

OPEN

Probing and predicting ganglion cell responses to smooth electrical stimulation in healthy and blind mouse retina

Larissa Höfling^{1,2,3,4}, Jonathan Oesterle³, Philipp Berens^{3,5,6} & Günther Zeck^{1,5*}

Retinal implants are used to replace lost photoreceptors in blind patients suffering from retinopathies such as retinitis pigmentosa. Patients wearing implants regain some rudimentary visual function. However, it is severely limited compared to normal vision because non-physiological stimulation strategies fail to selectively activate different retinal pathways at sufficient spatial and temporal resolution. The development of improved stimulation strategies is rendered difficult by the large space of potential stimuli. Here we systematically explore a subspace of potential stimuli by electrically stimulating healthy and blind mouse retina in epiretinal configuration using smooth Gaussian white noise delivered by a high-density CMOS-based microelectrode array. We identify linear filters of retinal ganglion cells (RGCs) by fitting a linear-nonlinear-Poisson (LNP) model. Our stimulus evokes spatially and temporally confined spiking responses in RGC which are accurately predicted by the LNP model. Furthermore, we find diverse shapes of linear filters in the linear stage of the model, suggesting diverse preferred electrical stimuli of RGCs. The linear filter base identified by our approach could provide a starting point of a model-guided search for improved stimuli for retinal prosthetics.

Sensory neural prostheses attempt to replace nonfunctional elements of sensory organs and thereby to restore access to sensory information. For example, cochlear implants, which have been in clinical use for many years, replace the sensory receptors of the auditory system and can significantly improve auditory function and quality of life of implanted patients¹. Likewise, retinal implants aim to replace the function of the sensory receptors of the visual system, the photoreceptors, in retinopathies such as retinitis pigmentosa. While there are different approaches to retinal implants in terms of implantation site and technical realizations (subretinal²⁻⁴; epiretinal⁵; suprachoroidal⁶), all use short rectangular pulsatile stimuli to activate the remaining retinal neurons. Patients implanted with a retinal prosthesis experience benefits in mobility and navigation and some regain rudimentary but useful visual functions like the capability to identify, localize and discriminate objects^{2,7}.

Despite this success, there are a number of shortcomings in implant-aided vision, many of which are due to the fact that the electrical stimuli are very different from the physiological signals in the retina, which are much slower and smoother than the pulsatile electrical stimuli⁸⁻¹². Incidental activation of axons passing the stimulation target (e.g. a single retinal ganglion cell (RGC)) severely limits spatial resolution¹³⁻¹⁶, while desensitization of retinal neurons to repetitive electrical stimulation impedes high temporal resolution¹⁷⁻¹⁹. Acuity of prosthesis-mediated vision therefore reaches only a fraction of healthy visual acuity, leaving patients legally blind despite wearing a retinal implant²⁰. Another reason why bionic vision comes short of natural vision is that the image processing mechanisms of the intricate retinal network cannot be triggered selectively by retinal implants; pulsatile stimuli do not separately activate channels like the ON and the OFF pathway^{21,22}.

Different approaches have been suggested to tackle these problems. Placing the stimulation electrodes in different positions within the retina and varying the pulse durations confers control over which retinal elements are activated by pulsatile electrical stimulation²³. The activation of axons of passage can be avoided by increasing the

¹Natural and Medical Sciences Institute at the University of Tübingen, Reutlingen, Germany. ²Graduate Training Centre of Neuroscience, University of Tübingen, Tübingen, Germany. ³Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany. ⁴Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany. ⁵Bernstein Center for Computational Neuroscience, University of Tübingen, Tübingen, Germany. ⁶Department of Computer Science, University of Tübingen, Tübingen, Germany. *email: Guenther.Zeck@nmi.de

stimulus duration, thus improving spatial resolution of implant-aided vision¹⁴. Varying the shape of the stimulation waveform can increase the charge-efficiency²⁴ and the selectivity of the stimulus with respect to cell type and mode of activation^{23,25}. Cells in the different layers of the rabbit retina preferentially respond to increasing frequencies of sinusoidal stimulation as one moves from the outer nuclear to the ganglion cell layer²⁶. ON and OFF RGCs respond in different phases of low-frequency sinusoidal stimulation; however this phase-preference was assigned to activation of photoreceptors. The clinical relevance of sinusoidal stimulation thus appears to be limited²⁵. Varying stimulus parameters such as shape, repetition rate, duration and charge allowed to identify stimulus regimes that maximize the response ratio of ON and OFF RGCs^{27–31}.

However, the insight gained by such heuristic searches for optimal stimulus parameters is limited to a small area of the parameter space. One way to more exhaustively look for optimal stimuli for a given sensory modality is white noise analysis^{32–34}. This concept has been successfully applied to stimulation with the natural stimulus modality in different sensory systems, such as visual stimulation in the visual system³⁵, whisker deflection in the vibrissal system in rats³⁶, but also to artificial, electrical stimulation in the visual system^{33,37}. Recently, white noise composed of biphasic electrical current pulses with amplitudes drawn from a Gaussian distribution was used to map the spatiotemporal electrical receptive fields of rat retinal ganglion cells. Subsequently fit linear-nonlinear models accurately predicted RGC responses to electrical stimulation^{38,39}. Linear filters of mouse RGCs could also be recovered with a spatially uniform white noise stimulus consisting of normally distributed subthreshold voltage pulses^{40,41}. The results of these studies demonstrate the feasibility of the white noise analysis approach in electrical stimulation. However, these studies sampled subspaces of the stimulus space, varying pulse amplitude, but not duration or frequency. As white noise analysis can only identify the stimulus that is optimal within the space of provided stimuli, one should try to sample a space that is most likely to contain the true optimal stimulus.

Therefore, we developed and applied a smooth electrical Gaussian white noise current stimulus to cover a space of stimuli that more closely approximate the time scale of physiological signals in the retina. This smooth stimulus simultaneously probes the preferences of RGCs with respect to amplitude, polarity and frequency of an electrical stimulus. Using a high-density CMOS-based microelectrode array for stimulation, we were able to reliably activate RGCs in both wild-type mouse retina (*wt*) and retina from the *rd10* mouse model of retinal degeneration. We estimated linear filters of cells using two approaches to fitting a linear-nonlinear-Poisson (LNP) model: spike-triggered averaging (STA) and maximum likelihood estimation (MLE). Probing the light responses of RGCs in *wt* retina allowed us to relate electrical filter shapes to light response profiles. The LNP model accurately predicted RGC responses to electrical stimulation, demonstrating that it captures aspects of retinal processing of electrical stimuli relevant for response generation. The model may be useful for guiding the search for stimuli that improve spatial and temporal resolution of prosthetic-aided vision. The linear filters described here provide a starting point for this search.

Results

Simultaneous electrical stimulation and recording using a high-density CMOS-MEA. We used a smooth electrical current stimulus applied by a high-density CMOS-microelectrode-array (hdCMOS-MEA 5000) to stimulate *ex-vivo* flatmount preparations of wild-type and photoreceptor-degenerated retina in epi-retinal configuration. Our setup allowed us to simultaneously and continuously electrically stimulate on an arbitrary subset of the 1024 stimulation electrodes and record on 4225 recording electrodes (Fig. 1a,e,f). After the smooth stimulation waveform was removed from the recording, spike-sorting allowed to analyse the retinal ganglion cell responses to the stimulus at the level of individual cells (Fig. 1b–d). We evaluated the retinal response in wild-type retina ($n = 3$, *wt*) and blind retina ($n = 5$, *rd10*) to electrical stimulation by computing the temporal linear electrical filters of the RGCs in a model-based approach.

Reliable RGC responses to smooth electrical Gaussian white noise stimulation. Retinal ganglion cells in healthy and blind mouse retina responded reliably to stimulation with smooth electrical Gaussian white noise (Fig. 2b,e). For some cells, firing patterns were nearly identical across trials (Fig. 2b, fourth and fifth row (*wt*); and e, second row from below (*rd10*)); others were more variable in their response (Fig. 2b, fifth row from below (*wt*); and e, third row (*rd10*)). We quantified the reliability of the RGC response to the electrical stimulus by computing a *reliability index*, *RI* (see Methods). In *wt* retina, the majority of RGCs ($N = 53/84$, 63%) were entrained to the stimulus with a *RI* larger than a threshold of 0.15 (Fig. 2d). In *rd10* retina, a smaller percentage of cells were above threshold ($N = 26/126$, $RI > 0.15$, Fig. 2g); however, the levels of reliability among above-threshold cells were comparable between *rd10* and *wt* retina. A simulation of lateral and vertical current spread, taking into account different retinal thickness in *wt* and *rd10* retina, showed that the effect of this difference in thickness on stimulation intensity was negligible (see Fig. 1e,f).

One major challenge in bionic vision is the selective activation of different information channels such as the ON and the OFF channel^{42,43}. In order to be able to evaluate whether RGCs belonging to these retinal processing pathways are differently affected by our stimulus, we classified RGCs from healthy retina as ON, OFF or ON-OFF based on their light response. We presented light flashes (0.5 s OFF, 0.5 s ON; see Methods and Fig. 2a) and recorded the RGC responses to light stimulation. Of the cells that could be identified across recordings with light and electrical stimulation (47/84), 10 cells increased their spiking during incremental light stimuli (ON response, Fig. 2a, rows 1–3, row 6), 23 increased their activity during light decrements (OFF response, Fig. 2a, row 4, last five rows), and the remaining 14 cells had no clear preference for ON or OFF stimuli (ON-OFF or unknown, Fig. 2a, row 5). Only one of the 10 cells with ON response passed the reliability criterion of $RI > 0.15$ during electrical stimulation (see Fig. 2a,b, row 6), while 19/23 cells with OFF response and 8/14 cells with ON-OFF response responded reliably to electrical stimulation.

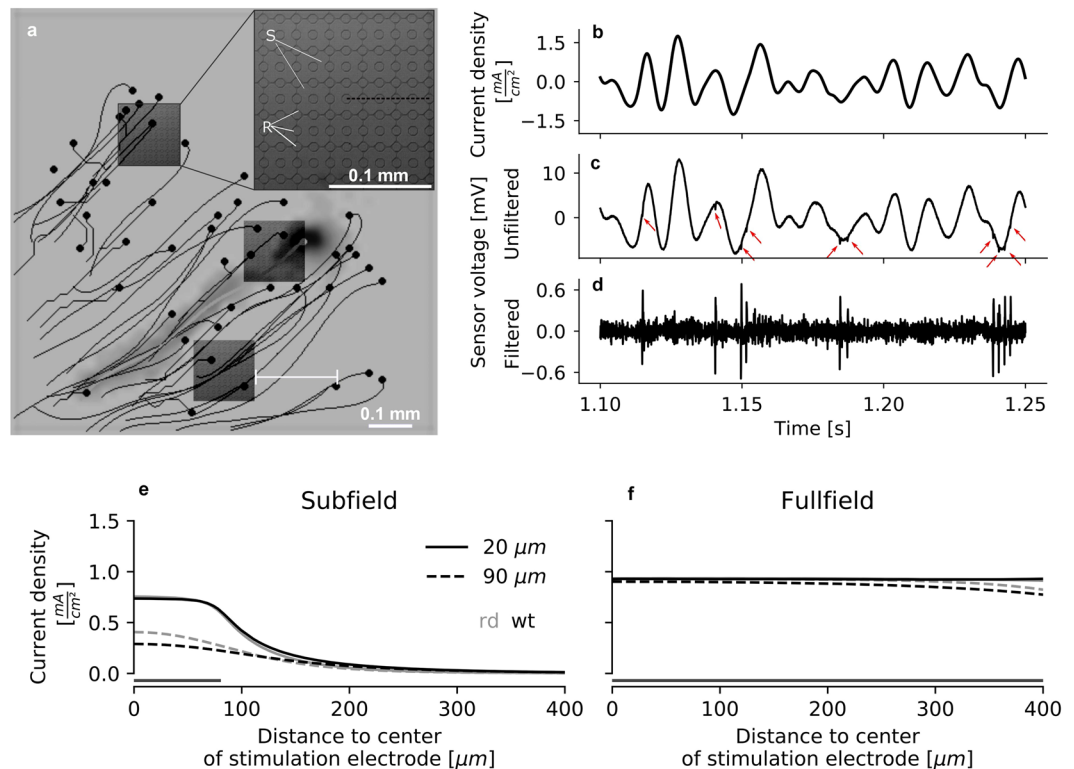


Figure 1. Simultaneous stimulation and recording of RGCs using a hd CMOS-MEA (a) Schematic of soma location (black dots) and axon traces (black lines) of RGCs as measured by the stimulation and recording software, and configuration of active stimulation areas (dark gray squares) in one example recording. The retina (here from *bl6* mouse) was flat-mounted on a hd CMOS-MEA (gray background). The horizontal white line illustrates how the distance of a cell's soma from the closest stimulation area was determined. The inset shows the grid of stimulation electrodes (large elements, labelled S) and recording electrodes (small elements, labelled R). In most recordings, only a subset of the stimulation electrodes were active (i.e. delivering the stimulation current). The black dashed line indicates the path from the center of the stimulation area to the edge along which current density was simulated (see panel (e)). Cell activity was recorded simultaneously on recording electrodes. (b) Expected current density of the smooth electrical Gaussian white noise stimulus, calculated as the derivative of the voltage command (see Methods, Eq. (1)). (c) Raw recording signal upon stimulation with the stimulus shown in (b). The stimulus causes an artefact in the raw recording orders of magnitude larger than the signal of interest, the spikes, indicated by red arrows. (d) Signal after filtering with a 2nd order band-pass Bessel filter between 1000 and 9500 Hz and artefact subtraction. The artefact is removed from the signal and spikes are clearly detectable. (e,f) Simulation of the current density at different heights above the stimulation electrodes in *wt* (black) and *rd10* (gray) retina in the subfield and fullfield condition, respectively (see Methods). The solid lines represent simulated current at a height of 20 μm , corresponding to the ganglion cell layer^{67,68}, and dashed lines represent simulated current at a height of 90 μm , approximately corresponding to the inner nuclear layer. The extent of the stimulation area is indicated by the line parallel to the x-axis. The distance of the path from the center to the edge was measured at a 0° angle (dashed black line in inset in panel (a)).

