

Visual coding with a population of direction-selective neurons

Michele Fiscella,¹  Felix Franke,¹ Karl Farrow,² Jan Müller,¹ Botond Roska,³
Rava Azeredo da Silveira,^{4,5} and Andreas Hierlemann¹

¹Bio Engineering Laboratory, ETH Zurich, Basel, Switzerland; ²Neuro-Electronics Research Flanders IMEC, Leuven, Belgium; ³Neural Circuits Laboratory, Friedrich Miescher Institute, Basel, Switzerland; ⁴Department of Physics, Ecole Normale Supérieure, Paris, France; and ⁵Laboratoire de Physique Statistique, Centre National de la Recherche Scientifique, Université Pierre et Marie Curie, Université Denis Diderot, Paris, France

Submitted 17 November 2014; accepted in final form 13 August 2015

Fiscella M, Franke F, Farrow K, Müller J, Roska B, Azeredo da Silveira R, Hierlemann A. Visual coding with a population of direction-selective neurons. *J Neurophysiol* 114: 2485–2499, 2015. First published August 19, 2015; doi:10.1152/jn.00919.2014.—The brain decodes the visual scene from the action potentials of ~20 retinal ganglion cell types. Among the retinal ganglion cells, direction-selective ganglion cells (DSGCs) encode motion direction. Several studies have focused on the encoding or decoding of motion direction by recording multiunit activity, mainly in the visual cortex. In this study, we simultaneously recorded from all four types of ON-OFF DSGCs of the rabbit retina using a microelectronics-based high-density microelectrode array (HDMEA) and decoded their concerted activity using probabilistic and linear decoders. Furthermore, we investigated how the modification of stimulus parameters (velocity, size, angle of moving object) and the use of different tuning curve fits influenced decoding precision. Finally, we simulated ON-OFF DSGC activity, based on real data, in order to understand how tuning curve widths and the angular distribution of the cells' preferred directions influence decoding performance. We found that probabilistic decoding strategies outperformed, on average, linear methods and that decoding precision was robust to changes in stimulus parameters such as velocity. The removal of noise correlations among cells, by random shuffling trials, caused a drop in decoding precision. Moreover, we found that tuning curves are broad in order to minimize large errors at the expense of a higher average error, and that the retinal direction-selective system would not substantially benefit, on average, from having more than four types of ON-OFF DSGCs or from a perfect alignment of the cells' preferred directions.

coding; direction-selective system; microelectrode array; retina; retinal ganglion cells

DETECTING MOTION is a fundamental task in vision, and a large part of the visual system is devoted to processing motion information. Aside from photoreceptors, possibly the most-studied visual neurons are the retinal direction-selective ganglion cells (DSGCs) (Barlow and Hill 1963). DSGCs encode the direction of a moving object by responding vigorously to motion along a so-called “preferred direction” and weakly along the opposite direction, the so-called “null direction.”

DSGCs have been extensively studied in the mammalian retina, mostly in rabbits (Amthor et al. 1984; Oyster et al. 1972; Wyatt and Daw 1975; Yang and Masland 1994) and recently in mice (Weng et al. 2005). Furthermore, DSGCs have

been investigated in fish (Nikolaou et al. 2012), amphibians (Kühn and Gollisch 2013), reptiles (Jensen and DeVoe 1983), and insects (Maisak et al. 2013).

Several studies have focused on motion coding in the retina (Berry et al. 1999; Frechette et al. 2005; Schwartz et al. 2007), and a great deal of work has been devoted to understanding the circuit mechanisms responsible for direction selectivity (Ariel and Daw 1982; Barlow and Levick 1965; Briggman et al. 2011; Euler et al. 2002; Fried et al. 2002; Jeon et al. 2002; Yonehara et al. 2013; Yoshida et al. 2001), but less attention has been given to examining the coding properties of DSGCs.

Four ON-OFF DSGCs are present in the mammalian retina: each comes with a preferred direction along one of the four cardinal body axes (temporal, superior, nasal, and inferior) (Oyster and Barlow 1967). The axons of retinal ON-OFF DSGCs project to the superior colliculus and the thalamus (Levick et al. 1969; Pu and Amthor 1990), and, additionally, as recently discovered, there is a circuit linking retinal ON-OFF DSGCs to neurons in the visual cortex (Cruz-Martin et al. 2013). Neurons that respond asymmetrically to motion have also been identified in the thalamus (Levick et al. 1969; Marshel et al. 2012; Piscopo et al. 2013) and in the visual cortex (Hubel and Wiesel 2009; Ohki et al. 2005) of various species, and recently a number of studies have quantified the coding of directional information by populations of cortical and thalamic neurons (Graf et al. 2011; Stanley et al. 2012). ON-OFF DSGC population coding studies have been carried out in the retina by Amthor et al. (2005) in which simultaneous extracellular recordings from neighboring ON-OFF DSGCs were used to investigate stimulus-dependent correlations. Furthermore, Nowak et al. (2011) studied how direction, speed, and luminance of the stimulus could be separated by quantifying cell responses in terms of spike count and by using an encoding model based on multiplication of three independent functions. We simultaneously recorded from, on average, populations of ~12 direction-selective cells, featuring all four different preferred directions, and decoded their stimulus-dependent concerted activity.

In the present report, we address the coding of directional information by populations of ON-OFF DSGCs; its content is twofold. We present population recordings of ON-OFF DSGC outputs. We quantify the coding performance of the population by various theoretical methods. We then first derive the precision of directional coding using an optimal decoder and examine the dependence of this precision upon stimulus parameters (velocity, object size, angle of moving object) and decoder

Address for reprint requests and other correspondence: M. Fiscella, Bio Engineering Laboratory, ETH Zurich, Basel, Switzerland (e-mail: michele.fiscella@bsse.ethz.ch).

properties (complexity of decoders, noise assumption of responses, timescale to quantify responses). In a second step, we investigate the way in which the coding performance depends upon single-cell properties (shape of the tuning curves) and population properties (relative arrangements of tuning preferences, precision of tuning preferences), as all ON-OFF DSGCs were recorded from simultaneously.

Since the spiking outputs of retinal ON-OFF DSGCs were recorded simultaneously, we have access to the statistics of noise correlations, which may be relevant to coding. In the present work, we assessed the effect of such noise correlations on the decoding performance by randomly shuffling the trial order among multiple cells. However, a more detailed analysis of noise correlation is, because of the required comprehensive treatment, the subject of a separate article (Franke et al., manuscript in preparation).

MATERIALS AND METHODS

Retina Preparation

All animal experiments and procedures were approved by the Swiss Veterinary Office. Female New Zealand albino rabbits (2.5–3 kg) were obtained from Charles River Laboratories (L'Arbresle Cedex, France). The eyes were dissected under dim red light conditions in Ames solution (Sigma, A1420) continuously equilibrated with 5% CO₂-95% O₂. The vitreous was removed, and a retina patch (~2 × 2 mm²) was isolated between 1-mm eccentricity and 4-mm eccentricity, along the inferior direction. To keep track of the retina orientation, we marked the inferior region of the eye during dissection. We then isolated a retina patch that was marked in order to also track retinal inferior coordinates on the microelectrode array. The retina patch was placed ganglion cell side down on the electrode array and perfused with Ames solution (pH 7.4, 36°C) equilibrated with 5% CO₂-95% O₂. Retinal ganglion cell extracellular activity was recorded for 7–8 h.

Light Stimulation

For retinal light stimulation, we used the light projector Acer K10 (60 Hz refreshing rate) in a previously developed setup in which extracellular electrophysiological measurements under light projections can be performed on a microscope stage. More details about the microscope and the optics for focusing the light stimulus on the retina can be found in Fiscella et al. (2012).

We simultaneously used blue (460 ± 15 nm) and green (523 ± 23 nm) projector LEDs for stimulating retinal photoreceptors, and we applied the following light stimuli in our experiments.

Moving bar. This stimulus was used to characterize directional tuning properties of the ganglion cells. The bar length (dimension parallel to the movement direction of the bar) was always 1 mm in order to separate ON and OFF responses. The bar completely traversed the receptive fields of the cells. In experiments aimed at decoding motion direction, the bar moved at a constant velocity of 1.6 mm/s and thus traveled 2.6 mm in the time of stimulus presentation (1.6 s). The length of the leading edge of the bar was between 0.5 and 1 mm, and the receptive fields featured diameters between ~0.3 and 0.5 mm. Background irradiance was 0.2 μW/cm², and bar stimulus irradiance was 2.2 μW/cm².

