that for high SNR the spatial or temporal filter should be more biphasic, and at low SNR the filter should be more monophasic, which they verify in the second-order interneurons of the fly eye. The authors also criticize the redundancy reduction hypothesis as being qualitative, and put forward predictive coding as a quantitative instantiation of the efficient coding hypothesis.

“To further our understanding of retinal inhibitions repertoire, and to help define the function of inhibition in known networks, we have developed a new analytical approach, predictive coding. This simple procedure, initially developed for transmission of video data, reduces the signal amplitude by removing predictable, and hence redundant, components.”

1.1.2 Atick 1992, Could Information Theory be an Ecological Theory of Sensory Processing?

J.J. Atick identifies two sources of inefficiencies in information processing, 1) the unequal use of symbol values, and 2) statistical correlation between different symbols in a given message. Consider a message $w = (m_1, \dots, m_l) \in M$ where each m_i takes a value from a set of cardinality N . Then

$$\underbrace{H(M) \leq \sum_{i=1}^l \sum_{m_i=1}^N p(m_i) \log_2 p(m_i)}_{\text{statistical correlations between symbols}} = \underbrace{\sum_{i=1}^l H(i) \leq \max_{\{p(m_i)\}} \sum_{i=1}^l H(i)}_{\text{unequal use of } N \text{ symbol values}} = C, \quad (1)$$

where Atick defines C to be the source capacity (note that this is quite different from channel capacity). Each inequality represents the possible introduction of inefficiency into the source coding; the first inequality becomes an equality when symbol positions m_1, \dots, m_l are statistically independent, and the second inequality becomes an equality when the N values are used with equal frequency for each symbol.

Defining redundancy (or total inefficiency) as

$$R = 1 - \frac{H(M)}{C} \quad (2)$$

$$= \frac{1}{C} (C - H(M)) \quad (3)$$

$$= \underbrace{\frac{1}{C} \left(C - \sum_{i=1}^l H(i) \right)}_{\text{unequal use of } N \text{ symbol values}} + \underbrace{\frac{1}{C} \left(\sum_{i=1}^l H(i) - H(M) \right)}_{\text{statistical correlations between the } l \text{ symbols}} \quad (4)$$

we can see how the 2 types of inefficiencies contribute to redundancy in the source code.

Atick then references Laughlin, 1981 as an experimental verification that the nervous systems reduce the 1st kind of inefficiency (unequal use of N symbol values).

Introducing a simplified model of the retina where photoreceptor activity levels are mapped linearly and one-to-one onto retinal ganglion cells, Atick then derives the optimal filter. Looking at this derived filter in Fourier space, we see that at low frequencies where noise is low the filter flattens the frequencies of the source (which for naturalistic stimuli is $1/|f|^2$) - i.e., it's a whitening filter. Since whitening essentially removing the correlations in the stimulus, Atick interprets this optimal linear filter as the retina's method of removing the 2nd type of inefficiency (statistical correlations between symbols) from the visual world.

Atick does however admit that, "it is not clear in the retina what principle dictates the choice of the low-pass filter," since, although an optimal filter can be derived, a linear filter is not necessarily obtained when simply starting at the first principle that statistical correlations between symbols must be reduced. Putting this in a modern context, this uncertainty is well justified; Xaq Pitkow and Markus Meister's 2012 paper for instance established that less than a quarter of the correlations present in naturalistic visual stimuli are reduced by the retina's linear filtering. Instead, over half of the reduction in stimulus correlations is due to the nonlinear thresholding at the bipolar-retinal ganglion cell synapse. As Xaq has mentioned, "If the brain's computations could be carried out by linear filtering, and since the sum of many steps of linear filters is just another linear filter, we might as well not have a brain."¹

1.1.3 Fairhall 2001, Efficiency and ambiguity in an adaptive neural code

Fairhall et al. examine how the fly's retina adapts to changes in the visual stimulus statistics. Although this adaptation allows for a more efficient encoding of the visual environment in the face of a neuron's limited dynamic range, adaptation inherently makes the meaning of a pattern of spikes ambiguous. This paper investigates how these ambiguities are resolved.

