Crossover inhibition in the retina: circuitry that compensates for nonlinear rectifying synaptic transmission

Alyosha Molnar · Hain-Ann Hsueh · Botond Roska · Frank S. Werblin

Received: 20 January 2009 / Revised: 29 May 2009 / Accepted: 15 June 2009 / Published online: 28 July 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract In the mammalian retina, complementary ON and OFF visual streams are formed at the bipolar cell dendrites, then carried to amacrine and ganglion cells via nonlinear excitatory synapses from bipolar cells. Bipolar, amacrine and ganglion cells also receive a nonlinear inhibitory input from amacrine cells. The most common form of such inhibition crosses over from the opposite visual stream: Amacrine cells carry ON inhibition to the OFF cells and carry OFF inhibition to the ON cells (“crossover inhibition”). Although these synapses are predominantly nonlinear, linear signal processing is required for computing many properties of the visual world such as average intensity across a receptive field. Linear signaling is also necessary for maintaining the distinction between brightness and contrast. It has long been known that a subset of retinal outputs provide exactly this sort of linear representation of the world; we show here that rectifying (nonlinear) synaptic currents, when combined thorough crossover inhibition can generate this linear signaling. Using simple mathematical models we show that for a large set of cases, repeated rounds of synaptic rectification without crossover inhibition can destroy information carried by those synapses. A similar circuit motif is employed in the electronics industry to compensate for transistor nonlinearities in analog circuits.

Keywords Retina · Linearity · Rectification · Inhibition · Visual processing · Neural circuitry · Synaptic nonlinearity

1 Introduction

Retinal and cortical processing requires that some visual pathways retain a linear representation of visual signals. However, signal transmission across synapses and via spike generation is inherently nonlinear and rectifying. This rectification distorts the original visual signal and can destroy information in a given pathway because signals that fall below the rectification threshold are compressed. How, then, is linearity maintained in the face of rectification? In this paper we will show that in many cells, inhibition between the ON and OFF pathways acts to suppress the effects of synaptic rectification.

The retina parses the visual signal into two complementary visual streams; the ON and OFF pathways that first appear at the bipolar cells (Dacey et al. 2000; Kaneko 1970; Werblin and Dowling 1969). In these same cells, OFF cells receive ON inhibition and ON cells receive OFF inhibition.
(Molnar and Werblin 2007). A similar interaction takes place among amacrine cells (Hsueh et al. 2008) and ganglion cells (Roska et al. 2006). Recent work indicates this inhibition is predominantly glycinergic (Hsueh et al. 2008; Manookin et al. 2008; Molnar and Werblin 2007; Wassle et al. 1986). We refer to this class of inhibition as “crossover inhibition” because it is carried by amacrine cells that carry signals across the ON/OFF boundary of the IPL. There is evidence that additional forms of crossover inhibition continue in the lateral geniculate and at different levels in the visual cortex as well (Hirsch 2003).

The commonality of crossover inhibition, dominating more than half of the cell types in the inner retina, suggests that it performs an important function in the retina; yet crossover inhibition seems redundant: When light levels increase, the excitation to OFF cells intrinsically decreases, and it is not obvious how reinforcing this reduction with increased inhibition from the ON system would enhance the response of the cell.

Several possible functions for cross-over inhibition have recently been suggested. Recent work by Manookin et al. (2008) suggests that cross-over to OFF ganglion cells enhances the sensitivity (gain) of these cells to small changes in brightness at low light levels. Renteria et al. (2006) suggest that this same cross-over may act to suppress spurious ON signals that otherwise appear in the OFF pathway. We demonstrate here that crossover inhibition acts to suppress the distorting effects of synaptic rectification. The excitatory and inhibitory currents that impinge upon a given cell are rectified, but in opposing directions, so their recombination generates a membrane voltage representation of the visual signal that is more linear than either of its inputs. This is especially important in the subset of retinal pathways that respond to the average light intensity across their receptive field, distinct from the (nonlinear) pathways that selectively respond to local contrast. We further suggest that maintaining linearity in the temporal domain, at least in some retinal outputs, is necessary to prevent information loss due to interleaved rectification and adaptation. A similar “crossover” circuitry is routinely implemented in analog electronic circuits to perform all of the suggested functions of crossover in the retina, including compensating for the nonlinearities introduced by transistors.

In this study we look at crossover inhibition and rectification in several ways. First we review the basic evidence for, and spatial-temporal behavior of crossover inhibition. We then present evidence for the commonality of synaptic rectification, evidence that crossover inhibition suppresses the effects of this rectification, and a simple model that captures both of these effects. Using a combination of experimental and computational results, we then show why correcting for rectification at each stage of the retina is important even when this correction is followed by further rectification. Finally, we present some computational evidence for why a combination of synaptic rectification and crossover inhibition provides better signaling performance than simple linear synapses.

2 Methods

Eyes were extracted from rabbits and dissected as described previously (Roska et al. 2006). Ganglion cell recordings were made in flat mount (Fried et al. 2005; Roska et al. 2006), while amacrine and bipolar cell recordings were made from 250 μm slices (Molnar and Werblin 2007). Spiking in ganglion cells was recorded using loose attached patches. Excitatory and inhibitory currents were recorded under voltage clamp at -60 mV and 0 mV respectively. Membrane voltages were recorded under 0pA current clamp. Cell types were identified by imaging with Alexa Fluor 488 and comparing with previous work (MacNeil et al. 2004, 1999).

Slice mounted cells were stimulated by 200 μm wide light and dark flashes and sinusoidally varying intensities of ± 100%, relative to a set background illumination of 3×105 photons/μm2/s. Ganglion cell receptive field centers were found by flashing spots of varying size, from 100 μm to 1000 μm and recording spikes, excitation and inhibition. Sinusoids and gratings were sized to match the maximal spiking flash response for each cell. Gratings took the form of 50 μm, 75 μm and 100 μm stripes of ± 100% relative to background, and were flipped every 0.5 s. “Full” gratings were used, where each stripe was either 100% or -100% relative to the background, and each stripe was inverted at each transition. In the case of these gratings, every location on the grating saw a change at every transition but the average intensity across the whole grating was held constant. We also used two “partial” gratings where half of the stripes were held at background, such that all the remaining stripes were either 100% or -100% relative to background, and switched at each transition, driving ½ of the receptive field, but with a change in average intensity. The two partial gratings use complementary subsets of stripes such that in combination every stripe was driven. Conceptually, the full grating is just the sum of the two partial gratings.

To characterize the amplitude and sign of flash responses, we took the average response 200 ms immediately before and immediately after the onset and offset of each flash. Finding this value at the beginning and end of light and dark flashes yields 4 numbers:

\[d_{BL}, d_{EL}, d_{BD}, d_{ED}.\]
To extract the polarity of a given response, we computed a normalized term (Hsueh et al. 2008; Molnar and Werblin 2007).

\[ P = \frac{d_{BL} - d_{EL} - d_{BD} + d_{ED}}{|d_{BL}| + |d_{EL}| + |d_{BD}| + |d_{ED}|} \] (2)

Based on this metric, excitation to purely ON cells give a \( P = -1 \), and to OFF cells, \( P = +1 \). Similarly, ON inhibition has \( P = +1 \), and OFF inhibition has \( P = -1 \). ON-OFF responses give \(-1 < P < 1\), with the sign defining the dominant (ON vs OFF) input.

Cells were defined as receiving crossover inhibition, and so were included in the various analyses described below if the polarity numbers for excitation and inhibition had the same sign (implying opposite changes in conductance). We also excluded cell types for which electrical coupling degraded the effective voltage clamp to the point where we could not distinguish excitation from inhibition, specifically, we excluded all amacrine cells, which are readily identifiable by both anatomy and physiology (Hsueh et al. 2008) and which show strong electrical coupling to one another through gap junctions.

The measure of rectification, \( R \), was defined using the same data as:

\[ R = \frac{d_{BL} + d_{EL} + d_{BD} + d_{ED}}{|d_{BL}| + |d_{EL}| + |d_{BD}| + |d_{ED}|} \] (3)

Amplitude-modulated sine waves were analyzed by taking the Fourier transform of the response, extracting the phases for the fast and slow frequency terms, and taking the difference in phase between excitation and inhibition at each of these frequencies. Because phase is meaningless for cases where the slow (artifact) signal was below noise levels for either excitation or inhibition, we culled all cells from this analysis where either excitation or inhibition was less than the noise level at adjacent frequencies; this resulted in culling approximately \( \frac{1}{2} \) of cells from the modulated sine-wave analysis.

When quantifying spatial linearity in ganglion cells, we counted total spikes in response to six transitions each for partial (\( P_1 \) and \( P_2 \)) and full (\( F \)) gratings and subtracted the baseline number of spikes for the same period without stimulus (\( B \)). A measure of linearity, \( L \), was defined as the difference between responses to partial and full gratings, normalized by the partial response alone:

\[ L = \frac{P_1 + P_2 - F - B}{P_1 + P_2 - 2B} \] (4)

A more detailed description of the Methods can be found in the study of Molnar & Werblin (Molnar and Werblin 2007).