In experiments aimed at decoding motion direction, the bar moved at a constant velocity of 1.6 mm/s along 36 equidistant angular directions radially spaced at 10°. Three different bar widths (extension orthogonal to the movement direction of the bar) were tested: 0.5, 0.8, and 1 mm. In addition, for experiments aimed at decoding motion direction across different stimulus velocities (0.4, 0.6, 0.8, 1.2, 1.6

mm/s), because of experimental time limits, the bar moved along 18 equidistant angular directions radially spaced at 20° intervals. Bar width was 0.5 mm.

Finally, for measuring cell responses to different stimulus parameters, such as velocity ($n = 5$) and bar width ($n = 3$), because of experimental time limits, the bar moved along eight equidistant angular directions radially spaced at 45° intervals.

Moving flashing square. This stimulus was used to map ganglion cell receptive fields. A positive-contrast square stimulus (0.08 × 0.08 mm²) was turned on for 1 s and turned off for 1 s. The stimulus moved sequentially in discrete nonoverlapping steps over an area of ~1 mm² with a background irradiance of 0.2 μW/cm² and a square stimulus irradiance of 2.2 μW/cm².

Microelectrode Array Recordings

For extracellular recordings we used a microelectronics-based high-density microelectrode array (HDMEA) (Fiscella et al. 2012; Frey et al. 2009). The microelectrode array features 11,011 platinum electrodes with diameters of 7 μm and electrode center-to-center distances of 18 μm over an area of 2 × 1.75 mm². One hundred twenty-six electrodes can be arbitrarily selected for simultaneous recording. Extracellular action potentials were recorded at a sampling frequency of 20 kHz.

In Fig. 1, we show the strategy for recording from a defined population of retinal ganglion cells by the microelectronics-based HDMEA.

1) We let the retina patch settle down on the HDMEA for 30 min before starting light stimulation (Fig. 1A, *step 1*).

2) We recorded light-induced ganglion cell activity by high-density electrode blocks (3,161 electrode/mm²) that were placed below the area where the stimulus was shown. Importantly, we sorted single-cell activity online during the experiment to find a set of electrodes that recorded ON-OFF DSGC activity (1–2 h; Fig. 1A, *step 2*).

3) Once we knew the location of ON-OFF DSGCs, we assigned five to seven electrodes per cell and recorded the light-induced concerted activity of a population of ON-OFF DSGCs (4–5 h; Fig. 1A, *step 3*).

Spike sorting was carried out manually by using the software UltraMegaSort2000 (Hill et al. 2011) on nonoverlapping groups of five to seven electrodes. All recorded traces were band-pass filtered (500 Hz–3 KHz), and all resulting multichannel spike clusters were manually inspected (Fig. 1B). To verify that each cluster came from a single retinal ganglion cell, the percentage of refractory period violations was estimated by using the number of interspike intervals < 1.5 ms (Fig. 1C). Only clusters with refractory period violations of <2% of all spikes were used for subsequent analysis. Duplicate units were removed by comparing the spatial distributions of extracellular action potentials and receptive field location and by temporal cross-correlation of spike trains. All data analysis steps were performed with the software application MATLAB.

Selection of Retinal ON-OFF Direction-Selective Ganglion Cells

To estimate the cell's directional tuning curve, its responses to all stimuli were averaged. We computed a weighted vector sum, D , of the different motion directions (Taylor and Vaney 2002):

$$D = \left\| \frac{\sum_i \vec{v}_i}{\sum_i r_i} \right\|$$

where S is the number of stimuli, \vec{v}_i is a vector pointing in the direction of the moving bar of stimulus i , while the length of \vec{v}_i represents the cell's spike count, r_i . The resulting scalar D then

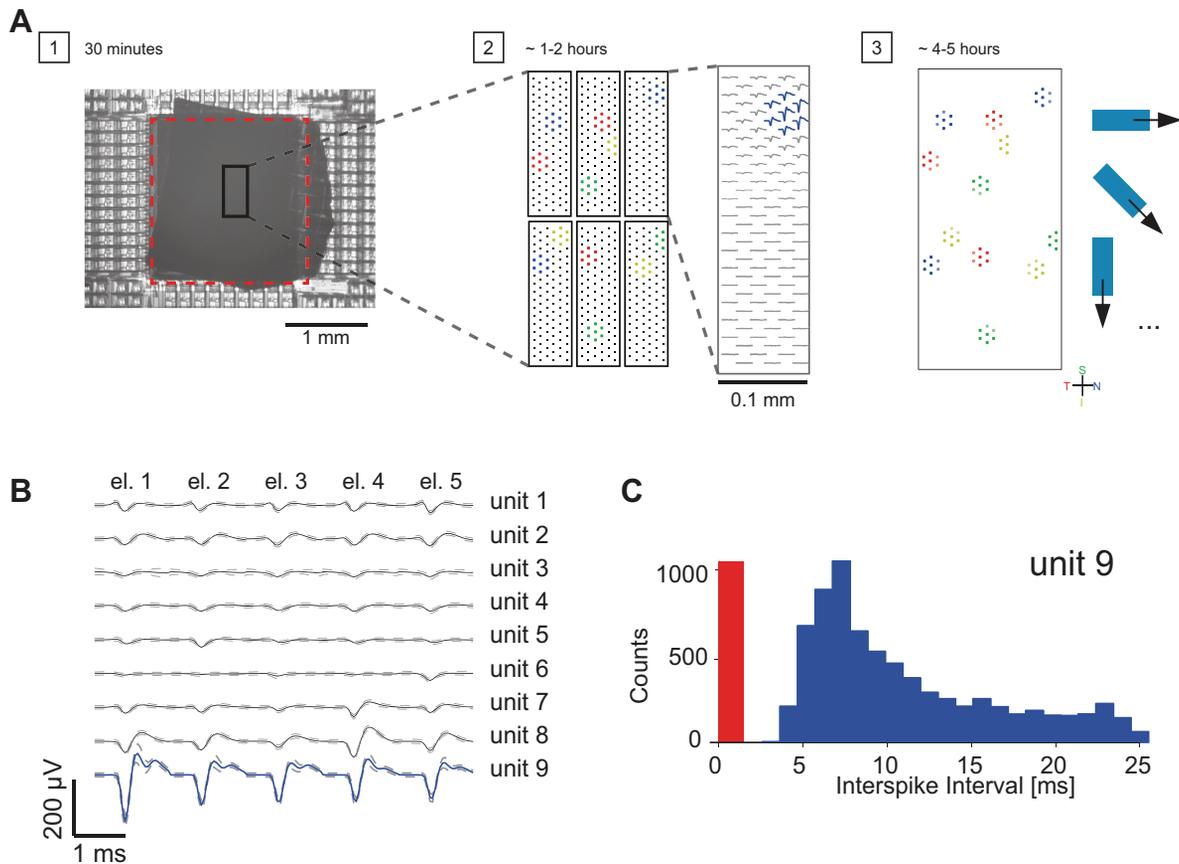


Fig. 1. Method for recording from a defined population of retinal neurons. *A: step 1:* rabbit retina patch, ganglion cell side down on the electrode array (sensor area is shown by dashed red rectangle). *Step 2:* scanning of the ganglion cell layer by high-density electrode blocks in order to find locations of defined cell types (indicated by different colors). *Step 3:* assignment of 5–7 electrodes per ganglion cell for simultaneous recording of light-induced activity from identified and defined populations of retinal neurons. *B:* average multichannel spike waveforms of 9 spiking units, isolated from 5 neighboring electrodes (el. 1–el. 5) (solid line, mean voltage; dashed line, SD voltage). Electrodes were chosen so as to feature the largest-amplitude signals of the neurons of interest (here, e.g., all 5 electrodes recorded large signals for *unit 9*). *C:* interspike interval (ISI) distribution for *unit 9* shown in *B*. The red region indicates the ISI violation time (0–1.5 ms).

corresponds to the direction selectivity index (DSI). The DSI ranges from 0 to 1, where 0 indicates a symmetric response to all motional directions of the stimulus and 1 indicates one specific preferred direction.

We also defined an ON-OFF index (OOI) in order to be able to identify ON-OFF cells:

$$OOI = (LE - TE) / (LE + TE)$$

where LE and TE are the responses to the leading and trailing edges of a bar moving in the cell's preferred direction, respectively. We used a positive-contrast bar stimulus; therefore the leading edge evoked an ON response and the trailing edge an OFF response. OOIs ranged from -1 to 1, where -1 indicated an OFF response and 1 an ON response.