The reliability of the RGC response demonstrates that smooth electrical Gaussian white noise can efficiently activate RGCs in healthy and blind mouse retina. The different percentages of activated cells in *wt* and *rd10* retina can be explained by a larger spatial extent of the sensitivity to electrical stimulation in *wt* compared to *rd10* retina.

Localized response to local electrical stimulation. Previous studies have investigated the spatial extent of the *electrical receptive field* (spatial eRF), i.e. the area in which electrical stimulation evokes a response in a cell. In healthy rat retinas, eRF with diameters of about 250 μm have been reported, while the eRF in blind rat retinas was found to span about 200 μm in diameter⁴. Recent studies in healthy and blind mouse retina found eRF diameters of around 400 μm for RGCs in healthy mouse retina, and diameters of around 350 μm for RGCs in blind mouse retina^{44,45}. These eRF sizes were found upon network-mediated activation of RGCs in response to subretinal stimulation with pulsatile stimuli. A different study investigated how the activation threshold of rabbit RGCs depends on the distance between the RGC and the stimulation electrode upon epiretinal electrical stimulation⁴⁶. The threshold was found to increase with distance; this increase was much steeper for direct activation compared to network-mediated activation.

Here, we investigated the spatial extent of the sensitivity of RGCs to smooth electrical stimulation applied epiretinally in *wt* and *rd10* retina by stimulating subfields of the electrode array and analyzing how RGC responses

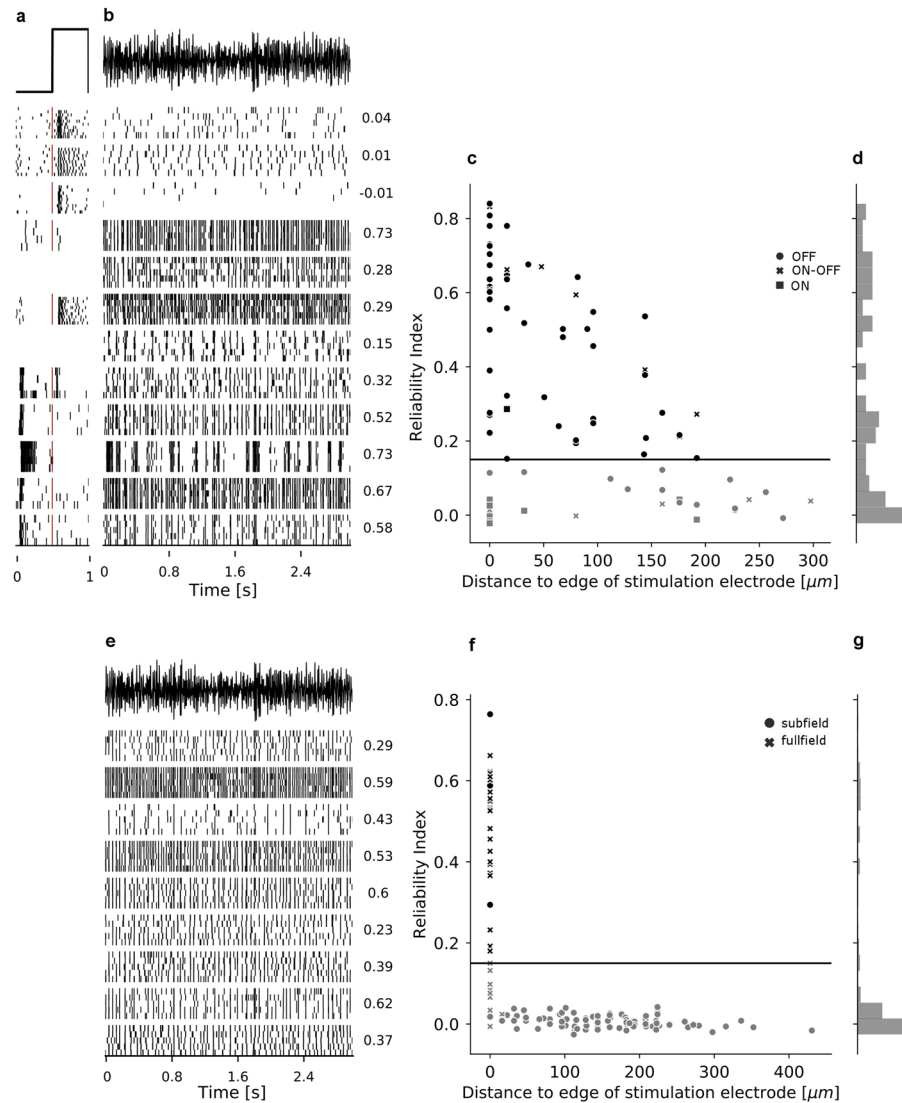


Figure 2. RGC responses to light flashes and smooth electrical stimulation. **(a)** Raster plots of the responses of RGCs from *wt* mouse retina to a fullfield light flash stimulus. The time course of the stimulus is indicated in the first row, and light onset is marked by a red vertical line in every raster plot. Not every cell was recorded both during electrical and light stimulation. **(b)** Raster plots of the responses of RGCs from *wt* mouse retina to an excerpt from the smooth electrical Gaussian white noise stimulus (stimulus shown in the first row). Numbers next to each row indicate the reliability index *RI* of the cell's response to electrical stimulation. **(c)** Reliability Index as a criterion for reliability of the response is plotted against distance to the edge of the closest active stimulation electrode. Black horizontal line indicates reliability threshold $RI > 0.15$ for inclusion in further analysis. Different markers indicate ON, ON-OFF and OFF cells, determined by their Bias Index as described in the Methods section. **(d)** Histogram of the reliability indices of *wt* RGCs. **(e) to (g)** same as **(b) to (d)** but for RGCs from *rd10* retina. Different markers in **(f)** indicate cells recorded in the subfield and fullfield condition (see Methods).

vary with distance from the stimulation areas. A simulation of the lateral current spread at different heights above the active stimulation electrodes shows that the stimulation in the subfield condition is indeed locally confined (Fig. 1e; see Methods). Simulated current density drops with distance from the center of the stimulation area; it does so more steeply at the level of the RGCs compared to the level of the INL. Simulated stimulation intensity is comparable between subfield and fullfield condition at the RGC level, whereas it is stronger in the fullfield condition at the INL level (Fig. 1e,f).

We found that the reliability of the RGC response to the electrical stimulation decreased with increasing distance from the closest active stimulation electrode (Fig. 2c,f). In *wt* retina, the majority of cells within a radius of $200 \mu\text{m}$ of an active stimulation electrode were entrained to the stimulus ($RI > 0.15$). Cells outside this radius did not respond robustly to the stimulus ($RI \leq 0.15$). This is largely in agreement with eRF sizes reported previously in studies using subretinal pulsatile stimulation. We hypothesized that the small fraction of ON cells responding to electrical stimulation was due to the fact that by chance, the recorded ON cells were located further away from

active stimulation electrodes. However, this was not the case; ON cells were recorded at distances from 0 to $\approx 200 \mu\text{m}$ from the closest active stimulation electrode (Fig. 2c).

In contrast, RGCs from *rd10* retina responded only in the fullfield condition (Fig. 2f, crosses), or if they were located directly adjacent to or above an active stimulation area in the subfield condition (Fig. 2f, dots; see Methods). Specifically, 23 of 35 (65.7%) RGCs from *rd10* retina responded reliably in the fullfield condition, while 3 of 4 cells located at zero distance from the stimulation electrodes in the subfield condition responded reliably to electrical stimulation. RGCs located further away from the stimulation electrodes did not respond to the stimulus (total number of cells in subfield condition: 91). Thus, the spatial extent of the effect of our stimulus is very localized in *rd10* retina and more spread out in *wt* retina. This difference in spatial extent suggests that different mechanisms underlie the RGC response to smooth electrical stimulation in healthy compared to blind mouse retina.

Temporal electrical linear filters of RGCs. We computed temporal electrical linear filters by fitting a linear-nonlinear-Poisson model using two different methods, spike-triggered averaging (STA) and maximum-likelihood-estimation (MLE). We did this for all cells which responded reliably to the stimulus (reliability index $RI > 0.15$, $N = 53/84$ in *wt* retina, $N = 26/126$ in *rd10* retina). While filters of RGCs from *rd10* retina were all monophasic negative (Fig. 3m,n), filter shapes of RGCs from *wt* retina were more variable (Fig. 3i,j).

To quantify the difference between the filter shapes found in wild-type retina, we performed a hierarchical clustering on the projection of the filters onto three principal components (PC). For STA filters, the first three PCs (3d–f) jointly explained 95% of the variance (PC1: 72.1%, PC2: 18.1%, PC3: 5.1%). For the MLE filters, the first five PCs were needed to explain 95% of the variance; we projected onto PC1, PC3 and PC4, explaining 69.1%, 9% and 3% variance, respectively (Fig. 4d,e,f), leaving out PC2. We did this because the variance explained by PC2 (12.4%) was mostly due to variance between different recordings, which was not of interest here (Fig. 4i–k).

We identified two distinct clusters for both STA and MLE filters based on the dendrogram generated by the hierarchical clustering algorithm (see Figs. 3g and 4g). These two clusters were separated along the axis of the first PC, and one of them could be split into two sub-clusters that were separated along the axis of the second (STA) or third (MLE) PC. The clusters corresponded to three filter shapes, monophasic negative (green, Figs. 3 and 4a), monophasic positive (red, Figs. 3 and 4b) and biphasic (violet, Figs. 3 and 4c). Assignment to clusters agreed between the STA and the MLE filter estimates for all except one cell. The estimates of the filters of the reliably responding example cells from *wt* retina (Fig. 2b rows 4–12) fall into these three categories (Fig. 3i,j). Filters from *rd10* retina were all monophasic negative (Fig. 3m,n, using the same example cells as in Fig. 2e). The acausal components observed in some filters are likely due to remaining correlations in the stimulus that could not be entirely removed by the whitening procedure (see Methods and Supplementary Fig. 1).

In order to compare the filters of RGCs from *wt* and *rd10* retina, we projected the LNP estimates of the filters from *rd10* RGCs (Fig. 3g, gray dots) onto the LNP PCs obtained from *wt* retina. The densely grouped projections in the lower part of the plot corresponds to projections of flat filters from non-responsive cells, while the more spread-out upper group corresponds to projections of filters from responsive cells (Fig. 3g, insets). Visual comparison of the filter shapes suggests that, while filters from *rd10* RGCs do not fall into one of the three clusters found in *wt*, they are more similar to monophasic negative and biphasic than to monophasic positive *wt* filters (Fig. 3i,j,m,n). The pattern of clustering described above confirms this observation.

We calculated the latencies of the (positive and/or negative) peaks of the STA and MLE filters of *wt* and *rd10* RGCs relative to the spike (see table 1). In *wt* retina, the latencies of the peaks of monophasic positive filters were around 2 ms longer than the latencies of the peaks of monophasic negative filters. For cells with biphasic filters, the spike occurred within 1 ms of the negative peak of the filter, preceded by 3 to 8 ms by the positive peak. For the monophasic negative filters of RGCs from *rd* retina, the spike occurred within less than 2 ms of the peak of the filter, comparable to the negative peak latencies observed for biphasic filters in *wt* retina.

Electrical linear filters vary with light response profiles of RGCs. Different RGC types, such as ON and OFF or sustained and transient types, constitute different information channels sending a preprocessed version of the visual scene to the brain^{43,47}. Preserving this first decomposition of the visual input into different information channels by separately activating RGCs of different types would significantly improve implant-aided vision⁴⁸. We therefore investigated whether there was a relationship between the light response profiles of RGCs from *wt* retina and their electrical linear filters. We characterized RGCs in terms of their preference for light polarity (ON, ON-OFF or OFF), and in terms of their response duration (transient, intermediate or sustained) by calculating a Bias Index and a Transiency Index (see Methods). We found transient and intermediate OFF and ON cells as well as sustained and intermediate ON-OFF cells (Fig. 5a). Transient OFF cells had biphasic or monophasic positive filters (Fig. 5f,h), while intermediate OFF cells and sustained ON-OFF cells had monophasic negative filters (Fig. 5g,d). Most ON cells did not respond reliably to the electrical stimulus ($RI < 0.15$) and could therefore not be assigned to a cluster based on their electrical linear filter. Electrical linear filters of most ON cells were rather noisy (Fig. 5b,e) due to their poor response to the stimulus; however, some ON cells had filters with a discernible negative deflection at spike time (Fig. 5c). The observed pattern suggests a difference in preferred electrical stimulus between RGCs with transient and sustained light responses.