**Modeling methods** In our simulations and derivations, we used three different models of rectification. The most general model is a sigmoidal function:

\[ \text{release} \propto \left( \frac{1}{1 + \exp((v - v_0)/9mV)} \right)^4 \] (5)

Chosen to roughly model the calcium curve of synaptic release (Hille 2001) assuming four protein segments each with three fundamental charges. The baseline point of operation is assumed to be on the convex portion of the curve. For ease of computation in derivations and proofs, we have simplified this model in one of two ways: Either we approximate the rectification with a second order polynomial:

\[ y = a_1 x + a_2 x^2 \] (6)

As can be seen in Fig. 1, this is accurate for small changes in input level, and captures the curvature of the response. For larger signals, and for strong rectification, such as spike generation, we use a piecewise-linear model:

\[ y = (x + |x|)/2 \] (7)

which can be seen in Fig. 1 to be accurate for larger signals.

For simulations where filtering and adaptation occur, we used two models. The more general model approximates receptors as having three states: closed, open, and desensitized. The proportion of channels in each state was then simulated using the following differential equations for open (\( x \)) and desensitized (\( y \)).

\[
\frac{dx}{dt} = u(t) k_1 (1 - x - y) - k_2 x - k_3 x + k_4 y
\] (8)

\[
\frac{dy}{dt} = k_3 x - k_4 y
\]

As an alternate, simpler model, a LTI model was used where:

\[
\frac{dx}{dt} = u(t) k_1 - k_2 x - k_3 x + k_4 y
\] (9)

\[
\frac{dy}{dt} = k_3 x - k_4 y
\]

In all cases, membrane voltage is modeled as a weighted, signed summation of excitation and inhibition.

Simulations were performed in MATLAB. For the simple simulations shown in Fig. 7 and Fig. 10, inputs were passed through a simple linear time invariant high-pass filter (as in Eq. (9)), as well as a difference-of-Gaussians model of the outer retina for Fig. 10. Differential equations were modeled using difference equations iterated across many time-steps: time steps were made very small.
relative to the time-constants of the differential equations to avoid numerical artifacts: \( dt = 0.001\tau \). The rest of the model consisted of a series of static computations (i.e., rectification, cross-subtraction). At each time-step, input was drawn from a previously arranged time series. The input was multiplied by +1 and -1 to generate the two (“ON” and “OFF”) pathways. Half-wave rectification was then applied to both signals, modeled by the Boltzmann distribution taken to the 4th power (Eq. (5)).

For the model in Fig. 12, similar components were used but in a slightly different order: the input was first parsed into ON and OFF pathways, then rectified using a Boltzmann distribution (Eq. (5)) as above. This was followed by an LTI filter (Eq. (9)), as above, and followed by a simple piece-wise linear rectifier (Eq. (7)), to model spike generation. Crossover inhibition (shown in Fig. 12(b)) was modeled by subtracting the two high-pass filtered signals before the final round of rectification.

Finally adaptation, modeled in Fig. 13, involved a similar process, but with each synapse including both a rectifier as in Eq. (5) and a linear-time-varying model of adaptation (as in Eq. (8)). The action of two such synapses were computed in series for each pathway, followed by piece-wise linear rectifier (Eq. (7)), to model spike generation. Crossover inhibition was modeled by subtracting the signals after each synapse model, but before the subsequent round of rectification. For simplicity, we decided not to model the additional effects of the intermediate cells and synapses and treated them as mirroring the action of the “bipolar” cells.

3 Results

3.1 Basic characterization of crossover inhibition

We characterized excitatory and inhibitory synaptic inputs to bipolar, amacrine and ganglion cells in the rabbit retina in response to flashed light stimuli. As we have described elsewhere (Hsueh et al. 2008; Molnar and Werblin 2007; Roska et al. 2006), and consistent with the findings of others (Chen and Linsenmeier 1989a, b; Cohen 1998; Renteria et al. 2006; Wassle et al. 1986) we found that most OFF cells, which showed a maximum excitatory input when light intensity decreased, received inhibitory inputs that were strongest when light intensity increased. This suggests that while excitation is derived from the OFF system, inhibition is derived from the ON system. We refer to this phenomenon as “crossover inhibition.” Similarly, in a subset of cells that received predominant ON excitation, we found that the dominant inhibition was strongest at light OFF, implying a crossover inhibition from the OFF system to the ON system. As described in “Methods”, cells receiving crossover inhibition were identified automatically, based upon the polarity of excitatory and inhibitory responses to light and dark flashes. The proportion of cells identified as receiving crossover inhibition is shown in Table 1. Example light responses of cells showing this behavior are shown in Fig. 2. Data presented in subsequent sections and figures is drawn from the set of cells identified as receiving crossover inhibition.

Pharmacological experiments performed by our lab and others (Hsueh et al. 2008; Manookin et al. 2008; Molnar and Werblin 2007) have confirmed that these signals cross over from ON to OFF pathways by using L-AP4, which selectively blocks the excitatory inputs to ON bipolar cells (Slaughter and Miller 1981). L-AP4 has been shown to reliably block inhibition, but not excitation to OFF bipolar, amacrine and ganglion cells. Similarly L-AP4 reliably blocks excitation to ON cells, but in the majority of cases does not block crossover inhibition, revealing OFF inhibition to ON bipolar, ON ganglion and especially ON amacrine cells. Examples of the effect of L-AP4 on cells receiving crossover inhibition are shown in Fig. 2(d).

Crossover inhibition is common, but not universal, and the distribution of this crossover inhibition across cell types is not symmetric: crossover inhibition was the dominant form of inhibition in almost all OFF bipolar and ganglion cells, and in nearly all ON amacrine cells, but was found to dominate in only about half of OFF amacrine cells, and in less than half of ON bipolar and ganglion cells. Crossover inhibition was also found in combination with other types of inhibition, especially in those cell types where it was not usually dominant (i.e. ON ganglion cells). The
distribution of crossover inhibition across different cell types, based upon previous work in our lab (Hsueh et al. 2008; Molnar and Werblin 2007; Roska et al. 2006), is summarized in the first three columns of Table 1.

Based upon these observations, one can infer the circuitry that underlies crossover inhibition. The types of circuits and signals most likely leading to the currents and voltage response of an OFF ganglion cell are shown in Fig. 3(a-f). OFF excitation to this OFF ganglion cell originates in OFF bipolar cells, which respond to a light flash with a transient hyperpolarization at light ON and a transient depolarization of similar magnitude at light OFF as shown in Fig. 3(a). OFF bipolar cells form excitatory synapses with OFF ganglion cells generating a strong increase in excitatory current at light OFF, and, because of synaptic rectification, a somewhat weaker decrease at light ON. The ON bipolar cell (Fig. 3(b)) follows a mirror-image voltage trajectory, depolarizing transiently at light ON and hyperpolarizing transiently at light OFF. ON bipolar cells drive ON amacrine cells and the membrane of the ON amacrine cell, shown in Fig. 3(d), also depolarizes at light ON and hyperpolarizes at light OFF. ON amacrine cells form inhibitory synapses with OFF ganglion cells, and so inject a strong inhibitory current at light ON. At the OFF ganglion cell (Fig. 3(f)) crossover ON inhibition combines with OFF excitation to generate a hyperpolarization at light OFF.

Table 1 Frequency of crossover inhibition and average rectification in various broad cell classes

| General cell class | fraction with crossover | R (degree of Rectification) |
|--------------------|-------------------------|-----------------------------|
|                    | raw # | By type | Excitation mean | std | p | Inhibition mean | std | p | Voltage mean | std | p |
| OFF BC             | 44/48 (48/48) | 5/6 (6/6) | -0.26 | 0.25 | 5.4E-08 | 0.37 | 0.35 | 4.6E-07 | -0.12 | 0.48 | 0.037 |
| ON BC              | 26/49 (38/49) | 2/5 (3/5) | -0.09 | 0.2 | 2.5E-02 | 0.17 | 0.51 | 8.4E-02 | -0.15 | 0.43 | 0.04 |
| OFF AC             | 38/86 (62/86) | 6/13 (11/13) | -0.61 | 0.32 | 7.7E-07 | 0.21 | 0.54 | 3.7E-02 | 0.1 | 0.58 | 0.16 |
| ON AC             | 90/119 (106/119) | 12/17 (16/17) | -0.27 | 0.29 | 1.1E-08 | 0.28 | 0.4 | 1.1E-05 | -0.1 | 0.4 | 0.014 |
| OFF GC             | 44/67 (67/67) | 4/5 (5/5) | -0.46 | 0.34 | 1.3E-03 | 0.23 | 0.42 | 6.3E-02 | N/A | N/A | N/A |
| ON GC             | 10/67 (54/67) | 1/5 (4/5) | -0.2 | 0.21 | 1.3E-01 | 0.1 | 0.39 | 1.0E+00 | N/A | N/A | N/A |

Frequency of crossover inhibition across all cells measured, pooled from previous work and more recent experiments: numbers indicate proportion where crossover inhibition is dominant over other types of inhibition, numbers in parentheses indicate proportion where crossover is present, but not necessarily dominant. Cells were also categorized into various sub-types and frequency of crossover across sub-types is also shown. Cell type categorization was based upon morphological classification using depth, width and diffuseness of IPL stratification, as described in (Hsueh et al. 2008; MacNeil et al. 2004, 1999; Molnar and Werblin 2007; Rockhill et al. 2002; Roska et al. 2006). Rectification (not previously reported) was analyzed across all cells of a given class that received crossover inhibition. Mean, standard deviation and significance (analyzed using a Wilcoxon signed-rank test) were derived for each subclass. All amacrine cells were explicitly eliminated from this pool, as poor space clamp through gap junctions made it very difficult to separate excitation from inhibition (Hsueh et al. 2008).