1.1.4 Baccus 2002, Fast and slow contrast adaptation in retinal circuitry

Similar to the findings of Fairhall 2001 in motion sensitive H1 blowfly cells, Baccus and Meister observe fast and slow components of contrast adaptation in all retinal ganglion cells and certain bipolar and amacrine cells.

Fast effects of a contrast increase included accelerated kinetics, decreased sensitivity, and a depolarization of the baseline membrane potential. Slow adaptation did not affect kinetics, but produced a gradual hyperpolarization. This hyperpolarization can account for slow adaptation in the spiking output of ganglion cells.

¹Pitkow, Xaq. "Efficient coding and beyond in the visual system." The Stanford Center for Mind, Brain, and Computation Theoretical Neuroscience Colloquium Talk on Thursday, March 7.

1.1.5 Hosoya 2005, Dynamic predictive coding by the retina

This paper demonstrated *pattern adaptation* of retinal ganglion cell receptive fields. The complex patterns they tested included

- full-field flicker vs. checkerboard of anti-correlated white noise
- horizontal gratings vs. vertical gratings
- time series where luminance values 60ms apart were anti correlated vs. same value.

They show that these adaptive changes are in the direction of rendering retinal ganglion cells more sensitive to novel stimuli, and argue that this adaptation of the RGC linear filters is due to activity-dependent synaptic plasticity between amacrine cells and RGCs.

Barry Wark et al., 2007 summarizes this work nicely as well: “This study explores adaptation to complex spatial or temporal correlations in salamander and rabbit retina, showing that the receptive fields of many RGCs change following stimulation with these correlated stimuli such that the response to novel stimuli is enhanced compared with the response to the adapting stimulus. The authors reject the hypothesis of fatigue of pattern matching cells in favor of a model of Hebbian plasticity at inhibitory amacrine-to-RGC synapses. This conclusion is intriguingly at odds with other studies that find adaptation in the retina to simpler correlations (mean, variance) in the absence of inhibitory transmission.”

Note that the paper does not deal with changes in contrast. Kastner and Baccus 2011 demonstrated that under low contrast (i.e., high noise) regimes, RGC linear filters lengthen, leading to sensitization in many cells.

1.1.6 Sharpee 2006, Adaptive filtering enhances information transmission in visual cortex

From Barry Wark et al., 2007 review: “This paper applies a novel information theoretic reverse correlation method [31] to obtain cat V1 receptive fields during viewing of a white noise stimulus and of natural movies. The receptive fields differed in the two cases in their low-pass properties, in such a way as to transmit similar low frequency power in the two cases. The information transmitted about the stimulus was determined to change on the order of ~ 100 s.

1.1.7 Nagel 2006, Temporal Processing and Adaptation in the Songbird Auditory Forebrain

This paper demonstrates that adaptation in the avian auditory forebrain also have fast and slow components similar to those seen in the H1 motion sensitive cell in the blowfly (Fairhall 2001) and the salamander retina (Baccus 2002).

1.1.8 Wark 2007, Sensory adaptation

This is a review paper that discusses the presence of adaptation in early sensory systems at short and long timescales, and the evidence for whether the observed adaptation supports the efficient coding hypothesis. One interesting point the paper makes is that, “the change in coding strategy cannot be shorter than the time required for the system to ‘learn’ the new distribution.”

Here are some of the observations in sensory adaptation that Wark et al. mentions:

- Smirnakis et al., 1997, showed that retinal ganglion cells exhibit an adaptive change in firing rate when the variance of a flickering gaussian light stimulus changes

- Brenner et al., 2000, demonstrated that the nonlinear gain function (in an LN model fitted to the fly motion sensitive neuron H1) adapted such that the scaling of the stimulus axis was normalized by the stimulus standard deviation and that this serves to maximize information transmission about the stimulus. In essence, this is a whitening filter.
- Fairhall et al., 2001, further showed that this maximization of information happens rapidly (in $\sim 100\text{ms}$) during continuous changes of the stimulus variance.