LNP model accurately predicts RGC responses to electrical stimulation. We predicted the firing rates of RGCs from *wt* and *rd10* mouse retina using the STA and the MLE fit of the LNP model (see Methods, Fig. 6a,b (*wt*); Fig. 7a,b (*rd10*)). The models were fit on 80% of the data (4 s of the stimulus), and the parameters recovered from this fit were used for the prediction of the remaining 20% of the data (1 s of stimulus). For RGCs from *wt* retina, prediction performance P^{wt} , evaluated as the correlation between true and predicted firing rate (Methods, Eq. (24)) ranged from 0.05 to 0.67 for the STA fit and from 0.1 to 0.7 for the MLE fit of the model

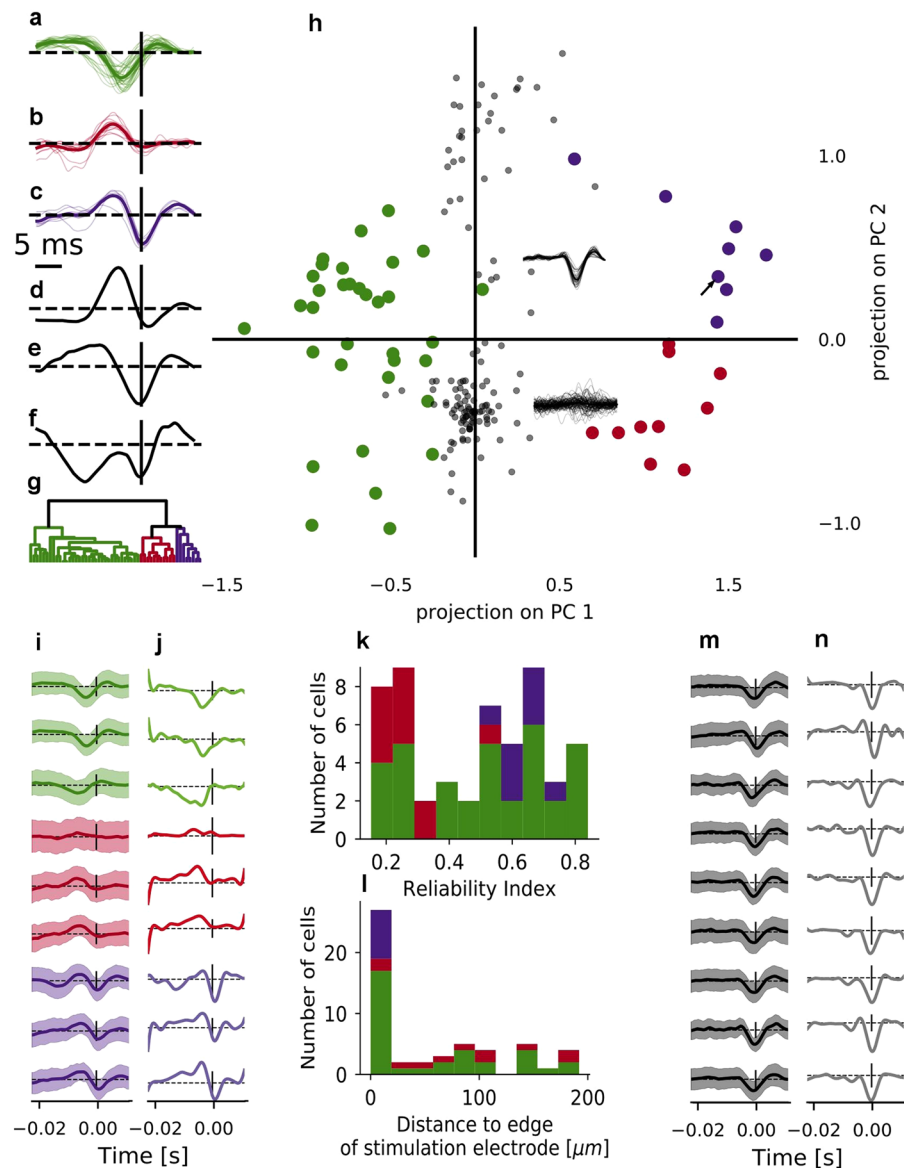


Figure 3. Hierarchical clustering of RGC electrical temporal filters estimated by STA. **(a–c)** Electrical temporal filters of all reliably responding RGCs from *wt* retina recovered from the STA fit of the LNP model, displayed separately for the three clusters identified by the hierarchical clustering algorithm. Thin lines are individual cell filters, thick lines indicate the average filter for one cluster. **(d,e,f)** 1st, 2nd and 3rd principal component (PC) recovered from principal component analysis of the ensemble of temporal filters from all cells. **(g)** Dendrogram showing the separation of consecutively joined clusters along the clustering metric (distance in euclidean space). **(h)** Scatter plot of the projections of the temporal filters onto the 1st and 2nd PCs (shown in **(d,e)**). Colors indicate cluster assignment. Grey dots indicate the projections of filters of RGCs from *rd10* retina, projected onto the same PCs. The lower inset shows the filters of all *rd10* RGCs whose filter projections are negative in PC2; the upper inset shows the filters of all *rd10* RGCs whose filter projections are positive in PC2. The black arrow marks the cell for which the assignment to clusters did not agree between STA and MLE estimate of the filters. **(i)** STA estimates of the filters of example RGCs from *wt* retina shown in Fig. 2b (rows 4–12, same order), with cluster assignment indicated by color. **(j)** Same as **(i)**, but for MLE estimates of the filters. **(k)** Stacked histogram of the distribution of reliability indices (RIs) for all cells with $RI > 0.15$, color-coded according to cluster assignment. **(l)** Distribution of distance to the edge of the closest active stimulation area of all responsive cells ($RI > 0.15$), color-coded according to cluster assignment. **(m)** STA estimates of the filters of example RGCs from *rd10* retina shown in Fig. 2e (same order). **(n)** Same as **(m)**, but for MLE estimates of the filters.

(Fig. 6c,d). The MLE fit performed slightly better than the STA fit (mean $\langle P \rangle_{STA}^{wt} = 0.46 \pm 0.15$, mean $\langle P \rangle_{MLE}^{wt} = 0.48 \pm 0.15$), and both models performed better for cells with monophasic negative and biphasic filter shapes compared to cells with monophasic positive filter shapes (Table 2 and Fig. 6e).

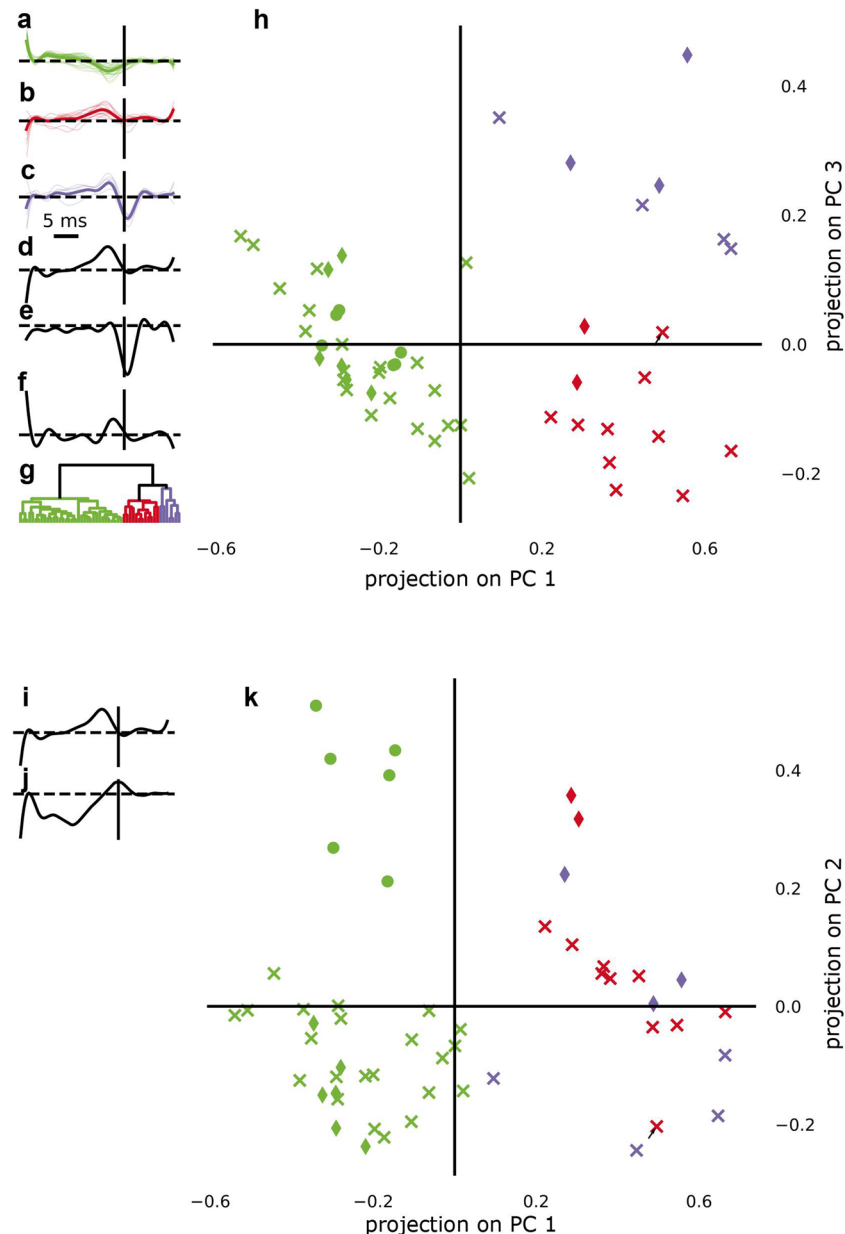


Figure 4. Hierarchical clustering of RGC electrical temporal filters estimated by MLE. (a–c) Electrical temporal filters of all RGCs from wild-type retina recovered from the MLE fit of the LNP model, displayed separately for the three clusters identified by the hierarchical clustering algorithm. Thin lines are individual cell filters, thick lines indicate the average filter for one cluster. (d,e,f) 1st, 3rd and 4th principal component (PC) recovered from principal component analysis of the ensemble of temporal filters from all cells. (g) Dendrogram showing the separation of consecutively joined clusters along the clustering metric (distance in euclidean space). (h) Scatter plot of the projections of the temporal filters onto the 1st and 3rd PCs (shown in (d,e)). Colors indicate cluster assignment. Different markers (filled circle, cross and diamond) indicate cells recorded in different sessions. The black arrow marks the cell for which the assignment to clusters did not agree between STA and MLE estimate of the filters. (i,j) 1st and 2nd PC recovered from principal component analysis of the ensemble of temporal filters from all cells. (k) Scatter plot of the projections of the temporal filters onto the 1st and 2nd PCs (shown in (i,j)). Colors indicate cluster assignment. Different markers (filled circle, cross and diamond) indicate cells recorded in different sessions. Note that the 2nd PC separates filter projections of cells from different recordings, but does not separate the red and the violet cluster.

We hypothesized that there was a positive relationship between the prediction performance of the LNP model and the reliability of a cell's response, indicated by its reliability index *RI*. Indeed, the performance of the model was generally better for more reliable cells (Fig. 6f). However, there were differences in model performance between cells with different filter shapes that could not be explained by differences in their response reliability; model performance was worse for cells with monophasic positive filters than for cells with monophasic negative

	latency of neg. peak [ms]	latency of pos. peak [ms]
monophasic neg. (green, N = 35)	-3.7, [-6.1, -1.3]	
	-3.3, [-5.4, 0.2]	
monophasic pos. (red, N = 11)		-5.5, [-6.2, -4.1]
		-4.7, [-6.6, -3.2]
biphasic (violet, N = 9)	0.2, [-0.1, 0.7]	-5.2, [-7.7, -4.5]
	-0.3, [-0.9, 0.2]	-3.1, [-7.3, -2.9]
<i>rd10</i> (N = 28)	-0.6, [-1.2, 0.3]	
	0.0, [-0.6, 0.8]	

Table 1. Peak latencies of filters The median and the range of the peak latencies relative to spike time for STA filters (top) and MLE filters (bottom). Negative values indicate that the peak occurred before the spike.