Fig. 2 Crossover inhibition is a common property of bipolar, amacrine and ganglion cells. Excitatory (red, top row) and inhibitory (blue, middle row) currents were measured in example ON and OFF (a) bipolar, (b) amacrine and (c) ganglion cells in response to a stepped, bright flash. These inputs combine to generate the membrane voltage (bottom row) for bipolar (a) and amacrine (b) cells, or spiking rate (c, for ganglion cells) denoted in black. (d) Inclusion of L-AP4 in bath eliminates excitation but not inhibition to ON cells (in this case an amacrine cell) and inhibition but not excitation to OFF cells (in this case an OFF ganglion cell), indicating that the inhibition truly “crosses over”. Washing L-AP4 back out reverses these effects. Width of the white regions of each response block denote timing of the 2 s light flash.
ON and depolarization at light OFF. Similar circuitry can be deduced for bipolar and amacrine cells, as shown in Fig. 3(g and h).

An alternate possible circuit that could generate the same basic excitatory and inhibitory responses in Fig. 3(c, e and f) would be a circuit involving two amacrine cells in series: For example an amacrine cell providing apparent ON inhibition could depolarize at light on due to direct excitation from an ON bipolar cell, or because of reduced inhibition from an up-stream OFF amacrine cell. In this second case the final inhibition would appear to be “ON” inhibition but would actually originate entirely from the OFF system. APB results from our lab and others (Hsueh et al. 2008; Manookin et al. 2008; Molnar and Werblin 2007) indicate that in most cases this is not the dominant circuitry supplying crossover inhibition. However, given the commonality of crossover inhibition to inner retinal interneurons, it is likely that most inhibitory signals come from interneurons (amacrine or, indirectly, bipolar cells) which themselves received crossover inhibition. Thus, crossover inhibitory signals are likely to reflect a combination of true crossover signaling and indirect within-layer cascaded inhibition.

In the basic flash responses shown in Figs. 2 and 3, excitation and inhibition do not appear to overlap: when excitatory conductance is maximum (as indicated by large inward currents) inhibitory conductance is at a minimum (as indicated by a small outward current), and when inhibitory conductance is maximum (large outward current), excitatory conductance is minimum (small inward current). This leaves the question: what function does crossover inhibition provide?

One possibility we investigated was that crossover inhibition only provided functional benefits at fine timescales. To explore this possibility, we stimulated the retina with sinusoidally varying intensities and compared the relative timing (phase) of excitation and inhibition. In bipolar cells (Molnar and Werblin 2007), we have previously shown that the phase difference for excitatory and inhibitory currents stayed consistently close to zero degrees (which implies that their associated synaptic conductances were consistently 180 degrees apart) across a five octaves of frequency. We have confirmed this result in amacrine and OFF ganglion cells where crossover is dominant, as shown in Fig. 4. ON ganglion cells showing pure crossover are sufficiently rare (see Table 1) that we were unable to compile meaningful population data, but individual cases were consistent with other crossover inhibition receiving cells. Note that in a linear system, stimulation by a sinusoid of a given frequency will generate a sinusoidal response of the same frequency, but with (frequency dependant) shifts in amplitude and phase. Thus, in the context of linear processing, the temporal interactions between excitation and inhibition are well described by their phase difference. Because the excitatory and inhibitory currents reliably maintained zero degree phase difference under sinusoidal

---

Fig. 3 Signal flow of crossover inhibition to an OFF ganglion cell, in response to a 2 s light step (white portion of each frame). Likely behavior of presynaptic interneurons is illustrated with example traces. OFF bipolar cells (a) transiently hyperpolarize while ON bipolar cells (b) transiently depolarize at light ON, and both transiently polarize in the opposite direction at light OFF (black voltage traces). OFF bipolar cells generate an asymmetric inward excitatory post-synaptic current (c) in the OFF ganglion cell. ON bipolar cells excite ON amacrine cells (d) which generate an outward, inhibitory post-synaptic current (e) in the ganglion cell. These two currents are each strongly rectified but combine, in the absence of spiking, to generate a symmetric voltage signal in the OFF ganglion cell (f). Similar circuitry also appears in (g) bipolar (Molnar et al. 2002) and (h) amacrine cells (Hsueh et al. 2008). Red arrows indicate measured excitatory connections, blue arrows indicate inhibitory connections, black indicate presumptive presynaptic connections.
stimulation, it seems highly unlikely that crossover inhibition performs significant linear signal processing in time beyond simply increasing the amplitude of the cell’s response.

We also investigated the possibility that crossover inhibition performed linear spatial center-surround antagonistic processing due to different receptive field sizes of excitation and inhibition. In ganglion and displaced amacrine cells, we flashed bright and dark spots of varying diameter from 100 μm to 1 mm (Roska et al. 2006), and recorded the amplitude of excitatory and inhibitory responses (see Methods). In each case, we characterized the size spot that elicited a maximum response. We found that for a given cell type, the spots that elicited maximum responses for crossover inhibition were similar in scale to those for excitation, with crossover typically showing a maximum response to spots slightly larger (by 100 um or less) than excitation. This held true for all ganglion cell types that showed clear crossover inhibition, including OFF beta, OFF alpha, OFF parasol, OFF coupled, and ON bistratified cells, as well as displaced ON amacrine cells (Fig. 5). In addition, this also appears true in most cells where crossover was not dominant, but made up a clear sub-component of inhibition, specifically, ON Beta, ON parasol and OFF Delta cells, although the presence of other forms of inhibition makes this less certain (Roska et al. 2006). This is compared to other forms of inhibition, such as the surround antagonism seen in DS and local edge detector cells, where inhibition tends to reach a maximum at spot diameters two to three times the maximum response size for excitation, or compared to the transient ON-OFF wide-field inhibition seen in many cells which is always greatest for very large spots. Such close matching in size tuning implies that crossover inhibition is unlikely to perform straightforward linear spatial processing, since both signals encode the same sizes and drive cell membrane with the same polarity. It also implies that the crossover signal is carried by relatively narrow field amacrine cells or networks of cells, since wider field cells would be expected to expand the receptive field of crossover inhibition.

**Linear model of crossover inhibition** Linear signal processing is defined as any operation whose output can be described as a weighted average of the input (light signal) across time and space. Under this approximation, the ON system encodes increments in light intensity and the OFF system encodes decrements, but both systems are capable of encoding both increments and decrements. A decrement in intensity would lead to a decrease in activity in the ON system as well as an increase in the OFF system. These pathways may be modeled linearly as representing the same signal (changes in intensity) but with opposite signs. Thus for an input light signal $u(t)$, the responses of the ON and OFF systems could be represented as $x_{ON} = u(t)$ and $x_{OFF}(t) = -u(t)$.

In this context, crossover inhibition can be modeled as cross-subtraction between the ON and OFF pathways, as shown in Fig. 3. Both pathways pass though multiple synapses, and are averaged across the dendrites of multiple cells, such that the input to a given cell is better modeled as

$$x_{OFF} = -u(t,x,y) * f_{OFF}(t,x,y)$$

where $f_{OFF}(t,x,y)$ describes the impulse response of spatio-temporal filtering seen by the OFF pathway, which is convolved with the input to generate a filtered output. A similar description can be applied to the ON pathway, and crossover inhibition in an OFF cell acts to subtract an ON, inhibitory current from an OFF, excitatory current, to yield an output voltage:

$$y_{OFF} = x_{OFF} - x_{ON}$$

$$= -u(t,x,y) * f_{OFF}(t,x,y) - (u(t,x,y) * f_{ON}(t,x,y))$$

$$= -u(t,x,y) * (f_{OFF}(t,x,y) + f_{ON}(t,x,y))$$
Fig. 5 Spatio-temporal responses of various Ganglion cell types to bright spots of diameter 100 μm, 200 μm, 300 μm, 500 μm, and 1000 μm (Roska et al. 2006). Each row is a given size, each column is a cell type. Average conductance for each cell type is shown (conductance is used rather than current to allow comparison of synaptic inputs in a compact format): red is excitation, blue is inhibition, scale bars are 1.6nS. Black is spiking (scale bar is 50spike/s). Responses enclosed in boxes are examples that show crossover inhibition. Note that the maximum amplitude of excitation and crossover inhibition occur for the same size spots.
However, the spatial and temporal results above seem to imply that the ON and OFF pathways leading to excitation and crossover inhibition maintain similar temporal delays and spatial scale such that

$$f_{OFF}(t,x,y) \sim f_{ON}(t,x,y) = f(t,x,y)$$

and therefore

$$y_{OFF} = -2u(t,x,y)*f(t,x,y)$$

Thus because both excitation and crossover inhibition represent the same intensity change with the same sign, crossover would seem to perform little in the way of real signal processing aside from increasing signal gain (Manookin et al. 2008). What, then, is the functional significance of crossover inhibition?