In addition to this fast contrast gain control adaptation, there is also a slower adaptive process that modulates the overall firing rate of retinal ganglion cells. The take home message here is the distinction between contrast adaptation (slow, seconds), which essentially modulates the firing rate given the mean contrast to reduce Atick's 1st type of inefficiency, and contrast gain control (fast, $\sim 100\text{ms}$), which essentially changes the transfer function for a particular retinal location and spatial frequency to whiten the signal.

- Maravall et al. 2007 and Nagel & Doupe 2006 have also shown evidence of filter and gain curve adaptation in rat barrel cortex (white noise stimulation of whiskers; covariance analysis used to estimate gain curves) and avian auditory cortex (broadband noise input; filters from a fitted LN model), respectively.
- Dean et al. 2005 used Fisher information to demonstrate that the population of cells in inferior colliculus (responding to sound amplitudes) as a whole shifted responses to best encode the high probability sounds (although individual neurons had a variety of response changes during adaptation)

In a sense, classic studies on the remapping of sensory cortex in the absence of input from a particular region is an extreme example of how populations of neurons shift to encode high probability signals. If suddenly there aren't any nonzero signals, it makes sense trivially from an information maximization sense to reallocate that neuronal population to a different region on the somatosensory map.

- Mante et al. 2005 found that adaptation of RGC filters to changes to the mean and variance (mean luminance and contrast) of drifting gratings occur independently. They also examined natural scene statistics and found independence between luminance and contrast.
- Bonin et al. 2005 examined how neurons in LGN respond to changes in the higher order statistics of random checkerboards and found no adaptation
- Hosoya et al. 2005 generalized a previous finding of rate adaptation to the spatial scale of a flickering checkerboard to demonstrate adaptation to a variety of arbitrary spatiotemporal correlations in visual stimuli in RGCs and showed that the new filters that evolve after exposure to these correlations act to remove the correlations and so perform predictive coding
- Sharpee et al. 2006 developed an information theoretic reverse correlation method that finds the stimulus dimensions that maximize mutual information between spiking responses and the stimulus - maximally informative dimensions - to estimate the receptive field under non-gaussian stimuli - i.e., for use with natural image stimuli.

1.1.9 Meister 2008, Rapid Neural Coding in the Retina with Relative Spike Latencies

This is a nice paper that

- hypothesizes a rapid neural code (latency of the first spike) for conveying information about an image's structure,
- provides evidence that this code contains more information about the stimuli than a neural code that uses spike count
- demonstrates that the minimal model to generate a relative latency code requires RGCs to have input from ON and OFF bipolar cells, and