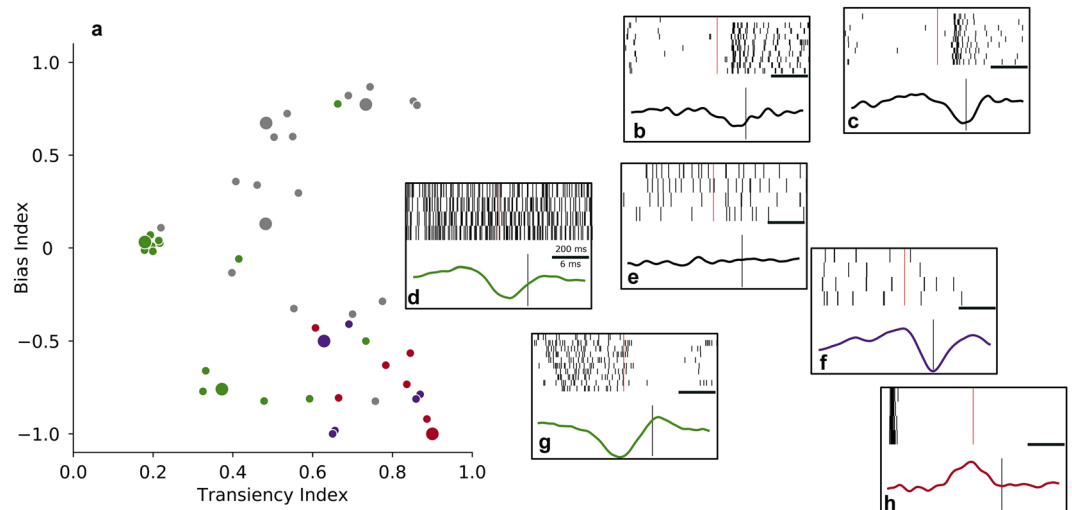


Figure 5. Electrical linear filters vary with light response profiles of RGCs. (a) Distribution of cells according to their light response profile. Color codes for cluster assignment of a cell based on its electrical linear filter. Grey dots indicate cells which were not included in the clustering due to below-threshold response reliability in the electrical stimulation condition ($RI < 0.15$). Bias Index quantifies the polarity of the light response from OFF (-1) through ON-OFF (around 0) to ON (1); Transiency Index quantifies the duration of the response to a cell's preferred light stimulus from sustained (0) to transient (1). (b-h) Example cells with different light response profiles, indicated in (a) by larger dots. Upper plot in each panel shows a raster plot of the cell's light response in 4–10 repetitions of the light step stimulus. Light onset is indicated by a vertical red line. Lower plot in each panel shows the cell's electrical linear filter as estimated by STA in the LNP model; color coding as in panel a. Vertical line indicates spike time. The timescale bar corresponds to 200 ms in the upper plot, and to 6 ms in the lower plot.

filters, even when the cells were equally reliable in their response. Thus, the difference in model performance must either be due to differences in the quality of fit of the linear stage, the nonlinear stage, or a combination of both. To disentangle the effects of the linear and the nonlinear stage on model performance, we compared the accuracy of only the linear stage to that of the full model in predicting the firing rate. The linear stage of the LNP model can be used to predict the firing rate by computing the dot product between the linear filter and the stimulus, without applying the static nonlinearity. The prediction performance of the linear stage was then determined in analogy to the full model (see Methods). Differences in performance between RGCs with monophasic negative and biphasic filters on the one hand and RGCs with monophasic positive filters on the other hand already emerged at the linear stage (Fig. 6g, x-dimension). The prediction performance further diverged when applying the nonlinearity (Fig. 6g, y-dimension); compared to the linear prediction, the nonlinear prediction was significantly better for RGCs with monophasic negative filters (Wilcoxon signed-rank test, $p = 3.67 \times 10^{-7}$, $\alpha = 0.05$) and for RGCs with biphasic filters (Wilcoxon signed-rank test, $p = 0.01$, $\alpha = 0.05$), but not for RGCs with monophasic positive filters (Wilcoxon signed-rank test, $p = 1$, $\alpha = 0.05$). Thus, differences in model performance between cells are a cumulative effect of differences in the reliability of the cells' response, but also of differences in the quality of the fit of the linear stage as well as the nonlinear stage.

Different filter shapes are the result of different response patterns that are elicited by the stimulus. Interestingly, both model fits also predict differences in response patterns between cells with different filter shapes (Fig. 6c,d, highlighted region). As noted before, linear filters of RGCs from *rd10* retina were most comparable to

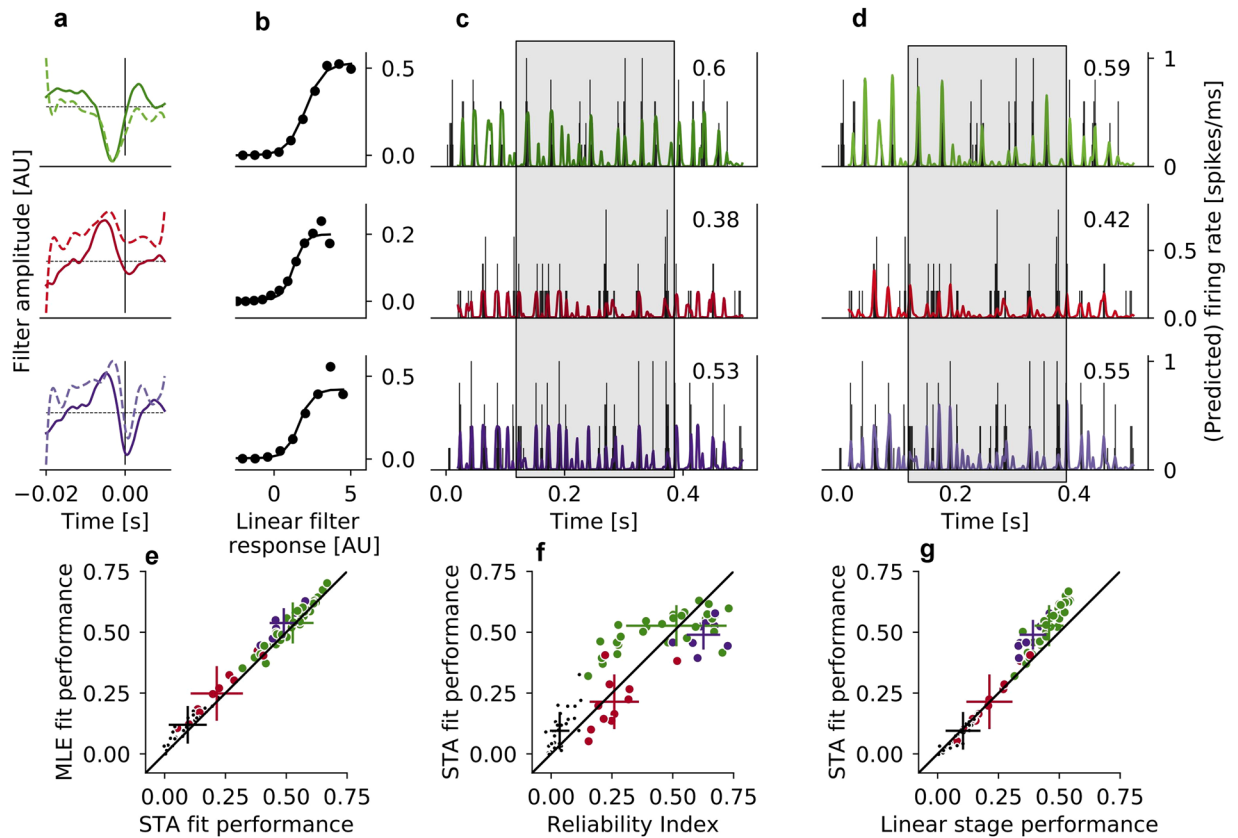


Figure 6. Model prediction of firing rates in *wt* retina. (a) Linear filters of three example cells from different clusters, derived from the STA fit (solid lines) and from the MLE fit (dashed lines); displayed at different scales in arbitrary units. To create the firing rate prediction, the filter response of the stimulus snippets with these filters was computed, and then the nonlinearity was applied. (b) Nonlinearity of the STA model; the nonlinearity of the MLE was the standard sigmoid function. (c) True firing rate, binned into 1 ms bins (black histogram) and the firing rate prediction of the STA fit of the model (colored traces) for three example cells; the highlighted region shows differential responses of cells with different types of filters, which are correctly predicted by the model. Decimal numbers in upper right corner indicate the performance of the model for that specific cell. (d) Same as (c) for the MLE fit of the model. (e) Prediction performance of the MLE fit plotted against the prediction performance of the STA fit; colors indicate cluster assignment. Small black dots represent RGCs that did not pass the reliability criterion of $RI > 0.15$, but are shown here for completeness. (f) Prediction performance of the STA fit plotted against reliability index. Colors indicate clusters assignment. Small black dots as in (e). (g) Prediction performance of the STA fit of the full LNP model plotted against the prediction performance of only the linear stage of the LNP model. Colors indicate cluster assignment. Small black dots as in (e).

monophasic negative and biphasic filters of RGCs from *wt* retina (Fig. 3h,i,m and Fig. 7a). Likewise, patterns of model performance were similar in *rd10* RGCs and *wt* RGCs with monophasic negative and biphasic filters. For RGCs from photoreceptor-degenerated retina, the performance of the STA fit of the model ranged from 0.35 to 0.66 ($\langle P \rangle_{STA}^{rd10} = 0.52 \pm 0.08$); again, the MLE fit performed slightly better (range 0.39 to 0.71, $\langle P \rangle_{MLE}^{rd10} = 0.58 \pm 0.09$, Fig. 7c–e). These performance values are very similar to those obtained for RGCs with monophasic negative filters from *wt* retina. Furthermore, the relationship between reliability of the response and LNP performance (Fig. 7f), as well as the relationship between linear stage performance and full model performance (Fig. 7g) in *rd10* retina were similar to these relationships for RGCs with monophasic negative and biphasic filters from *wt* retina (Fig. 6f,g).

Discussion

Retinal implants represent a promising treatment option for patients suffering from degenerative retinal diseases like retinitis pigmentosa or macular degeneration, but clinical efficacy remains limited due to non-physiological and therefore suboptimal stimulation strategies. The goal of this study was to investigate the response properties of retinal ganglion cells to a novel electrical stimulus that comes closer to the physiological signals in the retina. Identifying an electrical stimulus that is both physiologically plausible and technically feasible might significantly improve clinical outcomes in patients wearing retinal implants.

We electrically stimulated retinal ganglion cells in wild-type and photoreceptor-degenerated mouse retina with smooth Gaussian white noise currents and used a model-based approach to estimate linear filters of the

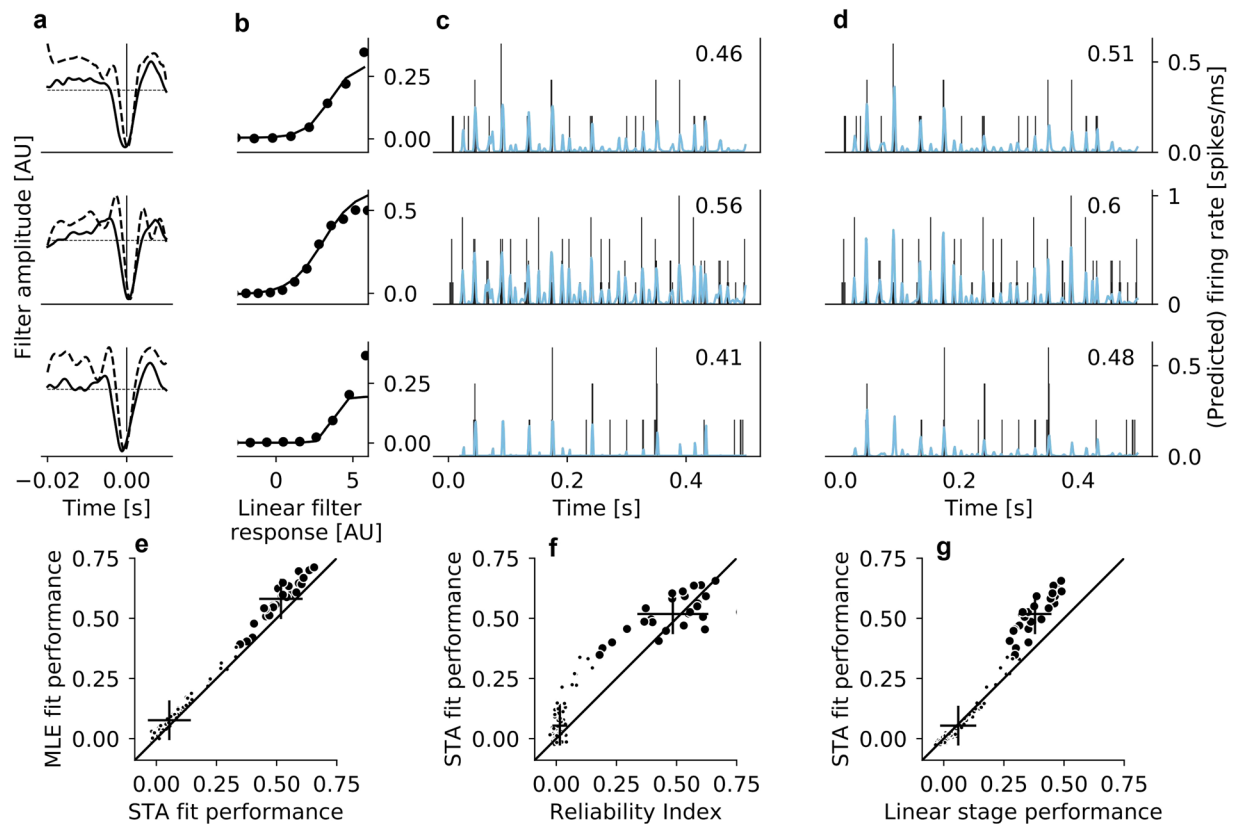


Figure 7. Model prediction of firing rates in *rd10* retina. (a) Linear filters of three example cells, derived from the STA fit (solid lines) and from the MLE fit (dashed lines); displayed at different scales in arbitrary units. To create the firing rate prediction, the filter response of the stimulus snippets with these filters was computed, and then the nonlinearity was applied. (b) Nonlinearity of the STA model; the nonlinearity of the MLE was the standard sigmoid function. (c) True firing rate, binned into 1 ms bins (black histogram) and the firing rate prediction of the STA fit of the model (blue traces) for three example cells. Decimal numbers in upper right corner indicate the performance of the models for that specific cell. (d) Same as (c) for the MLE fit of the model. (e) Prediction performance of the MLE fit plotted against the prediction performance of the STA fit. Small black dots represent RGCs that did not pass the reliability criterion of $RI > 0.15$, but are shown here for completeness. (f) Prediction performance of the STA fit plotted against reliability index. Small black dots as in (e). (g) Prediction performance of the STA fit of the full LNP model plotted against the prediction performance of only the linear stage of the LNP model. Small black dots as in (e).