3.2 Synaptic inputs are rectified, crossover inhibition compensates for this rectification

There is also the possibility that crossover is important for dealing with nonlinear processing. As Figs. 2 and 3 show, synaptic inputs to crossover cells are not simply linear representations of the light stimulus, which would show equal but opposite size steps in activity at the onset and termination of a stimulus. Instead, these inputs are rectified, encoding presynaptic depolarization much more strongly than hyperpolarization. A linear system such as described by Eq. (10) will encode the onset and offset of a flash symmetrically, whereas a pathway with rectification will respond to positive and negative signal changes with different magnitudes, distorting the linear response. As such, symmetry of a flash response can be used as an indicator of how linear the pathway that generated that response is: less symmetric responses imply a less linear, more heavily rectified signal pathway.

For example, looking again at Figs. 2 and 3, one can observe in the example cases that membrane voltages show roughly symmetric hyperpolarization and depolarization at light ON and OFF. However, the synaptic currents and spiking responses are much less symmetric, showing a large degree of rectification. In Fig. 3, the OFF excitatory signal showed a large increase in inward current at light OFF (a downward deflection) but only a weak decrease in this current at light ON. Similarly, The ON inhibitory current shows a large increase in outward current (an upward deflection) at light ON, but no decrease at light OFF. Thus, each synaptic input showed a complementary rectified response, with a large increase in conductance (and so current) when its presynaptic cells are depolarized, and a weak decrease in conductance when its presynaptic cells are hyperpolarized. The membrane voltage, however, showed a more symmetric response. We measured a depolarization at light ON, and equal but opposite hyperpolarization at light OFF. Thus, the combined currents, expressed as membrane voltage, create a representation of the stimulus that is much more symmetric, implying a more linear representation of the signal.

Rectification is a general property of crossover circuitry To test whether the symmetries and asymmetries seen in Figs. 2 and 3 were a general property of crossover circuits in the retina, we characterized the strength and polarity of rectification with a measure, R (see Methods). Histograms of this measure are shown in Fig. 6 for bipolar, amacrine and ganglion cells exhibiting crossover inhibition (for ganglion cells spiking is shown instead of voltage). Mean, standard deviation and significance of asymmetry are reported in Table 1 for each general class of cell (ON and OFF bipolar, amacrine and ganglion cells). In general, excitation was predominantly inward rectifying in both OFF cells and ON cells, shown in the top panels of Fig. 6. Inhibition, meanwhile, was predominantly outward rectifying in both OFF cells and ON cells, as shown in the middle panels of Fig. 6. Membrane voltage responses for bipolar and amacrine cells were on average quite symmetric in both ON and OFF cells, as shown in the bottom panels of Fig. 6(a-e) (as discussed below, spiking is much more rectified). As can be seen in Table 1 and in Fig. 6(e and g), the two cell types that show the least rectification in excitation are ON bipolar and ON ganglion cells. ON ganglion cells have been reported to have generally more linear inputs and outputs previously (Chichilnisky and Kalmar 2002; Demb et al. 2001; Zaghloul et al. 2003). However, Table 1 reveals some additional observations: ON bipolar and ON ganglion cells also show the least rectified crossover inhibition and, indeed, these two classes are the least likely to receive crossover inhibition in the first place. This may indicate cell types with more linear excitatory inputs require less correction from crossover inhibition.

The results shown in Fig. 6 and Table 1 indicate that in general excitation and crossover inhibition to inner retinal neurons are rectified: they respond more strongly to presynaptic depolarization than to presynaptic hyperpolarization. But, on average, the combination of excitation and inhibition results in voltage responses that are significantly more symmetric. However, these voltage responses are quite variable in their degree of rectification, as shown in Fig. 6. This variability is almost certainly due in part to experimental variation: bipolar and amacrine responses were recorded from retinal slices, which tend to cause some damage to the recorded cells’ dendritic and axonal trees as well as to the presynaptic cells which drive them. Such damage will result in a randomly reduced magnitude of the inputs which is not necessarily the same for
excitation and inhibition. This damage may also cause some leakage in post synaptic cells, shifting their resting potentials, and so altering the driving force associated with different synaptic channels. Thus, the relative weighting of excitation and inhibition in a given cell will almost certainly be somewhat shifted from its normal physiological state, increasing the apparent variability in membrane voltage rectification.

These average responses also encompass a wide variety of subtypes (as indicated in Table 1), not all of which are necessarily linear: indeed, known nonlinear OFF ganglion cell types have been shown to receive crossover inhibition (Zaghloul et al. 2003) indicating that crossover inhibition does not necessarily imply a linear response. Similarly, since a subset of ON ganglion cells have been shown to be linear (Chichilnisky and Kalmar 2002) even though most ON ganglion cells do not show purely crossover inhibition, linearity is possible without crossover inhibition, provided that the synaptic inputs are linear. This raises the possibility that the improvement in average voltage symmetry seen in Fig. 6 and Table 1, is a consequence of out-lying nonlinear cells, and that the more linear part of the population maintains this linearity because of intrinsically symmetric synaptic inputs, and not as a consequence of rectified crossover inhibition balancing rectified excitation.

To confirm that our population findings apply to linear cells (those with relatively symmetric voltage behavior in responses to flashed stimuli), we selected linear subsets of bipolar and amacrine cells (where we had significant amounts of voltage data) by only considering those cells whose voltage responses are reasonably symmetric. Specifically, we chose a threshold level of rectification, \( R_{thV} \) (where \( 0 < R_{thV} < 1 \)) and only considered cells for which \( |R(V)| < R_{thV} \), eliminating the cells with the most distorted voltages from consideration. The question we wanted to ask was: do cells with linear (symmetric) voltages show linear synaptic inputs as well, or are their inputs rectified, such that their linearity is a consequence of crossover inhibition.

In OFF bipolar, and ON and OFF amacrine cells, excitation and inhibition were consistently less symmetric (\(|R|\) was larger) than voltage, over the full range of possible values for \( R_{thV} \) with excitation consistently showing negative \( R \) numbers, and inhibition showing positive \( R \) numbers. More to the point, the magnitude of the rectification in excitation and inhibition was not significantly smaller for cells with symmetric voltage responses than for the population as a whole; even for the most linear subsets of cells, excitation and inhibition were still fairly asymmetric, with an average \( R \) greater than 0.2 in magnitude. The exception to this rule was in ON bipolar cells, where excitation was consistently more linear than voltage until \( R_{thV} \) was reduced below 0.3. This implies that 1) symmetric membrane voltage in response to light does not imply symmetric inputs from individual synaptic inputs (that is a linear voltage does not imply linear synaptic input), and 2) where voltage response is symmetric but synaptic input is not, crossover inhibition provides the balancing input to correct for synaptic rectification.

A simple model of rectification and crossover inhibition - Crossover inhibition has several properties that are closely analogous to differential electronic circuits. Such circuits encode signals in two parallel pathways which undergo similar filtering and amplification, and which are subtracted from each other at each stage of processing. Differential circuits, among other reasons, are often used specifically to suppress certain types of nonlinearity, in particular, rectifying nonlinearities from transistors. Since synapses and
spiking mechanisms have been shown to cause rectification in many cases (Katz and Miledi 1967), it seems reasonable to hypothesize that crossover inhibition might act in a similar fashion.

Analytically, nonlinear circuit components, such as transistors or synapses, are well approximated by a power series (such as Taylor series):

\[ f(x) = a_1x + a_2x^2 + a_3x^3 + \ldots \]  
\[ \text{(14)} \]

The even-order terms \((x^n, \text{where } n \text{ is even})\) of such a series describe the rectifying component of the nonlinearity, and are responsible for asymmetric responses such as seen in the currents shown in Figs. 1 and 2. This can be seen for even the simplest case of a second order power series:

\[ F(x) = a_1x + a_2x^2 \]  
\[ \text{(15)} \]

Which shows a larger change for positive changes in input \((x > 0)\) than for negative changes \((x < 0)\). Similar asymmetry can be seen for other models of rectification, such as a piecewise linear model, or a sigmoid operating at voltages below its inflection point (see Fig. 1, Methods).

The effect of crossover inhibition on rectification can be seen by starting with the same basic (linear) model as above, with two pathways with opposite signs (Fig. 7(a)) and applying the same second-order nonlinearity in [16] to both pathways (and taking \(x = y_{OFF} = -y_{ON}\)):

\[ F(y_{OFF}) = a_1x + a_2x^2 \]  
\[ \text{(16)} \]

and

\[ F(y_{ON}) = -a_1x + a_2x^2 \]  
\[ \text{(17)} \]

When one of these signals is subtracted from the other in a differential circuit (ie when one pathway is applied as excitation, and the other as inhibition), then, as illustrated in Fig. 7, the result is more linear than either input:

\[ F(x) - F(-x) = 2a_1x \]  
\[ \text{(18)} \]

And, in the ideal case, the even-order terms are eliminated (Magoon et al. 2002), yielding a perfectly linear output. As shown in Fig. 7(b), this simple model combined with some basic high-pass filtering predicts qualitatively similar results to those shown in Fig. 3. The elimination of even-order terms such as shown in Eq. (18) requires a perfect balance between the ON and OFF inputs, and so is unlikely to occur in real life. Nonetheless, even if they are imperfectly matched taking the difference between Eqs. of the form in (16) and (17) will lead to a result with a weaker 2nd-order term relative to the linear term. Similarly, even if crossover inhibition does not exactly match excitation, it will still lead to a suppression of rectification and its effects, as shown in Fig. 6.