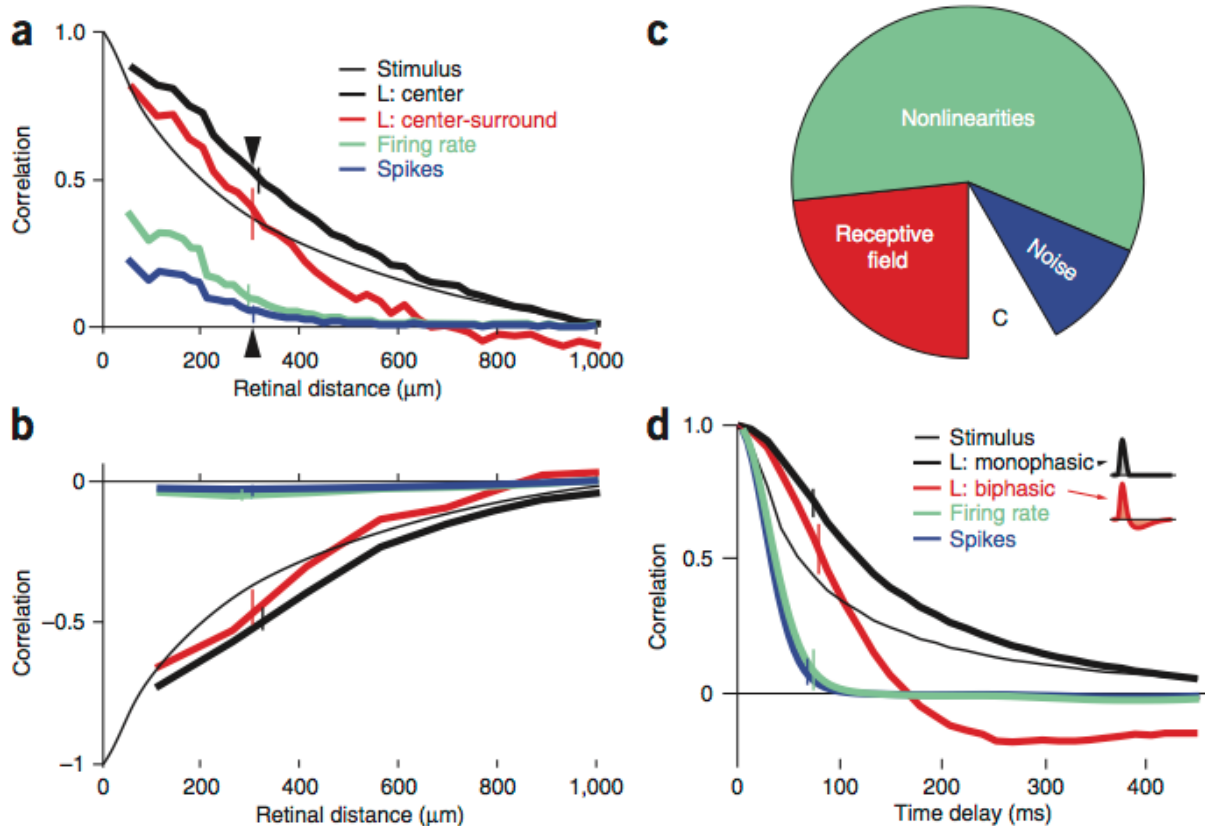


Figure 2 of Pitkow & Mesiter, 2012

- shows that by silencing ON bipolar cells, individual RGCs stop coding about half of the stimuli as predicted.

So the main results of the paper are that

1. the latency of a retinal ganglion cell's first spike relative to other RGCs first spike conveys information about an image's structure
2. and that input from both ON and OFF bipolar cells explains this

1.1.10 Pitkow 2012, Decorrelation and Efficient Coding in Retinal Ganglion Cells

Historically the field of sensory coding has pointed to center-surround, biphasic spatiotemporal filtering in the retina as the primary mechanism of reducing the redundancy in natural scenes and decorrelating neural responses. Pitkow and Meister investigate this claim by quantifying how each step of an LN model fit to salamander and macaque retina ganglion cells decorrelated natural images. Surprisingly, they find that the nonlinearity contributes the most to decorrelating natural images.

They then compute the optimal gain and threshold of a sigmoidal nonlinearity by assuming spike count distributions were Gaussian given the mean spike count (which is determined by the sigmoid nonlinearity), and then numerically optimizing bits/spike over the gain and threshold parameters. Interestingly, if gain and threshold are estimated from the firing rates of salamander retinal ganglion cells, they fall within the nearly optimally efficient regime. It would be interesting to do a more fine grained analysis between

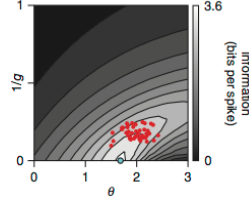


Figure 5 of Pitkow & Mesiter, 2012. Red dots are salamander retinal ganglion cells.

and within different retinal ganglion cell types. Presumably, the nonlinearity would strongly contribute to decorrelation between different cell types since direction-selective cells and object motion sensitive cells have different nonlinearities, for instance, while perhaps linear filtering is the best way of decorrelating cells of the same type. Another thing to note is that the retinal ganglion cell receptive fields were measured using spatiotemporal white noise and then tested using natural images. Since we know that these filters adapt dramatically, we may be dramatically underestimating the ability of linear mechanisms to decorrelate natural images. A better approach would have been to use Maximally Informative Dimensions to measure the receptive fields under natural conditions.