	STA fit performance	MLE fit performance
monophasic negative (green, N = 34)	0.52 ± 0.09 , [0.32, 0.67]	0.54 ± 0.09 , [0.35, 0.70]
monophasic positive (red, N = 11)	0.21 ± 0.11 , [0.05, 0.41]	0.25 ± 0.11 , [0.10, 0.42]
biphasic (violet, N = 8)	0.49 ± 0.06 , [0.39, 0.58]	0.54 ± 0.06 , [0.45, 0.63]
<i>wt</i> all (N = 53)	0.46 ± 0.15 , [0.05, 0.67]	0.48 ± 0.15 , [0.10, 0.70]
<i>rd10</i> all (N = 26)	0.52 ± 0.08 , [0.35, 0.66]	0.58 ± 0.09 , [0.39, 0.71]

Table 2. Prediction performance of the LNP model (STA and MLE fit) Prediction performance (mean \pm s.d., [range]) of the STA fit and the MLE fit, measured as the correlation coefficient $C_{\hat{r}}$ (see Methods) between predicted and true firing rate at a resolution of 1 kHz, evaluated separately for the three identified clusters among *wt* filters, as well as for *rd10* filters.

RGCs. The estimates of the linear filters could be clustered into three different groups based on their shapes and constituted the linear stage of a linear-nonlinear-Poisson model, which accurately predicted retinal ganglion cell firing probability. We demonstrate that physiologically plausible electrical stimuli can be used to activate retinal ganglion cells in both wild-type and photoreceptor-degenerated retina; furthermore, the LNP model could be used to find stimuli that maximize the response in cells with one class of filters while minimizing the response in cells with a different class of filter.

System identification approaches to estimate preferred stimuli of neurons are always limited to finding the preferred stimulus within the subspace that is spanned by the presented stimuli; if the true preferred stimulus

is not in this subspace, it can only be approximated by the most closely matching stimulus within the subspace that was sampled^{35,49}. Our smooth electrical stimulus broadly, but certainly not exhaustively, samples stimulus space, simultaneously probing different stimulus dimensions like frequency, amplitude and polarity. However, the stimulus is low-pass filtered at 100 Hz, posing a limitation on our sampling of frequency space; we cannot exclude that the true preferred electrical stimuli of RGCs are of higher frequency. The linear filters shown here should therefore be interpreted as the preferred electrical stimuli of RGCs within the subspace sampled by our stimulus.

One of the main goals in retinal prosthetics is the selective activation of different information channels in the retina, most prominently the ON and OFF channels. One way of achieving this goal are cell-type specific preferred electrical stimuli, like it is the case for light stimulation of healthy RGCs^{43,47}. Indeed, investigating network-mediated electrical activation of rat RGC, ON cells were found to prefer anodic-first biphasic pulses, and OFF cells to prefer cathodic-first pulses, but the difference in preference was small³⁹. Antagonistic polarities of electrical filters in wild-type mouse ON and OFF cells have been reported as well⁴¹; however, the clinical applicability of these filters as stimuli is limited as they most likely arose due to photoreceptor activation⁴⁹.

In this study, we found three different classes of RGC filters in response to smooth electrical stimulation in wild-type, and one class of filter in photoreceptor-degenerated mouse retina. The most numerous of the three classes of RGC filters in *wt*, the monophasic negative filter, reflects RGC activation by cathodic currents. This is in agreement with previous findings showing that RGCs are preferentially activated by cathodic currents in epiretinal configuration, irrespective of whether the mode of activation is direct or through the network. Anodic currents also activate RGCs in epiretinal configuration, but at higher thresholds²³. Indeed, less RGCs had monophasic positive filters reflecting activation by anodic currents, and these cells also responded less reliably to electrical stimulation (see Figs. 3k and 6e–g). This suggests that at the given stimulation intensity, RGCs with a preference for anodic currents were activated just above threshold. Considering the reported preference of OFF cells for cathodic-first pulses³⁹, it is surprising that we find biphasic filters with opposite polarity (anodic-first) in OFF cells.

We investigated whether the different classes of electrical filters relate to light response profiles of RGCs in the healthy mouse retina. We found only few reliably responding ON cells, suggesting the conclusion that our stimulus more selectively activates OFF and ON-OFF cells. However, a large fraction of cells responding reliably to the electrical stimulus could not be identified in the light condition and could therefore not be characterized in terms of their light response profiles. As these reliably responding RGCs might be ON, ON-OFF or OFF, we can not conclude that our stimulus more selectively activates one or the other cell type. However, among the OFF cells, RGCs with sustained and transient light response preferred different shapes of electrical stimuli, suggesting that they could be selectively activated by stimuli derived from their respective filters.

While we found different classes of filters in *wt* retina, we found only one type of filter in *rd10* retina, which might cast doubt on the clinical relevance of our findings. Specifically, if our electrical stimulus activated RGCs in healthy retina via photoreceptors, and the different filter shapes were due to network mechanisms between photoreceptors and bipolar cells^{41,50}, selective activation of different pathways would not be achievable in blind retina. Therefore, identifying which retinal elements in the vertical pathway (photoreceptors, bipolar cells or RGCs) are activated by our stimulus is highly relevant. One way of determining the mode of activation is by using pharmacological blockers⁵¹. We did not perform experiments with pharmacological agents, but the temporal and spatial properties of the RGC responses, i.e. latency and spatial spread, can still inform us about the origin of the response. Because our stimulus is time-continuous, we cannot directly compute the response latency as the time difference between stimulus delivery and response. Instead, we used the latencies of the filter peaks relative to the spike time as a proxy for response latency. The filter peak latencies (see Table 1) are much shorter than the response latencies >20 ms usually reported for photoreceptor-mediated activation of ganglion cells in rabbit and rat retina. It is therefore unlikely that the RGC responses in *wt* retina were due to activation of photoreceptors. Rather, the latencies of both positive and negative monophasic filter peaks fall in the range usually reported for bipolar cell activation^{23,28,51}. Also, *wt* RGCs with (negative or positive) monophasic filters responded within a radius of ≈ 200 μm of stimulation electrodes, congruent with reports of the spatial extent of sensitivity to electrical stimulation in mouse retina upon network-mediated stimulation^{44–46,51}. Spatial integration over presynaptic inputs might lead to indirect activation of RGCs at distances of up to ≈ 200 μm from the stimulating electrodes. The comparably more shallow drop in current density outside of the active stimulation area at the level of the INL (relative to the level of the RGCs) (Fig. 1e) might facilitate this mode of activation at larger distances to the stimulating electrodes.

Conversely, the short latencies of the negative peak of biphasic filters in *wt* and monophasic filters in *rd10* retina suggest that these filters arose due to direct ganglion cell activation^{16,22}. Also, RGCs with this type of linear filter were only found at a cell-electrode distance of 0 μm (see Fig. 2f and 3l). A possible explanation for this finding is that these cells were activated directly and just above threshold, and that the stimulation intensity at the ganglion cell layer was too weak at larger cell-electrode distances to surpass this threshold (see Fig. 1e). One factor that could lead to a more focal activation in *rd10* retina compared to *wt* retina might be previously reported elevated activation thresholds in *rd10* retina⁵². However, recent studies found a large variability in direct activation thresholds in degenerated retina⁵³ and no clear difference in thresholds for direct and indirect activation between *wt* and *rd10* retina^{45,54}. Taken together, the temporal properties of biphasic filters in *wt* retina and filters in *rd10* retina suggest that these were the result of direct retinal ganglion cell activation, while both monophasic filter shapes found in *wt* retina likely reflect RGC responses due to bipolar cell activation. It is therefore conceivable that different retinal elements are activated in *wt* and *rd10* retina by smooth electrical stimulation at the given intensity: bipolar and ganglion cells in wild-type retina, and only ganglion cells in photoreceptor-degenerated retina.

If *wt* RGCs with monophasic positive and negative filters were indeed activated through the network as hypothesized above, then the short latencies of filter peaks in this study would indicate an improvement in temporal

precision of network-mediated electrical activation of RGCs by the use of smooth over pulsatile electrical stimulation. Epiretinal stimulation of wild-type mouse retina with subthreshold pulsatile Gaussian white noise yielded biphasic filters with peak latencies roughly 2 orders of magnitude longer than the filters described here^{40,41}. Filters estimated from network-mediated rat RGC responses to photovoltaic stimulation were about 10 times slower⁵⁰. Comparing to direct estimates (as opposed to the proxy of filter peak latencies) of network-mediated response latencies found using pulsatile stimulation, the peak latencies of filters reported here are roughly 2–5 times shorter and less variable^{39,55,56}.

Our stimulus could prove clinically relevant in several different ways. First, it could be used to directly activate individual RGCs in a spatially confined manner, avoiding axonal stimulation which causes the occurrence of elongated phosphenes experienced by patients wearing a retinal implant⁵⁷. In this case, however, differential activation of retinal pathways could probably not be achieved by using selective stimulus waveforms, but would have to occur through targeted stimulation of identified RGCs^{22,58,59}. Second, the short and precise latencies found here are a result of fast and precise RGC responses to electrical stimulation. Thus, a stimulus derived from the filters described here, implemented in a retinal prosthesis, might confer greater control over temporal activation patterns and thus help to overcome problems such as fading¹⁸.

Finally, our model predicts differences in electrically evoked responses of cells with different filter shapes (see Fig. 6c,d, highlighted region) and might therefore be used to identify stimuli that will maximize the response of cells with one type of filter while minimizing the response in cells with a different type of filter⁶⁰. Identifying such stimuli might potentially provide a tool for selective activation of certain cell types. However, different filter shapes were so far only identified in wild-type retina, likely upon network-mediated activation. One factor affecting the possibility to activate RGCs via the network is the degree of degeneration of the retina. In this study we report findings from *rd10* retina at different advanced stages of degeneration (mice aged 80 to 209 days). In earlier stages of degeneration, however, the network remains intact in the *rd10* model of retinal degeneration as well as in patients suffering from retinitis pigmentosa^{44,61–63}. Future studies should therefore investigate how the degree of degeneration affects RGC responses to our stimulus, and whether remaining intact circuitry can indirectly activate RGCs in degenerating retina upon smooth electrical stimulation.

In conclusion, future work will reveal to what extent the stimulus and modelling approach presented here will improve stimulation protocols for retinal prostheses in terms of spatial and temporal resolution and cell-type specificity of the stimulation.

Methods

All experimental procedures were carried out in compliance with §4 of the German law on animal protection and were approved by the Regierungspräsidium Tübingen (Registration No.: 35/9185.82-7). All the experiments were performed in accordance with the ARVO statement for the use of animals in ophthalmic and visual research.

Retina preparation. *Ex vivo* retina from five B6.CXB1-Pde6brd10/J (*rd10*) mice (2 females, 3 males; age between post-natal days 80 and 209 (p80 and p209) and from three C57BL/6J (*wt*) mice (2 female, 1 male; age between p87 and p274) was used. Animals were dark-adapted for approximately 30 minutes before the experiment, anaesthetised with carbon dioxide and euthanized by cervical dislocation. Both eyes were then enucleated, the retina was isolated and dissected in Ames' buffer under dim red-light conditions to prevent bleaching of remaining photoreceptors (further details have been described previously^{44,64}). A retinal portion of ca. 1–2 mm² was placed on a CMOS-based microelectrode array in epiretinal flat-mount configuration. Before each use, the microchip was cleaned with Terg-a-zyme (Sigma Aldrich, Z273287 dissolved in bidistilled water) and then coated with Poly-L-lysine (Sigma Aldrich, P2636, 1 mg/ml dissolved in bidistilled water). The recording chamber (2 ml) was perfused with warm, carbonated Ames' buffer (Sigma Aldrich, A1420, 36 °C, pH 7.4) at a flow rate of 4 $\frac{ml}{min}$.

High-density CMOS-based microelectrode array (hd CMOS-MEA). Stimulation and recording of retinal ganglion cell activity were performed with a high-density CMOS-based microelectrode array (CMOS MEA 5000, Multi Channel Systems MCS GmbH, Reutlingen, Germany). The MEA comprised 1024 capacitive stimulation electrodes and 4225 recording electrodes on an area of 1 mm² (see Fig. 1). Each stimulation electrode extended over an area of 688 μm^2 with 0.5 μm spacing between stimulation electrodes. The stimulation electrodes were made of a titanium nitride and were covered by a native oxide layer which insulated the metal from the electrolyte in the recording chamber. The pitch (center-to-center spacing) between two recording electrodes was 16 μm . The stimulation software (<https://www.multichannelsystems.com/products/cmos-mea-control>) allowed to define arbitrary stimulation areas by choosing a stimulation electrode to be active or inactive independently for each stimulation site (see Fig. 1a). The resolution of the Digital-to-Analog Converter was 16 bit. Details of the CMOS-based MEA have been reported previously^{65,66}.

Electrical stimulation. We designed an electrical stimulus consisting of Gaussian white noise low-pass filtered at 100 Hz with a 5th order Butterworth filter (Fig. 1b). The electrical stimulus was applied using the hd CMOS-MEA 5000 (see Fig. 1a). Applying voltages to the stimulation electrodes of the chip evokes capacitive currents in the electrolyte. The stimulation current density is proportional to the time derivative of the electrode voltage and scales with the specific electrode capacitance c :

$$i_{stim} = c \frac{dV}{dt} \quad (1)$$

Therefore, in order to obtain a certain current i_{stim} , the integral function of the desired current has to be applied to the chip as voltage command. For our electrical Gaussian white noise stimulus, this was achieved by generating $f_s \cdot T$ samples (where f_s is the sampling frequency and t is the total stimulus duration) from a standard normal distribution, and then calculating the cumulative sum over these samples. To reduce the drift in the random walk that is generated by this process, a reflection limit $L_{reflect}$ was introduced: whenever adding a new sample to the cumulative sum would raise the absolute value of the cumulative sum above the reflection limit $L_{reflect} = 10$, the step's sign was inverted. This yielded a sequence of values (in arbitrary units) describing a random walk, which was then low-pass filtered at 100 Hz with a 5th order Butterworth filter and rescaled to generate a voltage command $V_{command}$ that respects the safe input limits of the stimulus generator, $[0V \leq V_{command} \leq 2.5V]$. The introduction of the reflection limit reduces the power of the stimulus in the low frequency regime; apart from that, power spectral density of the resulting signal is relatively flat up to a frequency of 100 Hz and then drops off (Fig. 8b; note that frequencies < 4 Hz are not shown to improve visibility).