3.3 Crossover inhibition suppresses artifacts generated by successive stages of synaptic rectification

Figures 2, 3, 4, 5, 6 and 7 indicate that while synapses tend to rectify the responses to light flashes, crossover inhibition will tend to suppress the effects of that rectification. However, this suppression appears futile, since in each case, once the voltage is relinearized by crossover inhibition, the signal is rectified again at each subsequent stage of synaptic transmission, or by spike generation. For example, compare Fig. 3(f), which shows a linear ganglion cell voltage representation of a light flash, with Fig. 2(c), which shows the final (strongly rectified) spiking response of a similar cell. What benefit, then, does crossover inhibition provide by correcting rectification, if the signal is re-rectified at the next processing stage?

Work in electronic circuit design has shown that successive stages of rectification without intervening cross-subtraction can generate artifacts that cannot be removed by cross-subtraction further down the chain. In the simplest case, consider the effect of concatenating two instances of rectification described by the second-order power series from above:

\[ F(F(x)) = a_1^2x + (a_1^2a_2 + a_1a_2)x^2 + 2a_1a_2^2x^3 \]
\[ + a_2^3x^4. \]  
\[ \text{(19)} \]

Subtracting the two pathways after this cascaded rectification now gives:

\[ F(F(x)) - F(F(-x)) = 2a_1^2x + 4a_1a_2^2x^3 \]  
\[ \text{(20)} \]

Even though we cross-subtract in the end, there is now a (3rd-order) distortion term that is solely a consequence of the rectifying nonlinearities, but which is insensitive to the final cross-subtraction. This distortion term is especially important since it tends to amplify large signals more than small ones, effectively reducing the system’s gain (and so sensitivity) to small, hard-to-detect inputs compared to its gain on already large (and easily detected) inputs.

When cross-subtraction is included between rectifying stages, we have

\[ F(F(x) - F(-x)) - F(F(-x) - F(x)) = 4a_1^2x \]  
\[ \text{(21)} \]

and this distortion term is now removed, selectively increasing the gain of small signals relative to larger signals, as has been reported by Manookin et al. (2008).

A second reason rectification is suppressed in electronic circuits is that even-order nonlinearity generates artifacts not present in the original signal (Razavi 1998) that exist at dramatically different frequencies than in the
original input (Lee 2004). For example, an amplitude modulated signal

\[ x(t) = (1 + \cos(2\pi f_2 t)) \cdot \cos(2\pi f_1 t) \] (22)

only contains spectral components at frequencies \( f_1 \) and \( f_1 \pm f_2 \).

However, if this signal is subjected to a second order nonlinearity such as in Eq. (15) the strongest additional spectral component that appears is at frequency \( f_2 \) (the modulation frequency) which was not present in the original signal, and which may be at significantly lower frequencies than the input.

This becomes a problem in systems where subsequent filtering, such as described by Eq. (10) is intended to suppress some frequency bands and not others. Rectification plus filtering, then, can generate ambiguities between very different inputs, or, as we will show below, can even destroy information altogether. By correcting for rectification, and suppressing its artifacts before subsequent processing, cross-subtraction in differential circuits remove these ambiguities.

Similar effects can interfere with visual function. In the time domain, rectification leads to artifacts that can confuse contrast with brightness: in an OFF cell receiving a rectified input, a slow decrease in brightness or a slow increase in contrast will both lead to an increase in average excitation. This potential confusion can be seen by taking the mean of Eq. (15),

\[ \langle F(x) \rangle = (a_1 \bar{x} + a_2 \bar{x}^2) = a_1 \bar{x} + a_2 \bar{x}^2 \] (23)

which shows that the average output of \( F(x) \) will reflect not just the mean of \( x \), but also its variance (from the \( x^2 \) term). Since variance is a common way of defining the amount of contrast in a visual signal, this implies rectification will tend to confuse average brightness with contrast. Combining the results of Eqs. (23) and (18) implies that crossover inhibition can reduce or eliminate this confusion by suppressing the variance term.

We demonstrate that contrast artifacts are generated by synaptic rectification and then corrected by crossover inhibition in Fig. 8. Here we stimulated the retina with an amplitude-modulated sinusoid as in Eq. (22), and recorded from an OFF bipolar cell. The input, shown in Fig. 8(a), contains only high frequency components at the “carrier” frequency, \( f_1 \) (1.2 Hz) and “sidebands” \( f_1 \pm f_2 \) (0.9 and 1.5 Hz), generated by modulating the “carrier” with the envelope frequency \( f_2 \). Rectification of this signal generated a new component at the envelope frequency \( f_2 \) or 0.3 Hz (dashed lines in Fig. 8(b)), which appeared in both excitatory and inhibitory currents. Amplitude modulation is equivalent to changing the temporal contrast of the signal (Zaghloul et al. 2005), so the appearance of this artifact implies that rectification tends to confuse slow changes in contrast (the envelope of the fast signal) with slow changes in brightness. When these inputs are combined, however, the fast components of excitation and inhibition, representing the original signal, are in-phase, and so add constructively, whereas the components of excitation and inhibition generated by rectification are out-of-phase, and therefore cancel. Thus, crossover inhibition enhances the original
signal and suppresses rectification artifacts even when these signals are at very different time scales. The result is a reconstructing the original signal, as shown in Fig. 8(c). In order for crossover to effectively suppress rectification artifacts, these phase relationships must be maintained: if there is some form of filtering, and associated phase shift between artifact generation and crossover, then the crossover will not be able to suppress rectification.

We found this phase relationship between excitatory and inhibitory currents across OFF bipolar cells (N=9), ON (n=10) and OFF (n=9) amacrine cells and OFF ganglion cells (n=6). ON bipolar and ganglion cells, which generally show relatively linear excitatory inputs and less crossover inhibition, were not included in this analysis. We performed this analysis for three different combinations of fast- and slow-frequencies: 1.2 Hz x 4.8 Hz, 0.3 Hz x 1.2 Hz, and 0.3 Hz x 4.8 Hz.

**Crossover inhibition suppresses spatial artifacts generated by synaptic rectification**

Synaptic rectification can also confuse brightness with spatial contrast as shown by the experiment illustrated in Fig. 9. We recorded spikes from OFF ganglion cells while stimulating the retina with high-contrast gratings that covered the center of the ganglion cells’ receptive field. The grating stripes were inverted every 0.5 seconds: dark stripes became bright and bright stripes became dark, but the average brightness was held constant.
Spiking responses to these “full” gratings were compared to “partial” cases where only half of the stripes were switched between bright and dark, and the remaining stripes were held at background levels, resulting in a net change in brightness. When transitioning from light to dark, these partial gratings elicited strong spiking from OFF ganglion cells, whereas full gratings elicited little spiking at either transition, as shown in Fig. 9. The cell shown behaved linearly, similar to the behavior of X-cells in cats (Hamasaki and Sutija 1979); they responded only to net changes in brightness across the entire center of their receptive field, and not to local changes within their receptive field. This behavior is distinct from the behavior reported in nonlinear Y-type ganglion cells (Demb et al. 2001), which respond to the sum of the activity in their rectified subunits, and whose response to full gratings is approximately equal to the sum of the responses to the two partial gratings. In order for a ganglion cell to behave in a linear fashion, either its subunits must be linear (with rectification occurring only after synaptic inputs have been summed across the dendritic field) or else the effects of rectification in its subunits must be compensated by crossover inhibition, permitting ON activity in one subregion of the cell’s receptive field to suppress OFF activity in other regions.

We assessed the contribution of crossover inhibition to the linearity of the spiking response by blocking the ON pathway with 20 μM APB. In the presence of APB, spiking in response to partial gratings was unaffected, but the response of the cell to each transition of the full grating dramatically increased, as shown in Fig. 9(c). The fact that partial gratings responses were unaffected implies that the increased response to full gratings is a result of increased nonlinearity and not a consequence of changing baseline activity or spiking threshold. This nonlinear response is almost certainly due to rectification in individual bipolar cell subunits, followed by summation across the ganglion cell’s dendritic field, as we model in Fig. 10. With crossover inhibition intact, these artifacts of rectification were suppressed, and the cell remained silent during the transitions of the full grating. In the presence of APB, however, ON crossover inhibition was removed, and only the rectified, OFF-bipolar mediated pathways were active, leading to a nonlinear response. This nonlinear behavior in an X-type ganglion cell under APB is similar to the baseline responses previously reported in nonlinear Y-type ganglion cells (Demb et al. 2001): thus, crossover inhibition acts to differentiate linear OFF cells from nonlinear OFF cells.