Cells with $DSI > 0.2$ and an $|OOI| < 0.8$ were categorized as ON-OFF DSGCs.

Tuning Curve Parameterization

Several parameterizations of direction-selective tuning functions have been suggested, and we tested the most common ones in term of fit quality and decoding performance (Swindale 1998).

We analyzed the following different forms of model tuning functions, $F_i(\theta)$, of neuron i with preferred direction θ_i and stimulus θ .

1) Gaussian:

$$F_i(\theta) = b_i + \frac{c_i}{\sigma_i \sqrt{2\pi}} e^{-\frac{D(\theta, \theta_i)^2}{2\sigma_i^2}}$$

where b_i is a constant reflecting the baseline firing rate of the neuron, c_i controls the peak firing rate, σ_i controls the width of the tuning function, and $D(x,y)$ is the minimal angle between x and y .

2) von Mises function (Elstrott et al. 2008; Oesch et al. 2005):

$$F_i(\theta) = c_i e^{\frac{\kappa_i \cos(\theta - \theta_i)}{\sigma_i^2}}$$

where c_i and σ_i control the peak height and width of the tuning function.

3) Flat-topped von Mises function (Swindale 1998):

$$F_i(\theta) = c_i e^{\kappa_i \cos(\eta_i \sin(\theta - \theta_i) - 1)}$$

where c_i , κ_i , and η_i control the peak firing rate, the width, and the "flatness" of the tuning function, respectively. We fitted the parameters to the mean firing rates of the recorded neurons iteratively, one parameter at a time, by using a one-dimensional minimization procedure, until either convergence or a predefined maximum number of iterations was reached.

4) Half-wave rectified cosine function (Theunissen and Miller 1991):

$$F_i(\theta) = \left[\frac{c_i \cos(\theta - \theta_i) - b_i}{1 - b_i} \right]$$

where $[\cdot]$ is the half-rectification at 0 and c_i and b_i control the peak firing rate and width of the tuning function.

5) Bimodal von Mises:

$$F_i(\theta) = c_{1,i}e^{\frac{1}{\sigma_{1,i}^2}\cos(\theta-\theta_i)} + c_{2,i}e^{\frac{1}{\sigma_{2,i}^2}\cos(\theta-\theta_i+\pi)}$$

where $c_{1,i}$ and $\sigma_{1,i}$ and $c_{2,i}$ and $\sigma_{2,i}$ control the peak heights and widths of two von Mises functions, whose preferred directions are offset by 180°.

Noise Characteristics of ON-OFF Direction-Selective Ganglion Cells

Rather than through a simple Poisson variance-to-mean relation, we found that the variance-to-mean relation was well characterized by a power law: $\sigma_N^2 = \sigma_N^2(\mu) = a\mu^b$ (fitted with MATLAB'S curve-fitting toolbox, where σ_N^2 is the cell's variance and μ its mean response). We therefore modeled the noise characteristics of a population of N ON-OFF DSGCs with an independent multivariate Gaussian distribution with angle-dependent means, $F(\theta) = [F_1(\theta), \dots, F_N(\theta)]^T$ and variances, $C(\theta) = [\sigma_N^2(F_1(\theta)), \dots, \sigma_N^2(F_N(\theta))]^T$.

Decoders

We decoded the stimulus direction from the simultaneous responses (spike counts), $r = [r_1, \dots, r_N]$, of N neurons. For the analysis presented in Figs. 3–5, we quantified ON-OFF DSGC responses for each direction of stimulus motion as the total number of spikes that every cell generated during the entire stimulation time. For the analysis presented in Fig. 6, the response of each cell for a certain direction of stimulus motion was quantified by overlapping time bins ranging from 0.01 s to 0.3 s.

To estimate the coding performance, we computed the output of the decoders for all possible groups of four ON-OFF DSGCs cells with different preferred directions within the same experiment (Figs. 3–5). Results for those groups (“quadruplets”), the receptive fields of which were overlapping, are shown in Fig. 6.

The decoding error was defined as the error between the direction of the true stimulus, θ , and the estimated direction, $\hat{\theta}$, over all directions. The maximum error is equal to 180°. For every direction of motion we computed the root mean squared error (RMSE):

$$RMSE = \sqrt{E[(\theta - \hat{\theta})^2]}$$

Finally, we averaged the RMSE across all motion directions in order to get the average RMSE for every configuration of decoding cells. All decoding results were 10-fold cross-validated.

We tested several decoders, described below.

1) Population vector decoder (Georgopoulos et al. 1986):

$$\hat{V} = \frac{1}{N} \sum_{k=1}^N r_k V_k$$

where \hat{V} and V_k are the vectors pointing in the directions $\hat{\theta}$ and θ_k , respectively, and θ_k is the preferred direction of neuron k .

2) Optimal linear estimator (Salinas and Abbott 1994): For symmetric tuning functions, the optimal linear decoding is given by

$$\hat{V} = \frac{1}{N} \sum_{k=1}^N r_k D_k$$

with

$$D_k = \sum_j Q_{kj}^{-1} V_k$$

The matrix Q can be computed from the tuning functions and captures the overlap among tuning functions:

$$Q_{kj} = \sigma_k^2 \delta_{kj} + \int_0^{2\pi} F_k(\theta) F_j(\theta) d\theta$$

where σ_k^2 is the average variance of the spike count of neuron k , δ_{kj} is the Kronecker delta, and the integral is over the whole stimulus range (Salinas and Abbott 1994). For equidistant preferred directions and identical tuning functions, this decoding method is equivalent to the population vector; in general, this decoder can take into account variations among the different tuning curves.

3) Probabilistic decoder: The maximum-likelihood (ML) decoder is given by

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}(L(\theta, r))$$

where $L(\theta, r)$ is the likelihood of stimulus θ for a population response r . Since we assume in this work that neurons are independent, $L(\theta, r)$ can be written as

$$L(\theta, r) = \prod_k p(r_k|\theta)$$

where $p(r_k|\theta)$ is the probability of observing r_k spikes from neuron k when a stimulus θ is presented. $p(r_k|\theta)$ is the distribution of spike counts for neuron k . We modeled the spike distribution in two ways. In some of the analyses, we used a Poisson distribution:

$$p_{\text{poiss}}(r_k|\theta) = \frac{F_k(\theta)^{r_k} e^{-F_k(\theta)}}{r_k!}$$

In some of the analyses, however, we used a distribution that satisfied the observed sub-Poisson variance, $\sigma_{N,k}^2(\theta) = aF_k(\theta)^b$ (see *Noise Characteristics of ON-OFF Direction-Selective Ganglion Cells*); we modeled the ON-OFF DSGC response by a Gaussian distribution with the corresponding variance:

$$p_{\text{Gauss}}(r_k|\theta) = \frac{1}{\sigma_{N,k}(\theta)\sqrt{2\pi}} e^{-\frac{D(\theta, \theta_k)^2}{2\sigma_{N,k}(\theta)^2}}$$

The Bayesian decoder is similar to the ML decoder. The only difference is that the direction is decoded as the circular average of the likelihood function instead of its maximum.

Decoding Performance Bound

In parallel with decoders, we used the Fisher information in order to obtain a bound on the coding performance of a population of ON-OFF DSGCs. The Fisher information is a stimulus-dependent quantity, calculated as

$$I_F(\theta) = \left\langle \left(\frac{\partial \log P(r|\theta)}{\partial \theta} \right)^2 \right\rangle$$

where the brackets denote an average over the distribution of population responses. For the case of Poisson variability in the spike count, this expression simplifies to

$$I_F(\theta) = \sum_k \frac{F_k'(\theta)^2}{F_k(\theta)}$$

The Fisher information yields a lower bound on the standard deviation of an unbiased deterministic decoder, and thus also the RMSE, via the Cramér-Rao bound,

$$\operatorname{std}(\hat{\theta}) \geq \frac{1}{\sqrt{I_F(\theta)}}$$

We express our results in terms of the average over all stimulus directions, S :

$$\varepsilon_{\text{mean}} = \frac{1}{S} \sum_{i=1}^S \frac{1}{\sqrt{I_F(\theta_i)}}$$

and the maximum Cramér-Rao bound over all stimulus directions:

$$\varepsilon_{\max} = \max_{\theta} \sqrt{\frac{1}{I_F(\theta)}}$$

RESULTS

Spiking Output of Retinal ON-OFF Direction-Selective Ganglion Cells

We used a HDMEA (Fiscella et al. 2012; Frey et al. 2009) to record action potentials from given types of retinal DSGCs. Action potentials of a single cell were recorded over multiple electrodes (Fig. 2A), so that we could use the five to seven electrodes with highest signal amplitudes for a given cell (Fig. 2A) for multichannel spike sorting (Hill et al. 2011). We then distinguished the ON-OFF DSGCs from the other ~20 types of retinal ganglion cells by using two characteristic response properties of ON-OFF DSGCs: 1) ON-OFF DSGCs respond to both positive- and negative-contrast static stimuli (Fig. 2B) or to leading and trailing edges of the moving bar (Fig. 2F); 2) ON-OFF DSGCs respond asymmetrically to a moving object (Fig. 2C).