The stimulus applied in the experiments described here was generated using the parameters $f_s = 10$ kHz, $T = 5$ s and $L_{reflect} = 10$ (AU). The stimulation current was measured as described before^{45,55}. The maximal current density achieved with these parameters and the available chips was $i_{max} \approx 1.5 \frac{mA}{cm^2}$. Absolute current magnitude depended on the area of the active stimulation sites; for most recordings a subset of the stimulation electrodes was chosen (see Fig. 1a), while for one recording of *rd10* retina, the whole array was stimulated. These conditions are referred to as subfield and fullfield condition, respectively. Active electrodes in the subfield condition were chosen such as to cover areas with good contact between retina and CMOS-MEA as reflected by RGC activity detected on the recording electrodes. The same Gaussian white noise stimulus was played on all active electrodes throughout all conditions and experiments. We pool across conditions and set the distance of RGCs to stimulation electrode to 0 for all cells in the fullfield condition.

Simulation of current density. We simulated the current density in the retina with the finite-element method using the software COMSOL Multiphysics[®]. We modeled the retina as a cylinder with a radius of 1 mm and varied its height to simulate *wt* retina (height of 200 μm) and *rd10* retina (height of 100 μm). All other parameters were identical in the simulations for *wt* and *rd10* retina. A second cylinder with the same radius and a height of 1 mm was stacked on top to model the surrounding electrolyte. On the bottom of the retinal cylinder we centrally placed 5×5 or 32×32 electrodes to simulate subfield and fullfield stimulation conditions, respectively. The stimulation electrodes were modeled to be flat with shapes identical to the experimentally used electrodes (see Fig. 1a, inset). The comparably large and remote reference electrode was modeled by setting the top of the electrolyte cylinder to an electrical potential of zero. For the stimulation electrodes we assumed a homogeneous surface current density of $1.5 \frac{mA}{cm^2}$. All other boundaries were assumed to be perfect insulators with a normal current density of 0. The electrical resistivities of the retina and the electrolyte were assumed to be homogeneous and isotropic with values of $1000 \Omega \cdot cm$ and $65 \Omega \cdot cm$, respectively¹⁶. We modeled rectangular stimulation areas of different sizes (5×5 and 32×32) and computed the norm of the current density for different lateral distances to the center of the stimulation area between 0 and 400 μm in steps of 0.2 μm . Due to the shapes and placements of the electrodes on the chip, the current density at a fixed lateral distance varies with angle. Therefore, we also computed the current density for each lateral distance at angles between 0 and 90° in steps of 0.2°. The current density was then averaged across all angles. This was done for a vertical distance to the stimulation electrodes (height) of 20 μm (approximately ganglion cell layer^{67,68}) and 90 μm (approximately inner nuclear layer) (see Fig. 1e,f). For further details on the simulation, please see Oesterle *et al.*⁶⁹.

Light stimulation. To probe the light response of RGCs in wild-type retina, we presented full-field light flashes repeated 4 to 10 times. Light stimuli were generated by a selected LED (pE 4000 coolLED, peak wavelength 470–490 nm) and were flashed on for 500 ms, then turned off for 500 ms, thus creating a 1 Hz flicker. The light stimuli were projected onto a custom-made digital mirror display (μ -Matrix, Rapp Optoelectronic GmbH) mounted on an upright microscope (Olympus, BX51WI). The digital mirror display was focused onto the backside of the microscope objective.

Light response analysis. The light response profile of RGCs from *wt* retina was quantified by calculating a Bias Index (BI) and a Transiency Index (TI)⁴³. The Bias Index is calculated as

$$BI = \frac{R_{ON} - R_{OFF}}{R_{ON} + R_{OFF}} \quad (2)$$

where R_{ON} and R_{OFF} are the total number of spikes fired by a cell in the first 300 ms after light onset and light offset, respectively, summed across all repetitions. The Bias Index ranges from -1 for OFF cells through 0 for ON-OFF cells to 1 for ON cells. RGCs with $BI < -0.25$ are termed OFF cells, RGCs with $-0.25 \leq BI \leq 0.5$ are termed ON-OFF cells and RGCs with $BI > 0.5$ are termed ON cells. The thresholds were chosen based on the distribution of BIs in the RGC population recorded here. The Transiency Index is defined as

$$1 - \langle R_{pref} \rangle \quad (3)$$

where $\langle R_{pref} \rangle$ denotes the average normalized histogram (binsize 50 ms) in the 500 ms window following a cell's preferred light stimulus (light onset or light offset). TI ranges from 0 (sustained; maximum firing rate maintained throughout duration of preferred stimulus) to $1 - \frac{1}{\text{numberofbins}}$ (transient; only fires in the first bin after stimulus

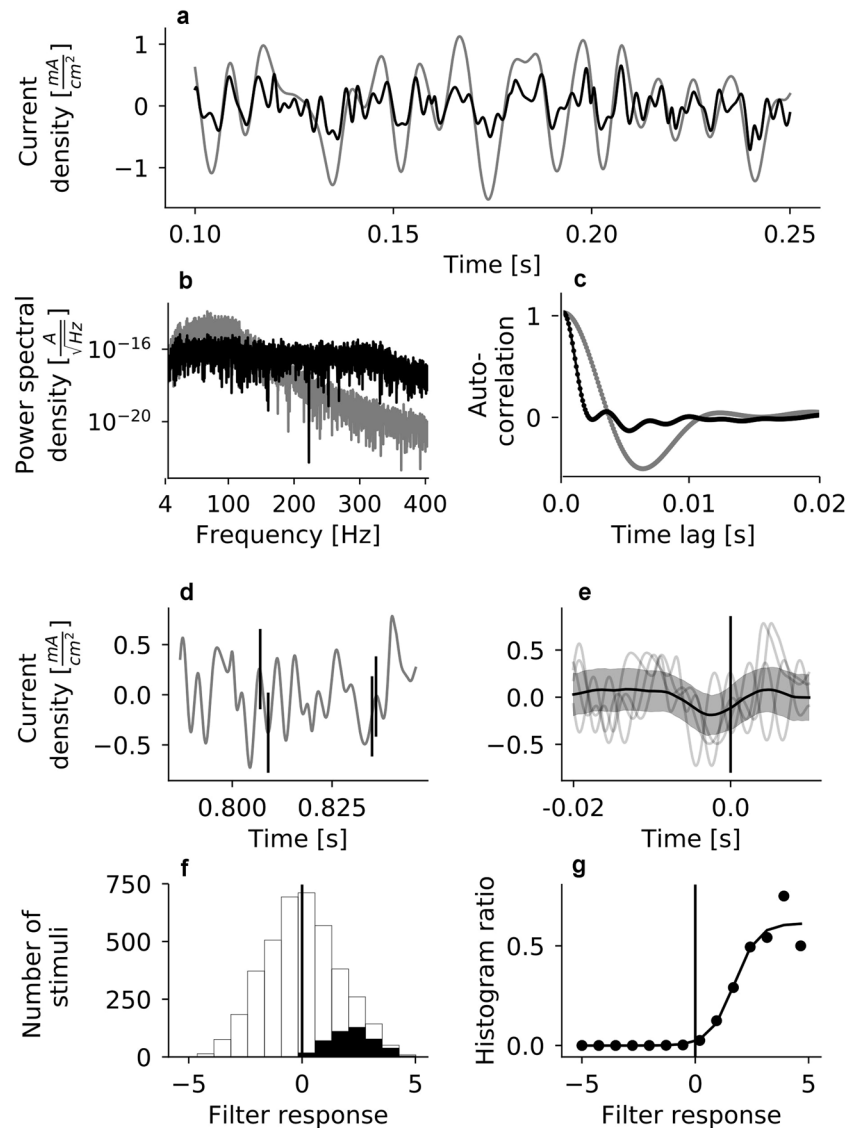


Figure 8. LNP model: Stimulus whitening and filter estimation by spike-triggered averaging. (a) Example segment of the original stimulus in grey, and the corresponding segment of the whitened stimulus in black. (b) Power spectral density (PSD) of the original (grey) and whitened (black) stimulus. While the PSD of the original stimulus declines at frequencies >100 Hz due to filtering, this effect is undone by whitening. (c) The autocorrelation function of the original (grey) and the whitened (black) stimulus at time lags of up to 20 ms. The temporal extent of the autocorrelation of the whitened stimulus is much reduced compared to the original. (d) For every spike of a RGC (black vertical lines), a segment of the whitened version of the stimulus, spanning 20 ms before and 10 ms after the spike, is added to the spike-triggered ensemble (STE). (e) The spike-triggered average (black trace) is computed by averaging across all elements of the STE (the STE elements from panel (a) are shown as thin gray lines). Gray shaded area indicates ± 1 standard deviation of the STE. Vertical black line indicates time of spike. (f) The elements of the full stimulus ensemble, consisting of all 300 sample long stimulus snippets taken 10 samples apart, are projected onto the linear filter of a cell and binned to yield a histogram (empty bars). The same is done for the elements of the STE (black histogram). (g) The nonlinearity (black dots) is estimated as the ratio between the histogram of the projected elements of the spike-triggered ensemble to the histogram of the projected elements of the full stimulus ensemble (black and the empty histograms from panel (f), respectively) and fit by a sigmoidal (Eq. (14)) or exponential (Eq. (15)) function (black trace, here sigmoidal).

onset; here $1 - \frac{1}{10} = 0.9$). For ON-OFF cells, the TI is defined as the mean between the TI for the OFF phase of the stimulus and the TI for the ON phase of the stimulus.

Artefact reduction and preprocessing. In the raw recording, spiking activity is masked by large artefacts caused by the stimulation current, making spike sorting and further analysis impossible (Fig. 1b,c). Therefore, a

series of data preprocessing steps was performed to remove the stimulation artefact. First, the raw traces of each recording electrode were band-pass filtered between 1000 and 9500 Hz with a 2^{nd} order Bessel filter. Then, because the time courses of the stimulation current and the artefact recorded on the sensors are aligned⁷⁰, the contribution of the stimulation artefact to the raw trace could be estimated by computing the dot product between the applied stimulus S and the raw trace r_{raw} (r_{raw} , S are N -dimensional column vectors), normalized to the dot product of the stimulation waveform with itself:

$$f = \frac{r_{raw}^T S}{S^T S}. \quad (4)$$

To remove the contribution of the stimulation artefact in the raw recording, the stimulus trace (Fig. 1b) was then subtracted from the raw recording (Fig. 1c), multiplied by the factor f :

$$r_{pre} = r_{raw} - f \cdot S. \quad (5)$$

In the resulting signal, spikes are clearly separated from the background signal and no longer masked by the stimulation artefact (Fig. 1d). Spike sorting was performed on this signal using the spike sorter based on convolutive independent component analysis⁷¹, implemented in the MultiChannel Systems CMOS-MEA-Tools Software (<https://www.multichannelsystems.com/software/cmos-mea-tools>, version 2.1.0).

Evaluating reliability of RGC responses to electrical stimulation. In order to quantify how reliably RGCs respond across repetitions of the same stimulus, we computed a *Reliability Index (RI)*. To compute the *RI*, individual RGC responses from 4 repetitions were binned into 2 ms wide bins, resulting in a histogram H . The same was done for the remaining 5th repetition, yielding a histogram h . Then the correlation coefficient between the two histograms was computed. This was repeated for all possible combinations of H and h , and the *RI* was computed as the average across the resulting correlation coefficients:

$$RI = \left\langle \frac{C_{Hh}}{\sqrt{C_{HH}C_{hh}}} \right\rangle, \quad (6)$$

where C denotes the covariance matrix of the histogram h of cell responses during four repetitions of the stimulus and the histogram h of the cell responses during the fifth repetition, and angular brackets $\langle \rangle$ denote the mean across all (five) possible combinations of H and h . A value of 1 indicates perfect reliability (identical response pattern in all repetitions), while a value close to 0 indicates no reliability in the response to the stimulus. Based on visual inspection of the distribution of the RIs of wild-type RGCs, we included RGCs in the analysis if their *RI* was larger than 0.15 (see Fig. 2d). This threshold was compatible with the distribution of RIs of RGCs from *rd10* retina (see Fig. 2g).

Fitting a linear-nonlinear-Poisson model to electrically evoked RGC responses. We modeled RGC responses to electrical stimulation with a Linear-nonlinear-Poisson (LNP) model³⁵. The LNP model assumes that the firing rate of a neuron in response to a stimulus s_t at time t can be modelled as an inhomogeneous Poisson process with instantaneous firing rate

$$\mu_t = \mathcal{N}(\mathbf{k}^T s_t) \quad (7)$$

where $\mathbf{k}^T s_t$ denotes the dot product of the cells linear filter and the stimulus at time t , and \mathcal{N} denotes a static nonlinearity.

We used two different approaches to fit the LNP, spike-triggered averaging with subsequent estimation of the nonlinearity, and maximum-likelihood estimation of a generalized linear model (GLM) with logit link function.

LNP model: Filter estimation by spike-triggered averaging. The spike-triggered average^{34,72} (STA) is calculated by averaging over the spike-triggered stimulus ensemble (STE), which is the collection of all stimuli that were followed by a spike within a defined time window (here 1 ms). For a discrete time shift τ_i before the spike the STA is given as

$$v(\tau_i) = \frac{1}{n} \sum_{k=1}^n s(t_k - \tau_i), \quad (8)$$

with

N the number of spikes, $s(t)$ the value of the stimulus at time t , τ_i the time shift relative to the spike.