We quantified linearity with a measure L (see Methods), reflecting the relative response to partial versus full

---

Fig. 9 Crossover circuitry maintains linearity in the presence of high spatial contrast. Columns (a) and (b) show spiking histogram responses to partial gratings. Column c shows spiking histogram response of full grating where the bars of the grating invert so that there is no change in overall luminance at the receptive field center. The top of each column shows successive frames of the stimulus, in each case, back areas turn white, and white areas turn black, while grey areas maintain a constant, intermediate brightness. Control: Partial gratings elicit OFF responses in a and b, but minimal response to full grating in c. APB: Partial gratings still elicit OFF responses, but without ON crossover inhibition the full grating (c) also elicits a strong, nonlinear response. Wash: Linear behavior is restored to the full grating response in c. The gray images at the top of columns a-c each show three successive frames of the 0.5 sec/frame movie. Ordinate: spikes per second; abscissa time in seconds. (d) Histograms of linearity coefficient L with and without APB (histograms correspond to the same rows as (a-c))
Histograms of $L$ across 9 OFF cells and three different spatial scales are shown in Fig. 9(d): $L=1$ corresponds to a perfectly linear response, $L \leq 0$ corresponds to an extremely nonlinear response. This sample included a variety of subtypes, of which 4 were relatively nonlinear ($L < 0.5$, corresponding primarily to alpha and delta cells) and the rest of which were quite linear (beta and parasol cells); However, regardless of their baseline linearity, all of these cells showed less linear responses under APB, with cells that started out less linear changing slightly less under APB (correlation coefficient between initial linearity and change in linearity under APB=0.3). Overall, the linearity of OFF ganglion cells decreased under APB by $-0.38 \pm 0.46$ (mean±s.d) $p=6 \times 10^{-4}$ (Wilcoxon signed-rank test), and recovered under wash to values equivalent to control ($p=0.46$). Thus, linearity in OFF ganglion cells is not inherent to the feed-forward pathway that drives them, but is actively maintained by crossover inhibition from the ON system.

This interaction is modeled in Fig. 10. We start by treating the input as having two regions, with each region feeding into ON and OFF pathways, both of which contain some adaptation (high-pass filtering). Thus for partial gratings, one of these regions is stimulated but the other is not, yielding very different responses (compare red vs pink for model bipolar cells, blue vs cyan for model amacrine cells). These signals are then all rectified, and summed in a single ganglion cell. For partial gratings, excitation and inhibition are largely non-overlapping, as in Figs. 2 and 3, and the resulting sum is re-rectified to generate spiking: inhibition has no effect on the output, and removing it has no effect. However, for full gratings, both
pathways are activated, resulting in an overlapping of excitation from one region with inhibition from the other. Thus, the inhibitory ON signal from one region suppresses the OFF inhibition from the other, and no spiking results. However, if this ON inhibition is suppressed, then excitation is no longer blocked, and significant spiking results.

Ganglion cells combine and average the inputs of many bipolar and amacrine cells. For purely linear pathways, this summation acts to average or “low-pass filter” the visual signal across space, responding selectively to features on the same spatial scale as the cell’s receptive field, and suppressing responses to smaller features. As seen in both Fig. 10 (a and b), if the averaging is performed across a set of signals with the same polarity of rectification, then this selectivity is lost, and the cell responds to stimuli with a variety of feature sizes. This difference is seen even though the final output is rectified by spike generation in both cases. Although the second round of rectification masks any hyperpolarizing effects of crossover inhibition, it does not mask the effects of crossover when low pass filtering occurs between stages of rectification. The basic implication is that low-pass filters, when inserted between rectifying stages, lose their selectivity if the preceding stage is not compensated by crossover inhibition.

If interleaving rectification with low pass filtering disrupts the selectivity of that filter, how does interleaving rectification with high- or band-pass affect the system response?

This question can be analyzed formally by modeling two rounds of rectification interspaced with a time-derivative (the simplest high-pass filter possible). This case is shown in Fig. 11, for both the ON and OFF pathways.

In each case two stages of piece-wise linear rectification are applied:

$$F(x) = \frac{x + |x|}{2}$$

Separated by a derivative with respect to time, such that the output is:

$$ON = F\left(\frac{d}{dt}(F(x))\right)$$
$$OFF = F\left(\frac{d}{dt}(F(-x))\right)$$

This circuit can be completely analyzed by looking at four input cases:
1) $x>0$, $dx/dt>0$, 2) $x<0$, $dx/dt<0$, 3) $x>0$, $dx/dt<0$, and 4) $x<0$, $dx/dt>0$. In the first case, $ON=dx/dt$, $OFF=0$, and in the second case, $ON=0$, $OFF=-dx/dt$, such that the input $x$ (really its derivative) is encoded in the difference of the outputs. In the other two cases, however, where the input and its derivative have opposite signs, both ON and OFF are zero. This implies that for half of all input scenarios, the output of the system conveys no information about its input.

Including cross-subtraction between the pathways implies:

$$ON = F\left(\frac{d}{dt}(F(x) - F(-x))\right) = F\left(2\frac{dx}{dt}\right)$$
$$OFF = F\left(\frac{d}{dt}(F(-x) - F(x))\right) = F\left(-2\frac{dx}{dt}\right)$$

such that one of the outputs always represents the input.

Thus, whereas low-pass filtering between rectifying stages leads to reduced selectivity and increased activity in both ON and OFF pathways, high-pass filtering interleaved with rectification has the capacity to entirely destroy information.

More generally, our models and measurements indicate that rectification creates artifacts not present in the original input, (Eq. (23), dashed lines in Fig. 7(b)). If this rectification is followed by adaptation (high-pass or band-
pass filtering), the system may adapt to these artifacts instead of the original signal level. In this case, baseline activity is reduced, and may fall below the threshold of subsequent rectifying stages, along with any weaker signals close to this baseline, destroying those signals. This can be demonstrated in the simulation (see Methods) shown in Fig. 12(a). Splitting the input into parallel pathways that are rectified and then high-pass filtered does not eliminate any of the transitions in the input. However, because much of the filter response now falls below baseline levels (dashed lines in 3rd row, Fig. 12(a)), successive rectification can eliminate those parts of the response, completely eliminating transitions present in the input from the final output. Thus, interleaved rectification and adaptation can destroy information. Crossover inhibition eliminates rectification artifacts and so stabilizes the baseline level of the system so that even weak changes in the input generate above-threshold responses in one of the two pathways, as shown in Fig. 12(b). Thus, in this model (and others, see Fig. 14) inclusion of crossover inhibition between rounds of rectification prevents loss of information even in the presence of intervening filtering. Since the retina shows rectification at each stage, and since many glutamatergic synapses show a degree of adaptation (Lukasiewicz et al. 1995; Tran et al. 1999), it seems likely that crossover inhibition acts to reduce information loss in the visual system, but future experimental work will be required to prove this assertion.

**Fig. 12** Numerical simulation of rectification interleaved with high pass filtering: (a) Without interleaved crossover, this leads to loss of information in the output; all traces plot signal level versus time. The input signal (top) is parsed into positive and negative replicas and rectified. These signals are then high-pass filtered and rectified again. Because rectification of the filtered signals occurs before the final crossover subtraction, several transitions in the latter part of the input do not appear in the output. (b) Simulations where crossover inhibition is included between rectifying stages. The input signal is parsed, rectified and filtered as in (a). These filtered signals are then cross-subtracted before they are rectified again. When these rectified signals are cross-subtracted (x) they generate an accurate high-passed version of the original input, representing every transition in input level.
Crossover inhibition provides circuitry that compensates for the nonlinear behavior of synapses, so that linear signal streams can be maintained through the visual pathway. This crossover inhibition combines rectified excitatory currents with rectified crossover inhibitory currents to reconstruct a linear, non-rectified membrane voltage in bipolar, amacrine and ganglion cells (Figs. 1, 2 and 3). This reconstruction suppresses artifacts in time (Fig. 9) and space (Fig. 10), maintaining linear activity throughout the retina. The crossover signals are inhibitory in the conventional sense: they are usually mediated by glycine and modulate chloride channel conductance (Hsueh et al. 2008; Molnar and Werblin 2007; Wassle et al. 1986). Functionally, however, crossover inhibition responds to stimuli by driving membrane voltage with the same polarity as excitation, hyperpolarizing the cell when excitation is weak while disinhibiting the cell when excitation is strong. So, from a linear perspective, crossover inhibition acts in concert with and serves to augment, rather than suppress excitation. This augmentation appears to hold across a variety of spatial and temporal scales, such that inhibition enhances excitation across a wide range of spot sizes and for stimuli ranging from 2 s flashes (Fig. 2 (a–c)) to 10 Hz sine waves (Fig. 4).

This matching in time is slightly surprising given that crossover pathways imply at least one additional neuron and synapse, which should cause some delay. It seems that either the cells and synapses involved are sufficiently fast compared to the bandwidth of the stimulus (10 Hz) so as to not affect this phase, or that the synapses act in some way to equalize the phase delay intrinsic to the crossover pathways. The speed of this transmission is not entirely surprising given that most evidence points to crossover inhibition being carried by glycinergic amacrine cells (Hsueh et al. 2008; Manookin et al. 2008; Molnar and Werblin 2007) such that inhibitory synaptic transmission will be relatively fast (Eggers and Lukasiewicz 2006). Furthermore since vertically oriented, glycinergic cells are generally thought to have narrow dendritic fields (MacNeil et al. 1999; Wassle et al. 1998; Wassle et al. 1986) an idea consistent with the size matching in Fig. 5, electrotonic signaling within these cells is also likely to be quite fast. It therefore seems likely that crossover inhibition primarily serves to maintain and enhance rather than to shape the basic light response.

Crossover inhibition may also serve to maintain a more constant input conductance to the cell: as the excitatory conductance increases, the inhibitory conductance decreases, and vice versa. In this way other synaptic inputs to the cell are less affected by the changes in membrane conductance of a given synapse. This process of crossover inhibition appears to be repeated at higher visual centers: similar “push pull” circuitry has been inferred at the lateral geniculate nucleus and visual cortex (Anderson et al. 2000; Hirsch 2003; Lauritzen and Miller 2003).