1.2 Energy Cost and Efficiency of the Brain

1.2.1 Laughlin 1998, The metabolic cost of neural information

This paper measures the amount of ATP required in the blowfly retina to transmit a bit of information and compares it to Landauer’s limit. Laughlin et al. conclude that it costs 10^4 ATP molecules to transmit a bit at a chemical synapse, and 10^6 - 10^7 ATP for graded signals in an interneuron or a photoreceptor, or for spike coding. Since the hydrolysis of 1 ATP to ADP releases about $25 kT$ (Landauer’s limit is kT/bit), information transmission in neural networks is orders of magnitude (2.5×10^6) more costly than the theoretical minimum.

This paper also observes that in noise-limited systems, low capacity channels transmit information more efficiently. This conclusion derived from their observation that the capacity of a neural signaling pathway is increased by adding more synapses. This results in an exponential increase in the cost per bit of information transmitted as the bit rate increases.

1.3 Methodologies

1.3.1 Strong 1998, Entropy and Information in Neural Spike Trains

This publication from Bialek’s group advocated the estimation of spike train entropies from the “binary word” approach.

First, divide up spiketrain into non-overlapping segments of T duration. Then choose some bin width Δt such that a maximum of one spike appears in each bin. Typically $\Delta t \approx 2$ ms, $T \rightarrow \infty$ approximated by taking $30 \text{ ms} < T < 100 \text{ ms}$ and extrapolating. For approximately $T < 20$ ms the information rate (bits/s) increases linearly as T becomes smaller. Estimating the entropy for various $T < 20$ ms allows for linearly extrapolating the entropy in the $T \rightarrow \infty$ limit.

2 Information Theory and Statistical Mechanics

2.1 Statistical Physics as Inference

2.1.1 Jaynes 1957, Information Theory and Statistical Mechanics I

Rather than viewing statistical mechanics as its own separate physical theory, we should view statistical mechanics as a form of statistical inference. Central to this shift in perspective is the idea that maximum entropy models are the least biased models that take into account our experimental constraints.

2.2 Physics of Information

2.2.1 Still 2012, Thermodynamics of Prediction

“All systems perform computations by means of responding to their environment.” In particular, how well systems respond to their environment - quantified by how much a system’s dynamics predict future environmental driving signals - determines the physical dissipation of the system as it’s driven by the environment. In a sentence, information processing inefficiency is equivalent to thermodynamic inefficiency.

$$\beta \langle W_{\text{diss}}[x_t \rightarrow x_{t+1}] \rangle = I[s_t; x_t] - I[s_t; x_{t+1}], \quad (5)$$

where s_t is the state of the system (for instance, the velocities and positions of gas molecules) at time t , x_t is the environmental driving signal (for instance, the position of a piston) at time t , and $\beta = 1/k_B T$.

This equality holds when the dynamics of the system are modeled by discrete time Markovian conditional state-to-state transition probabilities, $p(s_t | s_{t-1}, x_t)$. The paper also assumes that the system begins in thermodynamical equilibrium at time $t = 0$, and that a change in the external driving signal $x_0 \rightarrow x_1$ forces the system out of equilibrium.

3 Observations from Experimental Neuroscience

3.1 Observations about Early Visual Systems

3.1.1 Hubel 1988, Segregation of Form, Color, Movement, and Depth

This paper discusses the parallel pathways of visual information that are functionally distinct. In particular, the magnocellular pathway (originating in parasol retinal ganglion cells) is sensitive to depth and movement, while the parvocellular pathway (originating in midget RGCs) is sensitive to form and color.

The midget and parasol cells that comprise the parvo and magno cellular pathways in the retina, respectively, differ in color selectivity, contrast sensitivity, spatial resolution, and temporal properties. Specifically,

| | color selectivity | contrast sensitivity | spatial resolution | temporal properties |
|-------|--|--|--------------------|----------------------------|
| parvo | typically color-opponent (e.g. center R, surround G) | higher contrast | sharper resolution | slower and more persistent |
| magno | color-blind | more sensitive to low contrast (saturated at 10-15%) | poorer resolution | faster and more transient |

This raises an important point: when I'm using black and white stimuli, I may only be getting responses from parasol RGCs.