A cell's linear temporal filter is then defined as the vector

$$\mathbf{v} = (v(\tau_{max}), v(\tau_{n-1}), \dots, v(\tau_0)) \quad (9)$$

The spike-triggered average was calculated 20 ms into the past and 10 ms into the “future” relative to the spike, i.e. $\tau_{max} = -20$ ms and $\tau_0 = 10$ ms. The duration of the filters presented here is thus 30 ms, equivalent to 300 stimulus samples.

The stimulus was low-pass filtered at 100 Hz (see Methods) and therefore has a non-zero autocorrelation at time shifts ≤ 10 ms (see Fig. 8c). These autocorrelations affect the shape of the filters computed by spike-triggered averaging⁷³. To reduce the effect of the autocorrelations on the filters, the STA was calculated using whitened stimulus “snippets”. For the whitening procedure⁷⁴, the 50 000 samples (5 s) long continuous stimulus was sliced into snippets of 300 stimulus samples, taken 10 samples apart (1 ms). This yielded a stimulus ensemble containing 4701 stimulus snippets, and the covariance matrix was computed based on this ensemble. The square-root of the pseudo-inverse of the resulting covariance matrix, $C_{stim}^{-\frac{1}{2}}$, was computed, and each stimulus snippet was whitened by multiplying with $C_{stim}^{-\frac{1}{2}}$:

$$\mathbf{s}_i^{white} = C_{stim}^{-\frac{1}{2}} \mathbf{s}_i \quad (10)$$

with $\mathbf{s}_i \in \mathbb{R}^{300 \times 1}$ stimulus “snippets” of 300 frames each $C_{stim}^{-\frac{1}{2}} \in \mathbb{R}^{300 \times 300}$ the pseudoinverse of the stimulus covariance matrix C_{stim} .

The pseudoinverse is obtained from the eigendecomposition of the stimulus covariance matrix

$$C_{stim} = \sum_{i=1}^{K \leq M} \lambda_i \mathbf{e}_i \mathbf{e}_i^\top \quad (11)$$

by keeping the eigenvectors $\mathbf{e}_i \in \mathbb{R}^{300 \times 1}$, but inverting and taking the square root of the $L < K$ largest eigenvalues λ_i and setting the remaining eigenvalues to zero:

$$C_{stim}^{-\frac{1}{2}} = \sum_{i=1}^L \lambda_i^{-\frac{1}{2}} \mathbf{e}_i \mathbf{e}_i^\top \quad (12)$$

The whitened stimulus used for spike-triggered averaging (Fig. 8a,d) was reconstructed by concatenating the resulting whitened stimulus snippets and smoothing the edges with a hanning window.

The nonlinearity \mathcal{F} was estimated as the ratio between the histograms of the projection of the STE onto the cells linear filter and the projection of the raw stimulus ensemble onto the cells linear filter³⁵ (Fig. 8f,g):

$$\hat{\mathcal{F}} = \frac{H_{STE}}{H_{FE}} \quad (13)$$

H_{STE} denotes the histogram of projected stimulus snippets from the spike-triggered ensemble (STE) and H_{FE} denotes the histogram of projected stimulus snippets from the full ensemble (FE). The estimate of the nonlinearity was then fit by a sigmoidal or exponential function. The sigmoidal fit to the nonlinearity at time t is described by the equation

$$\hat{\mathcal{F}}_{sig}(\mathbf{v}^\top \mathbf{s}_t) = \frac{y_{max}}{1 + \exp(-g(\mathbf{v}^\top \mathbf{s}_t - x_0))} + y_{min} \quad (14)$$

with free parameters y_{max} (saturation amplitude of the sigmoidal curve), g (gain factor), x_0 (50% threshold of the saturation amplitude) and y_{min} (vertical offset). The exponential fit to the nonlinearity at time t is described by the equation

$$\hat{\mathcal{F}}_{exp}(\mathbf{v}^\top \mathbf{s}_t) = a \exp(g(\mathbf{v}^\top \mathbf{s}_t)) + y_{min} \quad (15)$$

with free parameters a (controlling the height of the curve), g (gain factor), and y_{min} (vertical offset). \mathbf{s}_t is a whitened stimulus snippet as defined in Eq. (10), containing the 300 stimulus samples preceding t .

LNP model: Filter estimation by maximum-likelihood estimation. In addition to the spike-triggered approach, the temporal electrical filter was also estimated by fitting a GLM with logit link function and an elastic net penalty to the data^{75,76}. Here, the filter coefficients are found by minimizing the negative log-likelihood of the observed data with respect to the coefficients. To constrain the coefficients and to prevent overfitting, a penalty term is added to the loss function. The GLM was fit on the raw (not whitened) stimulus, because the process of maximum likelihood estimation intrinsically performs correlation-correction on the filter estimate⁷⁷. We used the `scikit-learn` python package⁷⁸ for fitting the GLM and for cross-validation.

The firing probability μ_t at time t is assumed to be related to the stimulus by a sigmoid function

$$\mu_t = \frac{1}{1 + \exp(-z_t)} \quad (16)$$

where

$$z_t = \beta_0 + \sum_j \beta_j s_{tj} \quad (17)$$

with $\beta_0 \in \mathbb{R}$ a scalar offset $\beta \in \mathbb{R}^m$ the estimate of the m -dimensional linear filter is a linear transformation of the stimulus $\mathbf{s}_t \in \mathbb{R}^{300 \times 1}$ at time t . Assuming that a spike train $Y = \{y_1, y_2, \dots, y_t\}$, $y_t \in \{0, 1\}$ of one particular neuron is generated by a sequence of independent (but not identically distributed) Bernoulli trials with time-dependent firing probability μ_t , the probability of the spike train is given by

$$P(Y|\mu_1, \mu_2, \dots, \mu_t) = \prod_t \mu_t^{y_t} (1 - \mu_t)^{(1-y_t)} \quad (18)$$

The average log-likelihood function $\mathcal{L}(\boldsymbol{\mu}; Y)$ of the parameters $\boldsymbol{\mu}$ given the observed spike train Y is obtained by taking the logarithm of equation (18) and dividing by the number of spikes N :

$$\mathcal{L}(\boldsymbol{\mu}; Y) = \frac{1}{N} \sum_t y_t \log(\mu_t) + (1 - y_t) \log(1 - \mu_t) \quad (19)$$

Adding the elastic net penalty,

$$\lambda[(1 - \alpha)P_2 + \alpha P_1] \quad (20)$$

with

$$P_1 = \sum_j \beta_j \quad (21)$$

$$P_2 = \frac{1}{2} \sum_j \beta_j^2 \quad (22)$$

to the average negative log-likelihood function (Eq. (19)) yields the loss function

$$\mathcal{J} = -\frac{1}{N} \sum_t y_t \log(\mu_t) + (1 - y_t) \log(1 - \mu_t) + \lambda \left[(1 - \alpha) \frac{1}{2} P_2 + \alpha P_1 \right] \quad (23)$$

The values of the parameters β_0 and β which minimize the loss function were found by using the SAGA solver implemented in the `scikit-learn` package. The resulting β represents the linear filter of the LNP model estimated by maximum likelihood estimation (MLE). The hyperparameters λ (determining the strength of the regularization) and $\alpha \in [0, 1]$ (determining the relative contributions of the L1 (Eq. (21)) and L2 (Eq. (22)) regularization terms) were determined by a cross-validated parameter grid search, choosing λ from 5 logarithmically spaced values from the interval $[0.0001, 0.01]$, and α from $\{0, 0.01, 0.1, 0.5, 1\}$.

Model prediction and evaluation. Both the STA fit and the MLE fit of the LNP were used to predict RGC spiking activity in response to the stimulus: the filter response of the stimulus was computed, and the nonlinearity was applied to the filter response, yielding the firing rate prediction. To prevent overfitting and overestimation of the model prediction performance, the model was fit on four seconds of stimulus, and the parameters derived from that fit were used to predict the firing rate in the remaining leftout second of the stimulus. This was done for all five possible splits into four training seconds and one test second.

Prediction performance P was evaluated as the correlation coefficient between true (r) and predicted firing rate (\hat{r}) at a resolution of 1 kHz:

$$P = \frac{C_{r\hat{r}}}{\sigma_r \sigma_{\hat{r}}} \quad (24)$$

where C denotes the covariance matrix of the true and the estimated firing rates r and \hat{r} .

Hierarchical clustering. A hierarchical clustering algorithm was used to cluster the filters of all cells to determine whether there were different classes of filters. At first, the dimensionality of the filters was reduced from $N = 300$ features (30 ms at a sampling rate of 10 kHz) to $k < N$ features retaining 90% of the variance by applying principal component analysis (PCA) and projecting the filters onto the retained k principal components (PCs). Hierarchical clustering was then performed on these projections, using the `scipy.cluster.hierarchy` package with the average euclidean distance in PC space as clustering metric. The number of clusters was determined by visual inspection of the resulting dendrogram.

Data availability

The datasets generated and analysed during the current study, as well as the code used for analysis, are available from the corresponding author on reasonable request.

Received: 4 March 2019; Accepted: 19 February 2020;

Published online: 23 March 2020

References

- Peterson, N. R., Pisoni, D. B. & Miyamoto, R. T. Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *Restorative Neurology and Neuroscience* **28**, 237–250 (2010).
- Zrenner, E. *et al.* Subretinal electronic chips allow blind patients to read letters and combine them to words. *Proceedings of the Royal Society* **278**, 1489–97, <https://doi.org/10.1098/rspb.2010.1747> (2011).
- Wang, L. *et al.* Photovoltaic retinal prosthesis: Implant fabrication and performance. *Journal of Neural Engineering* **9** (2012).
- Lorach, H. *et al.* Photovoltaic restoration of sight with high visual acuity. *Nature Medicine* **21**, 476–482 (2015).
- Humayun, M. S. *et al.* Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Research* **43**, 2573–2581 (2003).
- Ayton, L. N. *et al.* First-in-human trial of a novel suprachoroidal retinal prosthesis. *PLoS ONE* **9**, 1–26 (2014).
- Stingl, K. *et al.* Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS. *Proc. R. Soc. B* **280**, 20130077 <http://www.ncbi.nlm.nih.gov/pubmed/23427175> (2013).
- Matsui, K., Hosoi, N. & Tachibana, M. Excitatory synaptic transmission in the inner retina: paired recordings of bipolar cells and neurons of the ganglion cell layer. *Journal of Neuroscience* **18**, 4500–4510 <http://www.ncbi.nlm.nih.gov/pubmed/9614227> (1998).
- Diamond, J. S. & Copenhagen, D. R. The contribution of NMDA and Non-NMDA receptors to the light-evoked input-output characteristics of retinal ganglion cells. *Neuron* **11**, 725–738 <https://www.sciencedirect.com/science/article/pii/0896627393900823?via> (1993).
- Field, G. D. & Rieke, F. Nonlinear Signal Transfer from Mouse Rods to Bipolar Cells and Implications for Visual Sensitivity. *Neuron* **34**, 773–785 <https://www.sciencedirect.com/science/article/pii/S0896627302007006> (2002).
- Euler, T. & Masland, R. H. Light-evoked responses of bipolar cells in a mammalian retina. *Journal of Neurophysiology* **83**, 1817–1829 (2000).
- Euler, T., Haverkamp, S., Schubert, T. & Baden, T. Retinal bipolar cells: elementary building blocks of vision. *Nature Reviews Neuroscience* **15**, 507–519, <https://doi.org/10.1038/nrn3783> (2014).
- Grosberg, L. E. *et al.* Activation of ganglion cells and axon bundles using epiretinal electrical stimulation. *Journal of Neurophysiology* **118**, 1457–1471, <https://doi.org/10.1152/jn.00750.2016> (2017).
- Weitz, A. C. *et al.* Improving the spatial resolution of epiretinal implants by increasing stimulus pulse duration. *Science Translational Medicine* **7**, 1–12 (2015).
- Nanduri, D. *et al.* Frequency and amplitude modulation have different effects on the percepts elicited by retinal stimulation. *Investigative Ophthalmology and Visual Science* **53**, 205–214 (2012).
- Eickenscheidt, M. & Zeck, G. Action potentials in retinal ganglion cells are initiated at the site of maximal curvature of the extracellular potential. *Journal of Neural Engineering* **11** (2014).
- Pérez Fornos, A. *et al.* Temporal properties of visual perception on electrical stimulation of the retina. *Investigative Ophthalmology and Visual Science* **53**, 2720–31 <http://www.ncbi.nlm.nih.gov/pubmed/22447863> (2012).
- Stingl, K. *et al.* Subretinal Visual Implant Alpha IMS - Clinical trial interim report. *Vision Research* **111**, 149–160, <https://doi.org/10.1016/j.visres.2015.03.001> (2015).
- Jensen, R. J. & Rizzo, J. F. III. Responses of ganglion cells to repetitive electrical stimulation of the retina. *Journal of Neural Engineering* **4**, 1–6 <http://stacks.iop.org/1741-2552/4/i=1/a=S01> (2007).
- Weiland, J. D., Walston, S. T. & Humayun, M. S. Electrical Stimulation of the Retina to Produce Artificial Vision. *Annual Review of Vision Science* **273–294** (2016).
- Sekirnjak, C. *et al.* High-Resolution Electrical Stimulation of Primate Retina for Epiretinal Implant Design. *Journal of Neuroscience* **28**, 4446–4456 (2008).
- Jepson, L. H. *et al.* Focal Electrical Stimulation of Major Ganglion Cell Types in the Primate Retina for the Design of Visual Prostheses. *Journal of Neuroscience* **33**, 7194–7205, <https://doi.org/10.1523/JNEUROSCI.4967-12.2013> (2013).
- Boinagrov, D., Pangratz-Fuehrer, S., Goetz, G. & Palanker, D. Selectivity of Direct and Network-mediated Stimulation of the Retinal Ganglion Cells with Epi-, Sub- and Intra-Retinal Electrodes. *Journal of Neural Engineering* **11**, 1–21 (2014).
- Lee, J. I. & Im, M. Non-rectangular waveforms are more charge-efficient than rectangular one in eliciting network-mediated responses of ON type retinal ganglion cells. *Journal of Neural Engineering* **15**, 4554–4554 (2018).
- Twyford, P. & Fried, S. I. The Retinal Response to Sinusoidal Electrical Stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **24**, 413–423 (2016).
- Freeman, D. K., Eddington, D. K., Rizzo, J. F., & Fried, S. I. Selective Activation of Neuronal Targets With Sinusoidal Electric Stimulation. *Journal of Neurophysiology*, **104**(5), 2778–2791, <https://doi.org/10.1152/jn.00551> (2010).
- Cai, C., Twyford, P. & Fried, S. The response of retinal neurons to high-frequency stimulation. *Journal of Neural Engineering* **10** (2013).
- Im, M. & Fried, S. I. Indirect activation elicits strong correlations between light and electrical responses in ON but not OFF retinal ganglion cells. *The Journal of Physiology* **593**, 3577–96 (2015).
- Im, M. & Fried, S. I. Temporal properties of network-mediated responses to repetitive stimuli are dependent upon retinal ganglion cell type. *Journal of Neural Engineering* **13**, 1–21 (2016).
- Im, M., Werginz, P. & Fried, S. I. Electric stimulus duration alters network-mediated responses depending on retinal ganglion cell type. *Journal of Neural Engineering* **15** (2018).
- Lee, J.-I. & Im, M. Optimal Electric Stimulus Amplitude Improves the Selectivity Between Responses of ON Versus OFF Types of Retinal Ganglion Cells. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **27**, 2015–2024 <https://ieeexplore.ieee.org/document/8822708/> (2019).
- Victor, J. D. Nonlinear systems analysis: comparison of white noise and sum of sinusoids in a biological system. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 996–8 <http://www.ncbi.nlm.nih.gov/pubmed/284424> (1979).
- Marmarelis, P. Z. & Naka, K.-I. White-Noise Analysis of a Neuron Chain: An Application of the Wiener Theory. *Science* **175**, 1276–1278 (1972).
- Chichilnisky, E. J. A simple white noise analysis of neuronal light responses. *Network: Computational Neural Systems* **12**, 199–213 (2001).
- Schwartz, O., Pillow, J. W., Rust, N. C. & Simoncelli, E. P. Spike-triggered neural characterization. *Journal of Vision* **6**, 484–507 <http://jov.arvojournals.org/article.aspx?articleid=2192881> (2006).
- Chagas, A. M. *et al.* Functional analysis of ultra high information rates conveyed by rat vibrissal primary afferents. *Frontiers in Neural Circuits* **7**, 1–17 <http://journal.frontiersin.org/article/10.3389/fncir.2013.00190/abstract> (2013).
- Freeman, D. K., Rizzo, J. F. & Fried, S. I. Electric stimulation with sinusoids and white noise for neural prostheses. *Frontiers in Neuroscience* **4**, 2–4 (2010).
- Maturana, M. I. *et al.* A Simple and Accurate Model to Predict Responses to Multi-electrode Stimulation in the Retina. *PLoS Computational Biology* **12**, 1–26, <https://doi.org/10.1371/journal.pcbi.1004849> (2016).
- Maturana, M. I. *et al.* Electrical receptive fields of retinal ganglion cells: Influence of presynaptic neurons. *PLoS Computational Biology* **14**, e1005997, <https://doi.org/10.1371/journal.pcbi.1005997> (2018).
- Sekhar, S., Jalligampala, A., Zrenner, E. & Rathbun, D. L. Tickling the retina: integration of subthreshold electrical pulses can activate retinal neurons. *Journal of Neural Engineering* **13**, 046004, <http://stacks.iop.org/1741-2552/13/i=4/a=046004?key=crossref.23132a21808a360924f1ec842690ff70> (2016).