Functional benefits of combining rectification and crossover inhibition The above results and analysis indicate that many synapses are rectifying, and that in many cases, throughout the inner retina, the effects of this rectification are suppressed by crossover inhibition. Stated simply, crossover inhibition acts to suppress the effects of synaptic rectification. However, the above evidence (especially in ON bipolar and ganglion cells) also indicates that synaptic rectification is not universal, and indeed, many synapses seem to behave quite linearly. This raises the question: why do many synapses in linear pathways operate in the rectifying part of the curve? Although this question is difficult to address experimentally, modeling of these synapses indicate several possible, mutually compatible benefits of rectification.

One possible answer is that a rectifying synapse has a lower level of tonic release than a synapse operating in the middle of the calcium curve. This reduced activity places less stringent requirements on the synapse in terms of energy requirements and size (a higher tonic release implies more vesicles). Although reducing synaptic release is unlikely to have a significant impact on the overall energy and size of the retina, it may provide significant local benefits in terms of metabolism.

A second likely benefit of a rectifying synapse is that it is less noisy. This has been demonstrated experimentally for the rod-bipolar synapse (Field and Rieke 2002; Rieke 2000) and seems likely to hold for other synapses. Assuming that the primary source of noise is vesicular release (and assuming this can be well described by a Poisson process), and assuming that the mean rate of release is a sigmoidal function of presynaptic voltage, rectifying synapses can be mathematically proven to provide lower noise than non-rectifying synapses (see appendix). In this context, the combination of rectification and crossover inhibition may act to maximize the dynamic range of the retina by simultaneously suppressing noise through rectification while extending the range of linear signaling through crossover inhibition.

Computational modeling also indicates that rectification, combined with crossover inhibition, permits contrast gain control (Baccus and Meister 2002; Beaudoin et al. 2007) in an otherwise linear pathway. As shown in Fig. 8, rectification tends to generate artifacts that are proportional to the contrast of a signal, as well as its brightness. This artifact can provide a signal for gain adaptation. This signal, if adapted to by a linear-time-invariant filter, and not compensated by crossover inhibition, can cause a shift in baseline and loss of information, as shown in Fig. 12. More
reasonable models of adaptation (such as Eq. (8)) cause changes in both absolute level (as occurs in a typical linear-time invariant high-pass filter) but also changes in gain. We modeled these combined phenomena with a two-part synapse model, shown in Fig. 13(a). The first part was a sigmoidal, static calcium curve, the second part was a 2nd-order differential equation to capture activation and desensitization of receptors (see Eq. (8)) methods). The overall model comprises two such stages each for ON and OFF pathways, followed by piecewise linear rectification to model spike generation. Crossover inhibition is modeled by cross-subtraction between the two pathways.

In order to test for contrast adaptation, we provided fast periodic inputs, whose amplitude dropped by ½ partway through the simulation. In the baseline case, with both rectification and crossover inhibition (Fig. 13(b)), we see that gain adapts, decreasing for high-contrast signals, and increasing when the contrast decreases. In particular, changes in contrast cause both a reduction in amplitude and in absolute level. This absolute level shift is analogous to the “after potential” reported in (Manookin and Demb 2006). Crossover inhibition suppresses this shift in absolute level, but retains the change in amplitude. Stated differently, crossover inhibition stabilizes the baseline activity under synaptic adaptation, permitting contrast gain control without loss of information. Shifting the quiescent voltage of each model cell to the linear (non-rectifying) portion of the calcium curve eliminates contrast gain control (Fig. 13(c)): rectification allows adaptation to these slow signals. Alternately, eliminating crossover inhibition permits the variation of baseline in response to changes in contrast. As a result, more of the signal is clipped by subsequent rectification, destroying information as shown in Fig. 13(d). Thus, the model indicates that effective contrast gain control should benefit from both rectification and crossover inhibition. One prediction of this model is that eliminating crossover inhibition would increase the magnitude of after hyperpolarizations due to high-contrast signals. This prediction is born out by pharmacological results presented in (Manookin and Demb 2006).

The nervous system is but one example of how cross-subtraction between parallel, complementary signaling pathways can compensate for nonlinear circuit elements. Differential circuitry (Leibowitz et al. 2005; Razavi 2000) is one of the hallmarks of modern analog electronic circuit design. Through circuitry that is similar to that in the retina, differential circuits parse signals into a pair of complementary streams, one positive, the other its negative counterpart, and information is carried in the difference between these streams. At each stage of processing, the negative stream is subtracted from the positive stream (and vice-versa) reinforcing the differential representation of the original signal, as diagrammed in Fig. 7, and equivalent to the circuits shown in Fig. 3. This subtraction suppresses signals that are introduced in common to both streams, including artifacts from the rectifying nonlinearity inherent to their constituent transistors (Magoon et al. 2002).

Fig. 13 Model of rectifying, adapting synapses with crossover inhibition. (a) schematic of model, blocks labeled “syn” contain sigmoidal rectification plus adaptation, blue lines indicate crossover subtraction paths, and blocks labeled “rect” contain a piece-wise linear rectifier to model spike generation. To better reveal the various effects of the model the final output is taken to be the difference the two rectified pathways, essentially reconstructing a processed version of the input. (b) Baseline behavior at input, after each synapse, after spiking rectification, and after recombining paths. Response shows contrast adaptation without information loss (c) same as b, but with synapses biased into the linear part of sigmoid: contrast adaptation is lost. (d) Same as (b), but without crossover inhibition: adaptation is extreme, information is lost when contrast us reduced.
This suppression is necessary partly because transistors are intrinsically nonlinear, and so always rectify their signals some degree. The degree of rectification is often enhanced however, because it is generally observed that transistors operated in their most nonlinear “subthreshold” mode provide optimal performance in terms of noise and efficiency (Steyaert and Sansen 1987). Furthermore, this mode of operation is useful for detecting signal amplitude and building variable gain circuits (Grey and Meyer 1993).

Based upon our modeling, above, it seems reasonable to hypothesize that the retina gains similar benefits from rectifying synapses.

In general, differential electronic circuits are used for their greater robustness in the face of the variations, imperfections and distortions introduced by their underlying components: rather than performing explicit signal processing, differential circuits mostly act to maintain signals in the presence of various forms of interference and disruption. That crossover inhibition plays a similar, multi-functional role in the retina is supported by the work of other groups as well as our own: crossover increases ganglion cell sensitivity to low contrast signals (Manookin et al. 2008), and suppresses apparent cross-talk between the ON-and OFF systems (Renteria et al. 2006).

Crossover inhibition is not universal, however, and is distributed asymmetrically between the ON and OFF systems (as outlined in Table 1). The functional significance of this asymmetry remains a mystery, and needs additional work to be understood. For example, it may be that the asymmetry in crossover inhibition and rectification are a function of the baseline adaptation state of the retina, with ambient light driving the OFF system into a rectifying state, and that under other illumination states this would change. Regardless, the apparent correlation between rectification and crossover inhibition would seem to indicate that they are connected. One possibility is that crossover inhibition is present because it suppresses rectification. Alternately, it may be that crossover is primarily present for some other reason, but once present, it permits synaptic rectification without signal distortion, and so permits lower noise and better adaptation.

All information processing systems are intrinsically limited by the behavior of the physical components of which they are comprised. It is generally the case that efficient, low noise signaling components (transistors, synapses, axons) also show significant nonlinearity. Since such nonlinearities can distort or even destroy the information carried by the system, modern electronic circuit topologies are designed to compensate for these limitations. Crossover inhibition in the mammalian retina is an elegant example of how the topology of inhibitory circuitry in the nervous system can also overcome the limitations of its constituent neurons and synapses.

**Appendix: Proof that rectification leads to maximal signal-to-noise ratio**

The proof is as follows: Assume that: 1) the dominant noise source in a synapse is vesicular release, and release is a Poisson process, such that the variance in the rate of release is proportional to the mean level of release; 2) that the average rate of release is a sigmoidal function of presynaptic voltage $F(V)$ such that $F(V)$ rises monotonically with voltage, starting at zero and saturating at a higher, positive level, and has a single inflection point. The behavior of a typical calcium channel is well described by such a sigmoid (Hille 2001).

Noise in a given electronic component (be it a synapse or a transistor) is important to the extent that it decreases a pathway’s signal-to-noise ratio (SNR). SNR is maximum when an input signal (a change in presynaptic voltage, $\Delta V$), generates an output signal (change in synaptic release $\Delta R$) that is maximized compared to added noise from the synapse. To first order, the expected change in release, $\Delta R$ is proportional to the input, $\Delta V$ times a gain term, $G$. $G$ is proportional to the slope of the Calcium curve at the quiescent (resting) potential, $V_0$ of the presynaptic cell:

$$\Delta R \propto \Delta V \left. \frac{dF(V)}{dV} \right|_{F=F_0}$$

(27)

At the same time, the variance (noise, squared) of release is proportional to the baseline level of release:

$$\text{var}(\Delta R) \propto F(V)$$

(28)

The signal to noise ratio, then is just the ratio of these signals:

$$\text{SNR} = \frac{\Delta R}{\sqrt{\text{var}(\Delta R)}} \propto \frac{\Delta V \left. \frac{dF(V)}{dV} \right|_{V=F_0}}{\sqrt{F(V)}}$$

(29)

![Fig. 14 Model calcium curve as well as expected gain and SNR (arbitrary scale). Peak gain and maximum linearity occur near $F(V) = 0.5$, peak SNR is always below this point](image)
SNR will be at a maximum when the baseline voltage, V is such that:

\[
\frac{d(SNR)}{dV} = 0 \Rightarrow 0 = \frac{d(SNR^2)}{dV} = \frac{d}{dV} \left( \frac{(dF(V))}{d\sqrt{F(V)}} \right) \text{ } (30)
\]

\[
= \frac{2 \frac{dF(V)}{dV} \frac{d^2F(V)}{d^2V} F(V) - \left( \frac{dF(V)}{dV} \right)^3}{F(V)^2}
\]

Since both F(V) and its derivative are strictly positive, SNR will only be maximized when its second derivative is also positive, i.e., when V is at a level where F(V) is rectifying, as illustrated in Fig. 14.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Anderson, J. S., Carandini, M., & Ffister, D. (2000). Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. Journal of Neurophysiology, 84, 909–926.