We recorded from 126 retinal ganglion cells, which obeyed these criteria, from a total of 10 rabbit retinas (Fig. 2D). As

expected, we found four groups of ON-OFF DSGCs, with preferred directions aligned with the temporal, superior, nasal, and inferior body axes (Fig. 2E). In Fig. 2F, we present an example data set with the spike trains of 12 ON-OFF DSGCs simultaneously recorded from the same retina patch. As expected from the minimal overlap of the dendritic fields of ON-OFF DSGCs of the same type, we observed that cells with the same preferred direction tiled the visual field (Fig. 2G) (Amthor et al. 1984, 1989; Oyster et al. 1993; Yang and Masland 1992).

ON-OFF DSGCs responded to movements along the preferred direction with 30.5 ± 19.8 and 34.7 ± 26.7 spikes for the leading edge and the trailing edge of the bar (Fig. 2H). ON-OFF DSGCs had an average OOI (see MATERIALS AND METHODS) equal to -0.06 ± 0.22 (Fig. 2I) and an average DSI (see MATERIALS AND METHODS) equal to 0.6 ± 0.2 (Fig. 2J). More than 93% of ON-OFF DSGCs recorded had a DSI > 0.4.

Coding Precision of Retinal ON-OFF Direction-Selective Ganglion Cells

To characterize the coding performance of a population of ON-OFF DSGCs, we asked with what precision a probabilistic

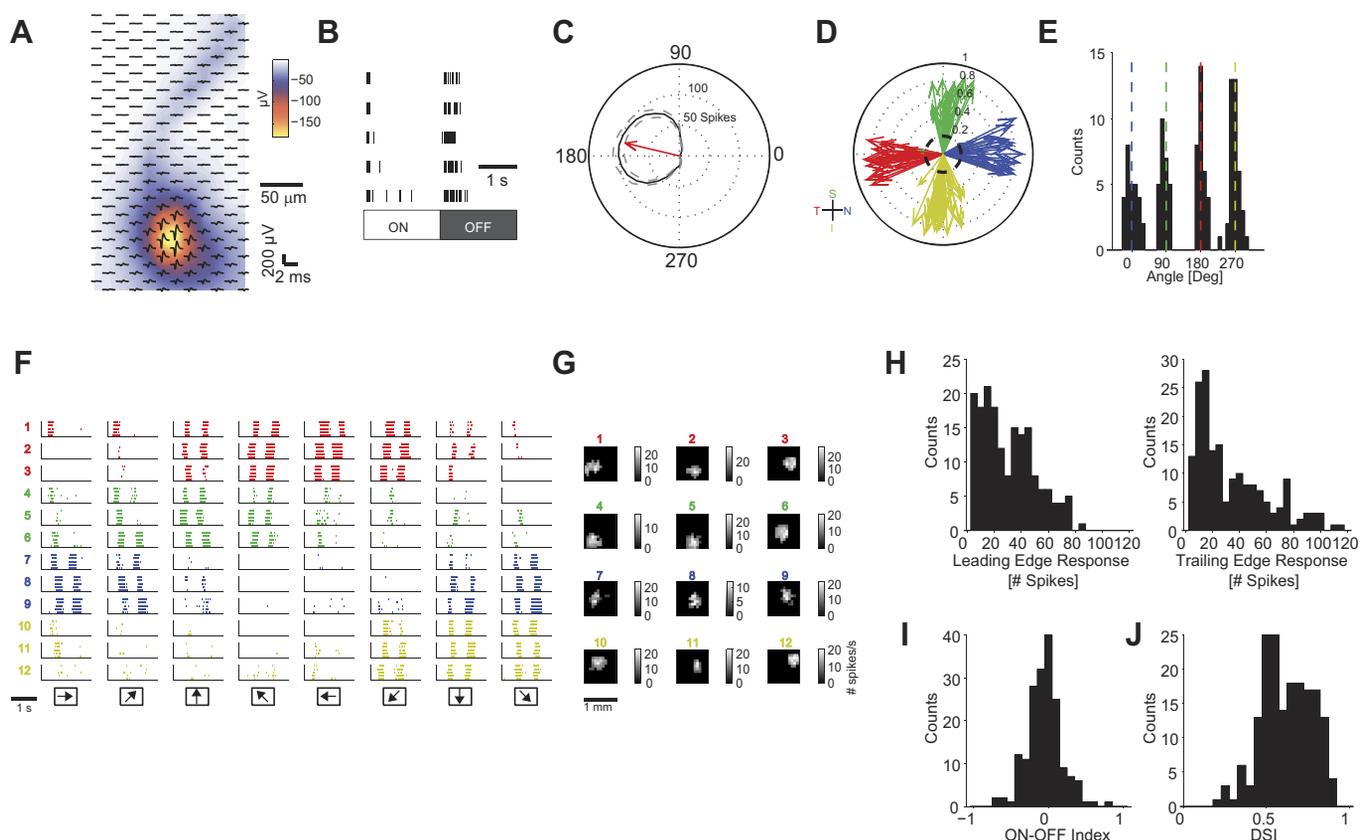


Fig. 2. Extracellular recording from retinal ON-OFF direction-selective ganglion cells (DSGCs). *A*: spatial profile of an action potential of a single ganglion cell. Red-yellow area, high amplitude signals; black waveform: average single-cell action potential on a single electrode (204 electrodes). Amplitude color scale refers to the minimum of the waveform. *B*: raster plot of ganglion cell responses to 5 ON and 5 OFF flashing static stimuli. *C*: tuning curve of ganglion cell responses to motion of a bar. Red arrow points in the preferred direction of the ganglion cell. Black solid line, mean response; dashed gray lines, SD. *D*: preferred directions of 126 ON-OFF DSGCs recorded from 10 retinas. The length of each vector indicates the magnitude of the direction selectivity index (DSI): 0 = not direction selective, symmetric response; 1 = direction selective, asymmetric response (see MATERIALS AND METHODS). T, temporal; S, superior; N, nasal; I, inferior. *E*: histogram of ON-OFF DSGC vector angles from *D*. Nasal = $1.4 \pm 14.1^\circ$, superior = $84.8 \pm 10.8^\circ$, temporal = $182.8 \pm 9.6^\circ$, inferior = $268.4 \pm 13.1^\circ$ (mean \pm SD). *F*: raster plot of ganglion cell responses to motion; simultaneous responses of 12 ON-OFF DSGCs. Responses are colored according to the ganglion cell's preferred directions (red, temporal; green, superior; blue, nasal; yellow, inferior). Each row shows the activity of a single ganglion cell for different directions of motion. Arrows at *bottom* indicate directions of motion. *G*: receptive field locations of ON-OFF DSGCs in *F*. *H*, *left*: distribution of cell responses for bar leading edge (ON response). *Right*: distribution of responses for bar trailing edge (OFF response). *I*: distribution of ON-OFF indexes. *J*: distribution of DSIs. In *H*, *I*, and *J*, statistics are computed from all 126 recorded ON-OFF DS cells across all stimulus parameters used (see MATERIALS AND METHODS).

decoder can retrieve directional information. In principle, if one knows the mean response of each cell as a function of motion direction (direction-tuning curves) (Fig. 3A) as well as the noise characteristics of the response (Fig. 3B), one can construct a ML or Bayesian decoder (see MATERIALS AND METHODS for precise definitions and details) (Dayan and Abbott 2001; Pouget et al. 2000). In this section, we decode

motion directions by using von Mises fits (Elstrott et al. 2008; Oesch et al. 2005) of direction-tuning curves and by assuming Poisson noise of the responses. Closer inspection of the data suggested, however, that the variance-to-mean relation of cell responses was better characterized by a power-law fit (see Fig. 3B). In *Influence of Tuning Curve Shape, Response Variability, and Noise Correlations on*