41. Sekhar, S., Jalligampala, A., Zrenner, E. & Rathbun, D. L. Correspondence between visual and electrical input filters of ON and OFF mouse retinal ganglion cells. *Journal of Neural Engineering* **14** (2017).
42. Barriga-Rivera, A., Baret, L., Goding, J., Aregueta-Robles, U. A. & Suaning, G. J. Visual prosthesis: Interfacing stimulating electrodes with retinal neurons to restore vision. *Frontiers in Neuroscience* **11**, 1–17 (2017).
43. Farrow, K. & Masland, R. H. Physiological clustering of visual channels in the mouse retina. *Journal of Neurophysiology* **105**, 1516–1530, <https://doi.org/10.1152/jn.00331.2010> (2011).
44. Stutzki, H., Helmhold, F., Eickenscheidt, M. & Zeck, G. Subretinal electrical stimulation reveals intact network activity in the blind mouse retina. *Journal of Neurophysiology* **116**, 1684–1693, <https://doi.org/10.1152/jn.01095.2015> (2016).
45. Corna, A., Herrmann, T. & Zeck, G. Electrode-size dependent thresholds in subretinal neuroprosthetic stimulation. *Journal of Neural Engineering* **15** (2018).
46. Jensen, R. J., Ziv, O. R. & Rizzo, J. F. Thresholds for activation of rabbit retinal ganglion cells with relatively large, extracellular microelectrodes. *Investigative Ophthalmology and Visual Science* **46**, 1486–1496 (2005).
47. Baden, T. *et al.* The functional diversity of retinal ganglion cells in the mouse. *Nature* **529**, 345–350, <https://doi.org/10.1038/nature16468> (2016).
48. Fine, I. & Boynton, G. M. Pulse trains to percepts: The challenge of creating a perceptually intelligible world with sight recovery technologies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370** (2015).
49. Rathbun, D., Ghorbani, N., Shabani, H., Zrenner, E. & Hosseinzadeh, Z. Spike-triggered average electrical stimuli as input filters for bionic vision - a perspective. *Journal of Neural Engineering*, <https://doi.org/10.1088/1741-2552/aae493Manuscript> (2018).
50. Ho, E. *et al.* Spatio-temporal characteristics of retinal response to network-mediated photovoltaic stimulation. *Journal of Neurophysiology* **116**, 00872.2016, <https://doi.org/10.1152/jn.00872.2016> (2018).
51. Eickenscheidt, M., Jenkner, M., Thewes, R., Fromherz, P. & Zeck, G. Electrical stimulation of retinal neurons in epiretinal and subretinal configuration using a multicapacitor array. *Journal of Neurophysiology* **107**, 2742–2755 (2012).
52. Jensen, R. J. & Rizzo, J. F. Activation of ganglion cells in wild-type and rd1 mouse retinas with monophasic and biphasic current pulses. *Journal of Neural Engineering* **6** (2009).
53. Cho, A., Ratliff, C., Sampath, A. & Weiland, J. Changes in ganglion cell physiology during retinal degeneration influence excitability by prosthetic electrodes. *Journal of Neural Engineering* **13**, 0, <https://doi.org/10.1088/1741-2560/13/2/025001> (2016).
54. Jalligampala, A., Sekhar, S., Zrenner, E. & Rathbun, D. L. Optimal voltage stimulation parameters for network-mediated responses in wild type and rd10 mouse retinal ganglion cells. *Journal of Neural Engineering* **14** (2017).
55. Jensen, R. J., Ziv, O. R. & Rizzo, J. F. Responses of rabbit retinal ganglion cells to electrical stimulation with an epiretinal electrode. *Journal of Neural Engineering* **2**, 16–21 <http://www.ncbi.nlm.nih.gov/pubmed/15876650>. (2005).
56. Tsai, D., Morley, J. W., Suaning, G. J. & Lovell, N. H. Direct Activation and Temporal Response Properties of Rabbit Retinal Ganglion Cells Following Subretinal Stimulation. *Journal of Neurophysiology* **102**, 2982–2993 (2009).
57. Beyeler, M., Nanduri, D., Weiland, J. D., Rokem, A. & Geoffrey, M. A model of ganglion axon pathways accounts for percepts elicited by retinal implant. *bioRxiv* <https://doi.org/10.1101/453035v2> (2018).
58. Jepson, L. H. *et al.* Spatially Patterned Electrical Stimulation to Enhance Resolution of Retinal Prostheses. *Journal of Neuroscience* **34**, 4871–4881, <https://doi.org/10.1523/JNEUROSCI.2882-13.2014> (2014).
59. Fan, V. H. *et al.* Epiretinal stimulation with local returns enhances selectivity at cellular resolution. *Journal of Neural Engineering* **16**, 025001 <http://stacks.iop.org/1741-2552/16/i=2/a=025001?key=crossref.1fc1dd5a07955fb38e33d939c0f38728> (2019).
60. Walker, E. Y. *et al.* Inception loops discover what excites neurons most using deep predictive models. *Nature Neuroscience*, **22**, <https://doi.org/10.1038/s41593-019-0517-x> (2019).
61. Gargini, C., Terzibasi, E., Mazzoni, F. & Strettoi, E. Retinal organization in the retinal degeneration 10 (rd10) mutant mouse: A morphological and ERG study. *Journal of Comparative Neurology* **500**, 222–238 (2007).
62. Jones, B. W. *et al.* Retinal remodeling in human retinitis pigmentosa. *Experimental Eye Research* **150**, 149–165, <https://doi.org/10.1016/j.exer.2016.03.018> (2016).
63. Marc, R. E., Jones, B. W., Watt, C. B. & Strettoi, E. Neural remodeling in retinal degeneration. *Progress in Retinal and Eye Research* **22**, 607–655 (2003).
64. Stutzki, H., Leibig, C., Andreadaki, A., Fischer, D. & Zeck, G. Inflammatory stimulation preserves physiological properties of retinal ganglion cells after optic nerve injury. *Frontiers in Cellular Neuroscience* **8**, 1–10 (2014).
65. Bertotti, G. *et al.* A CMOS-based sensor array for in-vitro neural tissue interfacing with 4225 recording sites and 1024 stimulation sites. *IEEE 2014 Biomedical Circuits and Systems Conference Proceedings* 304–307 (2014).
66. Zeck, G., Jetter, F., Channappa, L., Bertotti, G. & Thewes, R. Electrical Imaging: Investigating Cellular Function at High Resolution. *Advanced Biosystems* **1700107**, 1700107, <https://doi.org/10.1002/adbi.201700107> (2017).
67. Zeitler, R., Fromherz, P. & Zeck, G. Extracellular voltage noise probes the interface between retina and silicon chip. *Applied Physics Letters* **99** (2011).
68. Ferguson, L. R., Dominguez, J. M., Balaiya, S., Grover, S. & Chalam, K. V. Retinal Thickness Normative Data in Wild-Type Mice Using Customized Miniature SD-OCT. *PLoS ONE* **8**, 1–8 (2013).
69. Oesterle, J. *et al.* Bayesian inference for biophysical neuron models enables stimulus optimization for retinal neuroprosthetics. *bioRxiv*, 1–34 (2020).
70. Bertotti, G. *et al.* Artifact-compensated time-continuous recording from neural tissue during stimulation using a capacitively coupled in-vitro CMOS-MEA with 4k recording and 1k stimulation sites. In *Proceedings - 2016 IEEE Biomedical Circuits and Systems Conference, BioCAS 2016*, 256–259 (2016).
71. Leibig, C., Wachtler, T. & Zeck, G. Unsupervised neural spike sorting for high-density microelectrode arrays with convolutive independent component analysis. *Journal of Neuroscience Methods* **271**, 1–13, <https://doi.org/10.1016/j.jneumeth.2016.06.006> (2016).
72. Jones, J. P. & Palmer, L. A. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology* **58**, 1187–1211 (1987).
73. Bialek, W. & de Ruyter van Steveninck, R. R. Features and dimensions: Motion estimation in fly vision. *arXiv e-prints* <http://arxiv.org/abs/q-bio/0505003> (2005).
74. Aljadeff, J., Lansdell, B. J., Fairhall, A. L. & Kleinfeld, D. Analysis of Neuronal Spike Trains, Deconstructed. *Neuron* **91**, 221–259, <https://doi.org/10.1016/j.neuron.2016.05.039> (2016).
75. Nelder, J. A. & Wedderburn, R. W. M. Generalized Linear Models. *Journal of the Royal Statistical Society* **135**, 370–384 <http://www.jstor.org/stable/2344614> (1972).
76. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software* **33** (2010).
77. Paninski, L. Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems* **15**, 243–262 (2004).
78. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830, http://www.americanbanker.com/issues/179_124/which-city-is-the-next-big-fintech-hub-new-york-stakes-its-claim-1068345-1.html (2011).

Acknowledgements

This work was supported by the Baden-Württemberg Stiftung (NEU013), the German Ministry of Science and Education (BMBF, FKZ 01GQ1601 and FKZ 031L0059A) and the German Research Foundation through a Heisenberg Professorship (BE5601/4-1) and the SFB 1233 "Robust Vision" (project number 276693517).

Author contributions

L.H., P.B. and G.Z. designed the study. L.H. performed experiments and analyses. J.O. performed simulations of current spread. P.B. and G.Z. supervised the project. L.H. drafted the manuscript with input from P.B. and G.Z. All authors reviewed the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-61899-y>.

Correspondence and requests for materials should be addressed to G.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020