Baccus, S. A., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. Neuron, 36, 909–919.

Beaudoin, D. L., Borghuis, B. G., & Demb, J. B. (2007). Cellular basis for contrast gain control over the receptive field center of mammalian retinal ganglion cells. Journal of Neuroscience, 27, 2636–2645.

Chen, E. P., & Linsenmeier, R. A. (1989a). Centre components of cone-driven retinal ganglion cells: differential sensitivity to 2-amino-4-phosphonobutyric acid. Journal of Physiology, 419, 77–93.

Chen, E. P., & Linsenmeier, R. A. (1989b). Effects of 2-amino-4-phosphonobutyric acid on responsivity and spatial summation of X cells in the cat retina. Journal of Physiology, 419, 59–75.

Chichilnisky, E. J., & Kalmar, R. S. (2002). Functional asymmetries in ON and OFF ganglion cells of primate retina. Journal of Neuroscience, 22, 2737–2747.

Cohen, E. D. (1998). Interactions of inhibition and excitation in the light-evoked currents of X type retinal ganglion cells. Journal of Neurophysiology, 80, 2975–2990.

Dacey, D., Packer, O. S., Diller, L., Brainard, D., Peterson, B., & Lee, B. (2000). Center surround receptive field structure of cone bipolar cells in primate retina. Vision Research, 40, 1801–1811.

Demb, J. B., Zaghoul, K., Haarsma, L., & Sterling, P. (2001). Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (Y) ganglion cell in mammalian retina. Journal of Neuroscience, 21, 7447–7454.

Eggers, E. D., & Lukasiewicz, P. D. (2006). Receptor and transmitter release properties set the time course of retinal inhibition. Journal of Neuroscience, 26, 9413–9425.

Field, G. D., & Rieke, F. (2002). Nonlinear signal transfer from mouse rods to bipolar cells and implications for visual sensitivity. Neuron, 34, 773–785.

Fried, S. I., Munch, T. A., & Werblin, F. S. (2005). Directional selectivity is formed at multiple levels by laterally offset inhibition in the rabbit retina. Neuron, 46, 117–127.

Grey, P., & Meyer, R. (1993). Analysis and design of analog integrated circuits. New York: Wiley.

Hamasaki, D. I., & Sutija, V. G. (1979). Classification of cat retinal ganglion cells into X- and Y-cells with a contrast reversal stimulus. Experimental Brain Research, 35, 25–36.

Hamasaki, D. I., Tasaki, K., & Suzuki, H. (1979). Properties of X- and Y-cells in the rabbit retina. Japanese Journal of Physiology, 29, 445–457.

Hille, B. (2001). Ion channels of excitable membranes. Sunderland Massachusetts: Sinauer Associates.

Hirsch, J. A. (2003). Synaptic physiology and receptive field structure in the early visual pathway of the cat. Cerebral Cortex, 13, 63–69.

Hsueh H-A, Molnar A, & Werblin FS (2008). Amacrine to amacrine cell inhibition in the rabbit retina. J Neurophysiology.

Kaneko, A. (1970). Physiological and morphological identification of horizontal, bipolar and amacrine cells in goldfish retina. Journal of Physiology, 207, 623–633.

Katz, B., & Miledi, R. (1967). A study of synaptic transmission in the absence of nerve impulses. Journal of Physiology, 192, 407–436.

Lauritzen, T. Z., & Miller, K. D. (2003). Different roles for simple-cell and complex-cell inhibition in V1. Journal of Neuroscience, 23, 10201–10213.

Lee, T. H. (2004). The design of CMOS radio-frequency integrated circuits. Cambridge: Cambridge University Press.

Leibowitz, B. S., Boser, B. E., & Pister, K. S. J. (2005). A 256-Element CMOS Imaging Receiver for Free-Space Optical Communication. IEEE Journal of Solid-State Circuits, 40, 1948–1956.

Lukasiewicz, P. D., Lawrence, J. E., & Valentino, T. L. (1995). Desensitizing glutamate receptors shape excitatory synaptic inputs to tiger salamander retinal ganglion cells. Journal of Neuroscience, 15, 6189–6199.

MacNeil, M. A., Heussy, J. K., Dacheux, R. F., Raviola, E., & Masland, R. H. (1999). The shapes and numbers of amacrine cells: matching of photofilled with Golgi-stained cells in the rabbit retina and comparison with other mammalian species. Journal of Comparative Neurology, 413, 305–326.

MacNeil, M. A., Heussy, J. K., Dacheux, R. F., Raviola, E., & Masland, R. H. (2004). The population of bipolar cells in the rabbit retina. Journal of Comparative Neurology, 472, 73–86.

Magoon, R., Molnar, A., Hatcher, G., Zachan, J., & Rhee, W. (2002). A single-chip quad-band (850/900/1800/1900 MHz) direct-conversion GSM/GPRS RF transceiver with integrated VCOs and Fractional-N synthesizer. IEEE Journal of Solid-State Circuits, 37, 1710–1720.

Manookin, M. B., & Demb, J. B. (2006). Presynaptic mechanism for slow contrast adaptation in mammalian retinal ganglion cells. Neuron, 50, 453–464.

Manookin, M. B., Beaudoin, D. L., Ernst, Z. R., Flagg, L. J., & Demb, J. B. (2008). Disinhibition combines with excitation to extend the operating range of the OFF visual pathway in daylight. Journal of Neuroscience, 28, 4136–4130.

Molnar A & Werblin FS (2008). Inhibitory feedback shapes bipolar cell responses in the rabbit retina. J Neurophysiol.

Molnar, A., Magoon, R., Hatcher, G., Zachan, J., & Rhee, W., Damgaard, M., et al. (2002). A single-chip quad-band (850/900/1800/1900 MHz) direct-conversion GSM/GPRS RF transceiver with integrated VCOs and Fractional-N synthesizer. In: ISSCC (pp. 232–233). San Francisco: IEEE.

Razavi, B. (1998). RF Micerelectronics. Upper Saddle River: Prentice Hall.

Razavi B. (2000). Design of analog CMOS integrated circuits. McGraw-Hill.
Renteria, R. C., Tian, N., Cang, J., Nakanishi, S., Stryker, M. P., & Copenhagen, D. R. (2006). Intrinsic ON responses of the retinal OFF pathway are suppressed by the ON pathway. *Journal of Neuroscience, 26*, 11857–11869.

Rieke, F. (2000). Mechanisms of single-photon detection in rod photoreceptors. *Methods in Enzymology, 316*, 186–202.

Rockhill, R. L., Daly, F. J., MacNeil, M. A., Brown, S. P., & Masland, R. H. (2002). The diversity of ganglion cells in a mammalian retina. *Journal of Neuroscience, 22*, 3831–3843.

Roska, B., Molnar, A., & Werblin, F. (2006). Parallel Processing in Retinal Ganglion Cells: How Integration of Space-Time Patterns of Excitation and Inhibition Form the Spiking Output. *Journal of Neurophysiology, 95*, 3810–3822.

Slaughter, M. M., & Miller, R. F. (1981). 2-amino-4-phosphonobutyric acid: a new pharmacological tool for retina research. *Science, 211*, 182–185.

Steyaert, M., & Sansen, W. (1987). A micropower low-noise monolithic instrumentation amplifier for medical purposes. *Journal of Solid-State Circuits, 22*, 1163–1168.

Tran, M. N., Higgs, M. H., & Lukasiewicz, P. D. (1999). AMPA receptor kinetics limit retinal amacrine cell excitatory synaptic responses. *Visual Neuroscience, 16*, 835–842.

Wassle, H., Schafer-Trenkler, I., & Voigt, T. (1986). Analysis of a glycinergic inhibitory pathway in the cat retina. *Journal of Neuroscience, 6*, 594–604.

Wassle, H., Koulou, P., Brandstatter, J. H., Fletcher, E. L., & Becker, C. M. (1998). Glycine and GABA receptors in the mammalian retina. *Vision Research, 38*, 1411–1430.

Werblin, F. S., & Dowling, J. E. (1969). Organization of the retina of the mudpuppy, Necturus maculosus. II. Intracellular recording. *Journal of Neurophysiology, 32*, 339–355.

Zaghloul, K. A., Boahen, K., & Demb, J. B. (2003). Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *Journal of Neuroscience, 23*, 2645–2654.

Zaghloul, K. A., Boahen, K., & Demb, J. B. (2005). Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *Journal of Neuroscience, 25*, 860–868.