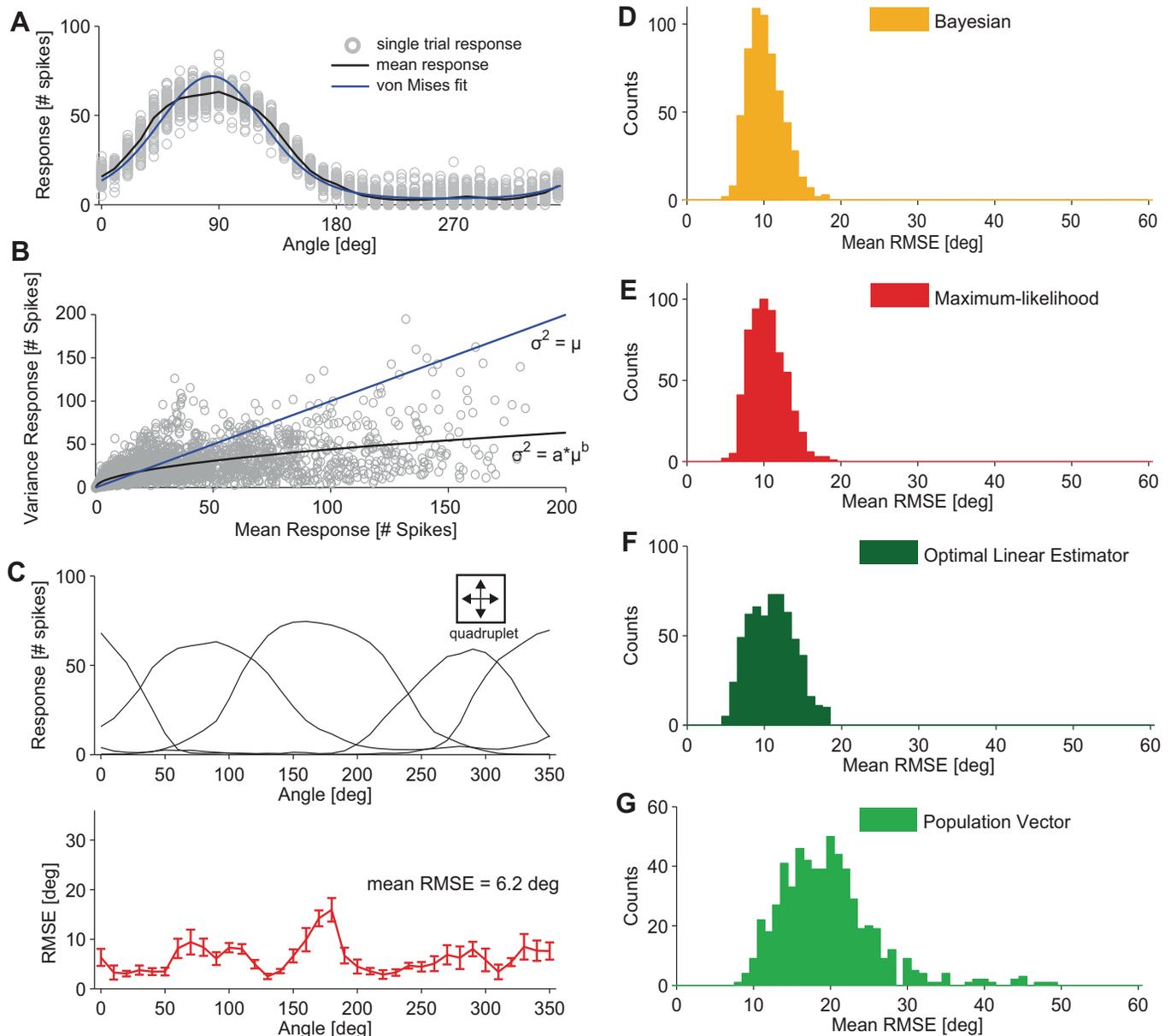


Fig. 3. Coding precision of retinal ON-OFF DSGC populations. *A*: response of a single ON-OFF DSGC to a bar moving along 36 directions, angularly spaced by 10° ; $n = 100$ sweeps per direction. Responses are quantified as the total number of spikes fired by the cell during each stimulus direction (gray circles). The mean response is displayed by the black curve; the von Mises fit to the mean response is displayed by the blue curve. *B*: characterization of the response variability of ON-OFF DSGCs ($n = 126$). The response was quantified as the total number of spikes for a given stimulus. The plot displays the correlation between the mean and variance of ON-OFF DSGC responses for all recorded cells and all motion directions across all applied stimulus parameters. Fit: power function: $\sigma_N^2 = \sigma_N^2(\mu) = a\mu^b$, with $a = 4.03$ and $b = 0.51$, where $\sigma^2 =$ variance and $\mu =$ mean (black line). Poisson variability of responses with a mean equal to the variance is indicated by the blue line. Fano factors of the cells (i.e., the variance of a cell's response to a stimulus divided by its mean) ranged from ~ 0.5 (for motion directions close to the preferred direction of the ON-OFF DSGC) to ~ 1.5 (for motion directions close to the null direction of the ON-OFF DSGC). *C*, top: example group of 4 ON-OFF DSGCs used for decoding motion direction (we refer to such a group as a "quadruplet"). Bottom: root mean squared error (RMSE) along 36 different motion directions angularly spaced at 10° . Error bars show SD due to 10-fold cross-validation (see MATERIALS AND METHODS). The general decoding performance of this quadruplet can be summarized by a mean RMSE equal to 6.2° . *D–G*: mean RMSE distributions results for 4 different decoders; each decoder type is indicated in the key [$n = 600$ quadruplets, bar size = $1 \text{ mm} \times 0.5 \text{ mm}^2$ (length \times width), stimulus velocity = 1.6 mm/s]. Responses were quantified as the total number of spikes fired by the cell during each stimulus.

Coding Performance below, we examine the influence of such sub-Poisson noise, as well as other tuning function models, on the coding performance.

We initially used groupings of four ON-OFF DSGCs with the four cardinal preferred directions; we refer to such groups of cells as quadruplets (Fig. 3C). For each data set, we chose all possible quadruplets, and we quantified the performance of decoding as the RMSE across all motion directions (Fig. 3C) (see MATERIALS AND METHODS). Assuming von Mises tuning functions and Poisson noise to model cell responses, we obtained an average RMSE of $10.2 \pm 2.3^\circ$ for the Bayesian decoder and $10.5 \pm 2.3^\circ$ for the ML decoder (Fig. 3, D and E). The performance of the Bayesian decoder was significantly higher than the performance of the ML decoder ($P < 0.05$, Mann-Whitney *U*-test).

As the brain may use simpler decoders than the probabilistic decoders, we investigated how the decoding precision degrades with reduced decoder complexity. In Fig. 3, F and G, we present decoding analyses carried out with different decoders, namely, the population vector (Georgopoulos et al. 1986) and the optimal linear estimator (Salinas and Abbott 1994). Simpler decoders, such as optimal linear estimator and population vector, on average, performed worse than the probabilistic decoders ($P < 0.01$, Mann-Whitney *U*-test).

Since probabilistic decoders outperformed linear decoders, we carried out the remaining decoding analyses with the Bayesian decoder. A summary of the decoding results is reported in Table 1.

Variability of Coding Precision with Respect to Stimulus Parameters

As the response characteristics of direction-selective neurons vary with stimulus properties, such as bar velocity and bar size (Fig. 4, A and E) (Oyster et al. 1972; Wyatt and Daw 1975), we investigated to what extent these variations influence the coding of directional information.

While the peak of the tuning curve depends strongly on the bar velocity (Fig. 4B) (decrease of 50.4% when bar velocity is varied from 0.4 to 1.6 mm/s), the width of the tuning function remains virtually unchanged (Fig. 4C) (2.9% decrease for the same variation in bar velocity).

Changing the width of the moving bar influenced tuning curve peak and width as well. Specifically, doubling the width of the moving bar from 0.5 mm to 1 mm caused an average 14.6% decrease in the directional peak magnitude (Fig. 4F) and an average 12.7% decrease in the tuning curve width (Fig. 4G). The decreases in peak response and tuning curve width are due to the fact that ON-OFF DSGCs have a large inhibitory

receptive field surround (Barlow and Levick 1965; Wyatt and Daw 1975).