3.2 Noise in Neural Systems

3.2.1 Stevens 1968, Synaptic noise and other sources of randomness in motoneuron interspike intervals

Poisson noise in spike trains assuming a rate code is on the order of 10% the mean interval length. This noise is due to synaptic noise originating from the temporal and spatial summation of PSPs.

Calvin and Stevens ruled out noise in the mechanism converting synaptic currents into spike trains as a primary source of neural noise.

Synaptic noise (fluctuations in membrane potential on the order of 2 mV peak to peak that vary quite a bit from cell to cell) could include 1) asynchronous arrival of PSPs, 2) spontaneous mini EPSPs, 3) thermal noise across the membrane impedance ($k_B T/C$), and 4) shot noise from ions moving through the membrane. But the primary source is the temporal and spatial summation of (1). Fluctuations in the membrane potential due to thermal noise is $< 5\mu V$ and fluctuations due to shot noise of ion movements is $< 1\mu V$.

Of course, if we don't believe the rate model, then reason (1) is not noise at all.

3.3 Activity-dependent plasticity

3.3.1 Lanahan 1998, Immediate-early genes and synaptic function

This is a brief overview of how immediate-early genes were first discovered, the scope of their influence on neurons, and four possible mechanisms that could describe why immediate-early genes tend to be localized at the synapse.

Following work that suggested long-term memory required new protein synthesis, it was discovered that rapid mRNA synthesis (in addition to new protein synthesis) was also required for long-term plasticity.

Immediate early genes like *c-fos* were originally found in assays looking at gene over-expression in neurons following seizures.

One particularly interesting immediate early gene, *zif268*, was induced in hippocampal neurons by the same synaptic stimuli that induce LTP. The induction of both LTP and *zif268* mRNA was dependent on activation of NMDA receptors and had the same threshold for stimulus intensity. It seems that most neuronal IEG mRNAs are induced in response to NMDA receptor activation. Does this suggest that neurons can change synaptic weights and otherwise alter information transmission by extracting information from sub-action-potential-threshold membrane potential voltages (i.e. even subthreshold NMDA current)?

Scope of IEGs: Neuronal immediate early genes encode growth factors and proteins involved in signal transduction, transcription, cytoskeletal rearrangement, and metabolism (metabolic enzymes). It is thought that IEGs may directly modify synaptic function as well, and IEGs tend to be localized at the synapse. The authors expect that IEGs will contribute directly to synaptic plasticity by regulating the structure, signal transduction capability, or spatial localization of critical receptors.

3.4 Feedback Connections

3.4.1 Sillito 2006, Always returning: feedback and sensory processing in visual cortex and thalamus

Feedback connections are ubiquitous throughout the brain; for instance, only 10% of the excitatory connections to the LGN are feedforward inputs from the retina. The lion's share of inputs to LGN come from V1 and other higher areas of cortex, including MT.

This review article explores feedback connections from MT to LGN and V1 that are thought to be involved in visual perception.

Interaction between V1 and LGN:

- feedforward from LGN to layers 4 and 6 of V1
- 30% of LGN inputs are feedback from layer 6 of V1
- feedback from V1 excites distal dendrites of LGN relay cells and inhibitory interneurons

MT has feedback connections to the V1 layer 6 cells that feedback to LGN. This, together with the observation that layer 6 cells seem to modulate the cells in layer 4, leads the authors to speculate that layer 6 cells serve as a “gatekeeper” of retinal input to the visual system.

Affects of feedback on LGN cells:

- modulates strength of center-surround interaction for moving stimuli, in particular making surround antagonism stronger (lowering average firing rates of LGN cells and making them relatively more sensitive to shorter bars)
- increases sensitivity to differences in grating orientation between center and surround (again only in the case of moving stimuli, not stationary flashing stimuli)