Varying the velocity of the moving stimulus modulated the average RMSE by $<1^\circ$, there was a nonsignificant drop from $9.5 \pm 1.3^\circ$ for a stimulus velocity of 1.6 mm/s to $9.3 \pm 1.7^\circ$ for a stimulus velocity of 0.4 mm/s ($P > 0.05$, Mann-Whitney *U*-test) (Fig. 4D). In contrast, the width of the moving bar influenced the decoding results more substantially. Increasing the width of the moving bar from 0.5 mm to 0.8 mm or 1 mm caused a significant increase of the average RMSE equal to 58% and 70%, respectively ($P < 0.01$, Mann-Whitney *U*-test) (Fig. 4H).

Although the velocity of the stimulus did not influence decoding precision, the increasing size of the bar significantly worsened decoding. Therefore, a single quadruplet is less accurate when decoding the movement direction of objects larger than its receptive fields.

Owing to the shape of the tuning curves, the decoding performance depended on the stimulus direction. Motion directions $\sim 45^\circ$ away from the ON-OFF DSGCs' preferred directions were decoded with lower average RMSE (32% significant decrease, for 0.5-mm bar widths; $P < 0.01$, Mann-Whitney *U*-test) compared with motion directions close to the ON-OFF DSGCs' preferred directions (Fig. 4I). However, for bar widths significantly larger than the receptive fields of the cells (0.8 mm and 1 mm), motion directions $\sim 45^\circ$ distant from the ON-OFF DSGCs' preferred directions were decoded with similar or higher RMSE compared with motion directions close to the ON-OFF DSGCs' preferred directions (Fig. 4I).

Finally, we asked how sharply tuned quadruplets (high DSI) decoded motion directions compared with more broadly tuned quadruplets (low DSI). We found that the decoding error was inversely proportional to the mean DSI of the respective quadruplet (Fig. 4J): sharply tuned quadruplets decoded motion directions with lower error than broadly tuned quadruplets. These results are in agreement with simulations reported in *Influence of Tuning Curve Width on Decoding Performance*. A summary of decoding results is reported in Table 2.

Influence of Tuning Curve Shape, Response Variability, and Noise Correlations on Coding Performance

The tuning curve of ON-OFF DSGCs has been commonly fitted by using von Mises functions (Elstrott et al. 2008; Graf et al. 2011; Oesch et al. 2005; Swindale 1998). We examined the extent to which the functional shape of the tuning curve affected the coding performance, by comparing the latter in the case of von Mises fits (4 free parameters, *fit 1*) to four other cases: Gaussian fit (4 free parameters, *fit 2*), half-wave rectified cosine (4 free parameters, *fit 3*), bimodal von Mises fit (6 free parameters, *fit 4*), and flat-topped von Mises fit (5 free parameters, *fit 5*) (Fig. 5A; see MATERIALS AND METHODS). In each case, we evaluated the goodness of the fit by computing the coefficient of determination, R^2 , for 126 ON-OFF DSGCs. The highest average R^2 (10-fold cross-validated) resulted for the flat-topped von Mises fit (0.97 ± 0.05 , *fit 5*) and was significantly higher than the average R^2 of the von Mises fit (0.93 ± 0.08 , *fit 1*) (Fig. 5B) ($P < 0.01$, Mann-Whitney *U*-test). The flat-topped von Mises functions, which represented the best fit

Table 1. Precision of decoders

Decoder	RMSE, $^\circ$	N Quadruplets 
Maximum likelihood	10.5 ± 2.3	600
Bayesian	10.2 ± 2.3	600
Optimal linear estimator	11.0 ± 2.9	600
Population vector	19.5 ± 6.2	600

Root mean squared error (RMSE) values are means \pm SD. See Fig. 3.

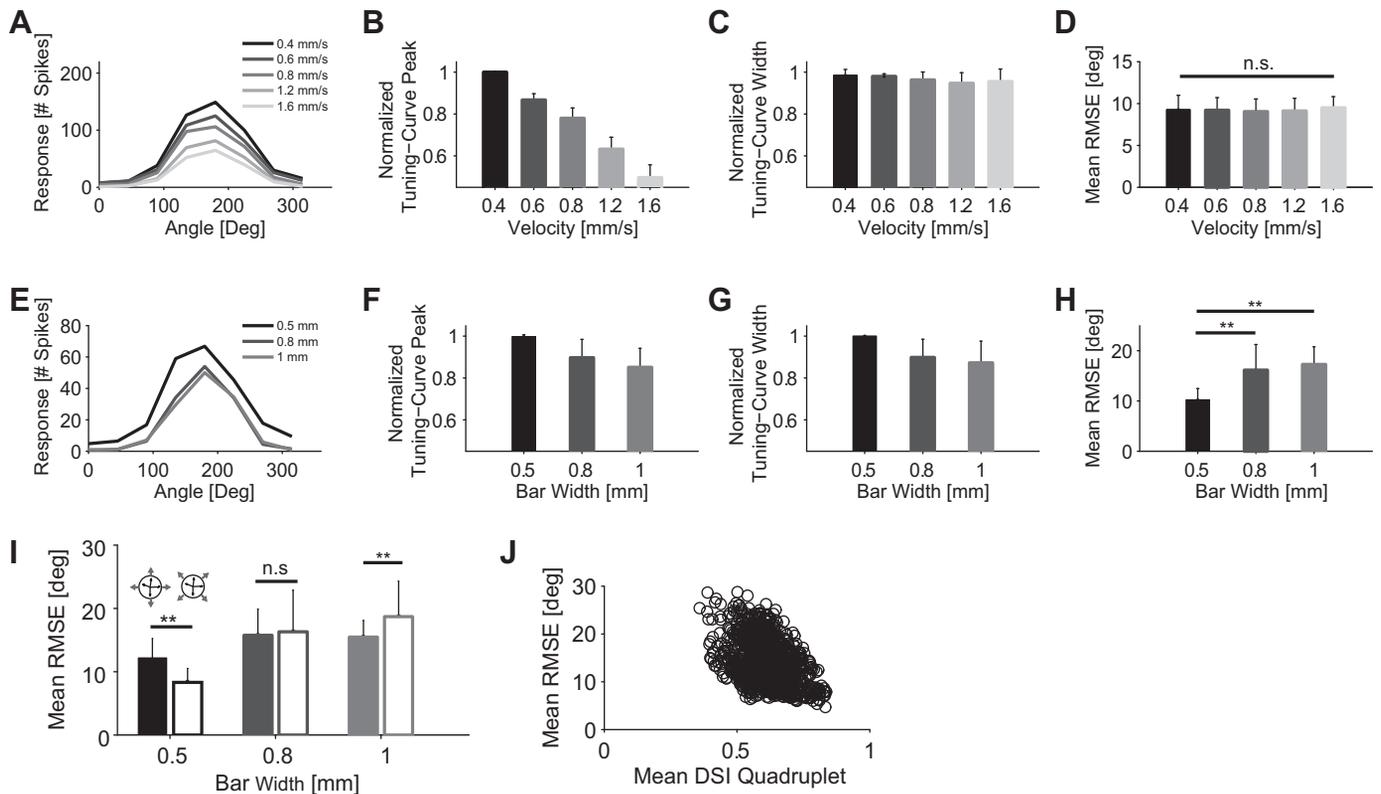


Fig. 4. Influence of stimulus velocity and size on decoding performance. **A:** tuning curves for stimulus velocities ranging from 0.4 to 1.6 mm/s. Bar size was 1 mm \times 0.5 mm (length \times width). The bar moved along to its long edge. **B:** normalized tuning curve peak as a function of stimulus velocity. For every cell, each tuning curve peak value, across all stimulus velocities, was normalized with respect to the highest tuning curve peak value (1 indicates highest tuning curve peak; $n = 10$). **C:** normalized tuning curve width (computed from von Mises fit) as a function of stimulus velocity. For every cell, each tuning curve width, across all stimulus velocities, was normalized with respect to the highest tuning curve width (1 indicates highest tuning curve width; $n = 10$). **D:** mean RMSE for 5 stimulus velocities ($n = 108$ quadruplets of ON-OFF DSGCs). Decoding performed by von Mises fit, Poisson noise, and a Bayesian decoder. Bar size was 1 \times 0.5 mm² (length \times width). Responses were quantified as the total number of spikes fired by the cell during each stimulus. **E:** tuning curves for bar widths ranging from 0.5 to 1 mm. Bar velocity was 1.6 mm/s. **F:** normalized tuning curve peak as a function of bar width (normalized as in **B**; $n = 10$). **G:** normalized tuning curve width (computed from von Mises fit) as a function of bar size (normalized as in **C**; $n = 10$). **H:** mean RMSE for bars with 3 different widths: 0.5 mm ($n = 600$ quadruplets of ON-OFF DSGCs), 0.8 mm ($n = 159$ quadruplets of ON-OFF DSGCs), and 1 mm ($n = 464$ quadruplets of ON-OFF DSGCs). Decoding performed by von Mises fit, Poisson noise, and a Bayesian decoder. Stimulus velocity was 1.6 mm/s. Responses were quantified as the total number of spikes fired by the cell for each stimulus direction. **I:** comparison of decoding performance between motion directions close to the cells' preferred directions (filled bars) and motion directions $\sim 45^\circ$ away from the cells' preferred directions (open bars). Comparison is reported for 3 different bar widths. Arrow at top indicates the decoded motion directions with respect to the cells' preferred directions. **J:** correlation plot between DSI and decoding accuracy. x -Axis displays mean DSI of a quadruplet. y -Axis displays mean RMSE obtained from a quadruplet (Pearson correlation coefficient = -0.4). The plot combines the decoding performance of quadruplets across all 3 bar widths (0.5, 0.8, and 1 mm; $n = 1223$). Error bars indicate SD in all panels. * $P < 0.05$; ** $P < 0.01$. n.s., Not significant.

to the tuning curves of ON-OFF DSGCs, also yielded the smallest average RMSE of $9.0 \pm 2.6^\circ$ compared with $10.2 \pm 2.3^\circ$ for the von Mises fit ($P < 0.01$, Mann-Whitney; Fig. 5C). For the sake of comparison, we also report the average RMSE

Table 2. Variability of coding precision with stimulus parameters

Velocity, mm/s	Bar Size (L \times W), mm	RMSE, $^\circ$	N Quadruplets
0.4	1 \times 0.5	9.3 ± 1.7	108
0.6	1 \times 0.5	9.2 ± 1.5	108
0.8	1 \times 0.5	9.0 ± 1.5	108
1.2	1 \times 0.5	9.1 ± 1.6	108
1.6	1 \times 0.5	9.5 ± 1.3	108
1.6	1 \times 0.5	10.2 ± 2.3	600
1.6	1 \times 0.8	16.1 ± 5.1	159
1.6	1 \times 1	17.3 ± 3.5	464

RMSE values are means \pm SD. See Fig. 4.

obtained from spline interpolation of the tuning curves, which amounts to $7.1 \pm 1.9^\circ$ (Fig. 5C).

Using sub-Poisson noise (Fig. 3B) influenced decoding performance and introduced relative changes between 2% and 6% among the tuning curve fits, although not significant among all tuning curve fits (Fig. 5D).

Finally, we measured the effect of noise correlations on decoding precision by randomly shuffling cell response trials between multiple cells (Averbeck et al. 2006; Cohen and Kohn 2011). For example, decoding of unshuffled responses from four cells means that we used the simultaneously recorded activity of these four cells during *trial 1*. Oppositely, decoding of shuffled responses from four cells means that we used the activity that is not recorded synchronously, and therefore cell responses are obtained by different trials (e.g., *trials 1, 5, 7, and 25* for *cells 1, 2, 3, and 4*, respectively). Random shuffling of the trials before decoding decreased the coding performance by ~ 4 – 5% for all tuning curve fit choices ($P < 0.05$, Mann-Whitney U -test) (Fig. 5E) This finding suggests that noise

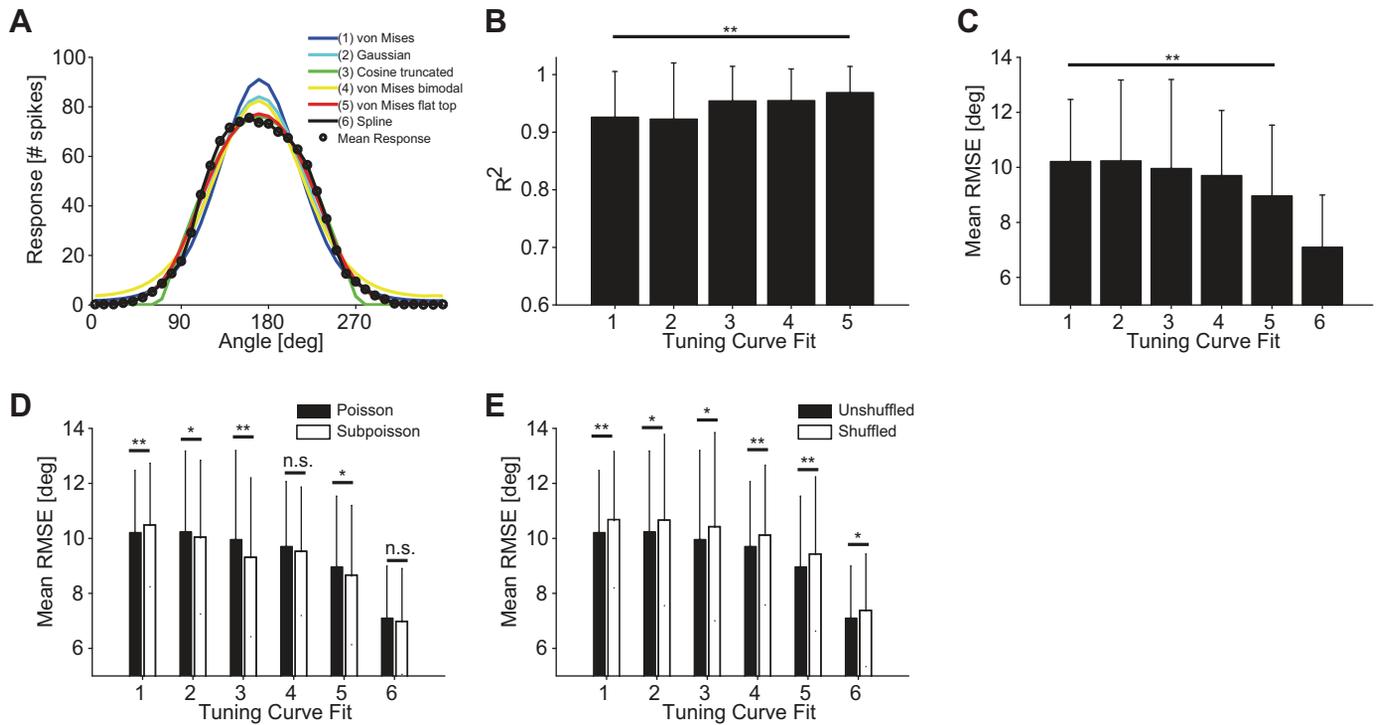


Fig. 5. Influence of tuning curve fit on decoding performance. *A*: various tuning curve fits. The different fit types are indicated in the key by different colors and numbers. *B*: mean coefficient of determination R^2 (10-fold cross-validated) for 126 ON-OFF DSGCs fitted by 5 different functions numbered 1–5 in *A*. *C*: average RMSE using 6 different functional fits of the tuning curves. Fits numbered as in *A* [$n = 600$ quadruplets, bar size = $0.5 \times 1 \text{ mm}^2$ (length \times width), stimulus velocity = 1.6 mm/s]. Decoding results were obtained by using a Bayesian decoder and Poisson noise assumption. Responses were quantified as the total number of spikes fired by the cell for each stimulus. *D*: comparison of decoding accuracies for Poisson and sub-Poisson noise assumptions across all tuning curve fits [$n = 600$ quadruplets, bar size = $0.5 \times 1 \text{ mm}^2$ (length \times width), stimulus velocity = 1.6 mm/s]. Decoding results were obtained by using a Bayesian decoder, and responses were quantified as the total number of spikes fired by the cell for each stimulus. *E*: comparison of decoding accuracies between unshuffled and shuffled trials across all tuning curve fits [$n = 600$ quadruplets, bar size = $0.5 \times 1 \text{ mm}^2$ (length \times width), stimulus velocity = 1.6 mm/s]. Decoding results were obtained by using a Bayesian decoder and Poisson noise assumption. Responses were quantified as the total number of spikes fired by the cell for each stimulus. Error bars indicate SD in all panels. $*P < 0.05$; $**P < 0.01$.

correlations are beneficial to the coding of motion directions in retinal direction-selective cells (Franke et al., manuscript in preparation). A summary of decoding results is reported in Table 3 and Table 4.

Influence of Time Bin Size on Decoding Performance

The results reported in the previous sections were obtained from spike counts within a time bin that covered the total duration of the stimulus presentation. To quantify how bin size influences decoding performance, we repeated the analysis with spike counts derived from overlapping time bins with sizes ranging from 10 ms to 300 ms (Fig. 6).

For a meaningful comparison, we always carried out the analysis on selected ON-OFF DSGCs with overlapping receptive fields and, therefore, with synchronous light-induced re-

sponses (Fig. 6, *A–D*). We decoded cell activity, according to their synchronously recorded spike trains, and obtained an error as function of time (Fig. 6*E*). With this approach, we could separate the decoding precision for the leading edge and the trailing edge of the moving bar (Fig. 6*E*). We found similar decoding accuracies for the leading edge and the trailing edge of the bar. We measured a mean RMSE between 10° and 15° for bin sizes ranging from 100 to 300 ms (Fig. 6*F*). However, for bin sizes smaller than 100 ms, decoding performance started to decrease with a mean RMSE between 20° and 50° for bin sizes ranging from 10 ms to 50 ms (Fig. 6*F*).

Influence of Tuning Curve Width on Decoding Performance

If any direction is to be encoded with ON-OFF DSGCs organized along the four cardinal directions, one expects

Table 3. Precision of tuning curve fits

Tuning Curve Fit	R^2	RMSE, $^\circ$
von Mises (1)	0.93 ± 0.08	10.2 ± 2.3
Gaussian (2)	0.92 ± 0.10	10.2 ± 2.9
Cosine truncated (3)	0.95 ± 0.06	10.0 ± 3.2
von Mises bimodal (4)	0.95 ± 0.05	9.7 ± 2.4
von Mises flat top (5)	0.97 ± 0.05	9.0 ± 2.6
Spline (6)		7.1 ± 1.9

RMSE values are means \pm SD. See Fig. 5.

Table 4. Effect of noise and trial shuffling

Tuning Curve Fit	% RMSE Change by Sub-Poisson Noise	% RMSE Change by Shuffling Trials
von Mises (1)	+3**	+5**
Gaussian (2)	-2*	+4*
Cosine truncated (3)	-6**	+5*
von Mises bimodal (4)	-2 (n.s.)	+4**
von Mises flat top (5)	-3*	+5**
Spline (6)	-2 (n.s.)	+4*

* $P < 0.05$; ** $P < 0.01$; n.s., nonsignificant; Mann-Whitney *U*-test. See Fig. 5.

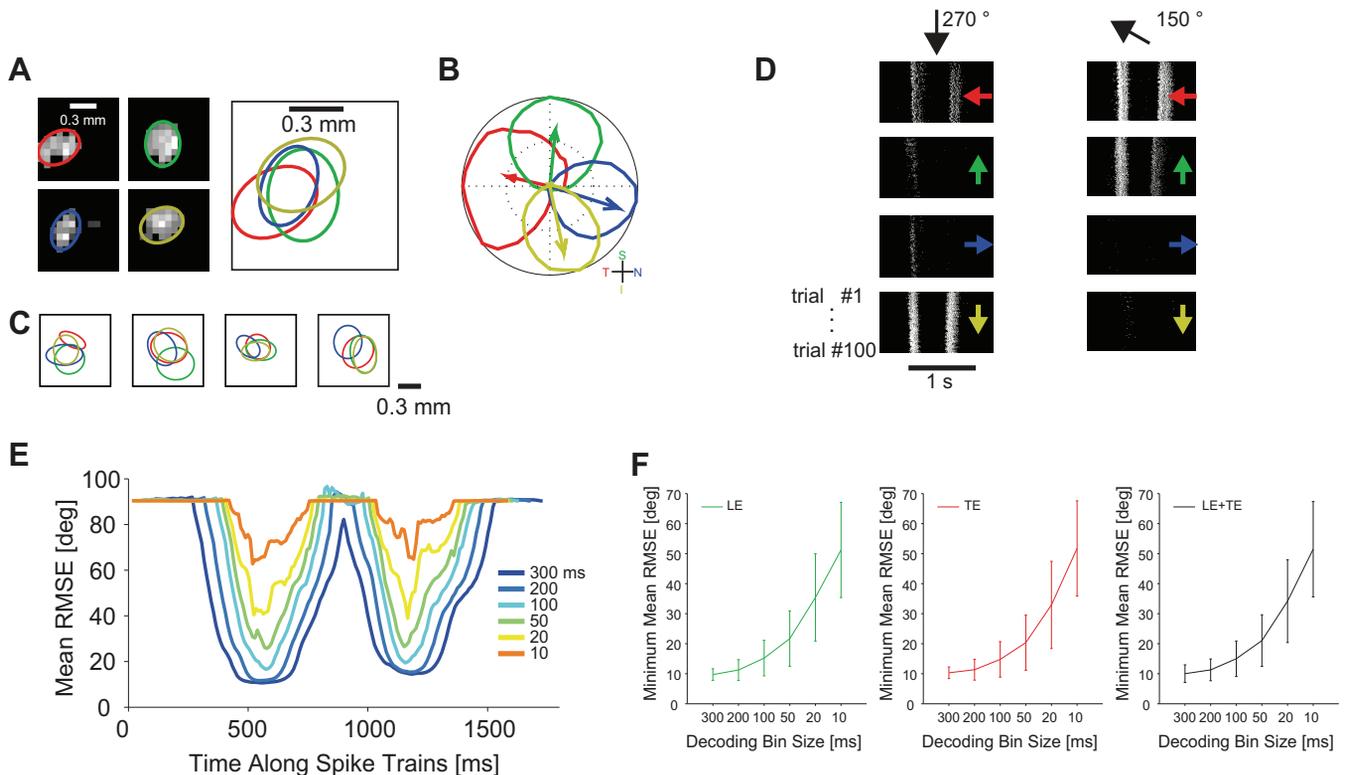


Fig. 6. Influence of time bin size on decoding performance. *A*: overlapping receptive fields of 4 simultaneously recorded ON-OFF DSGCs, fitted by ellipse function. *B*: tuning curves and preferred directions of the cells in *A*. *C*: additional examples of ON-OFF DSGC quadruplets with overlapping receptive fields. *D*: simultaneous responses of ON-OFF DSGCs in *A* for a bar moving in the direction indicated by black arrows at *top*. Each panel shows 100 responses of a single cell. Preferred directions of the respective cells are indicated by colored arrows within the panels. *E*: decoding performance at consecutive time points of cell responses. Each curve represents decoding results obtained with a different temporal bin size to count cell spikes. For decoding we used a von Mises fit for the tuning curves, Poisson variability of responses, and a Bayesian decoder [bar size = $0.5 \times 1 \text{ mm}^2$ (length \times width), stimulus velocity = 1.6 mm/s]. *F*: decoding performance as a function of bin size (ms) for the 5 quadruplets of ON-OFF DSGCs displayed in *A*. Decoding performance is reported separately for leading edge (LE; *left*), trailing edge (TE; *center*), as well as leading and trailing edges (LE+TE; *right*).

ON-OFF DSGCs' tuning curves to be broad. Here, we quantify the dependence of the coding performance on this width (see *Tuning Curve Parameterization*, σ , parameter for von Mises fit); we do so by assuming a von Mises form for the shape of the tuning curve and by establishing bounds, ϵ_{mean} and ϵ_{max} , on the decoding performance over a range of tuning curve widths (see *Decoding Performance Bound*). As we varied tuning curve width, we fixed the integral of the tuning curve, so that we compare cases that share a given total number of spikes (and, hence, a given metabolic cost) and in which each stimulus direction appear with equal probability. Quantifications on coding performance as function of tuning curve widths have been already performed in the context of visual cortical neurons (Seung and Sompolinsky 1993). We assigned tuning curves to four ON-OFF DSGCs with homogeneously distributed preferred directions (Fig. 7, *A* and *B*). The value of the integral of each model tuning curve was drawn from the distribution of the corresponding quantity in the data. We modeled 200 ON-OFF DSGCs quadruplets in which the tuning curve widths ranged from 10° to 70° .

For model tuning curves, the lowest obtained ϵ_{mean} value was 3.7° and corresponded to a narrow tuning curve with a width of 18° (Fig. 7*C*). We also derived the minimum value of the maximum estimation error over stimulus directions, ϵ_{max} , and compared these results with real data. The lowest ϵ_{max} value was 7.2° and corresponded to broader tuning

curves with a width of 36° (Fig. 7*D*). We found that the data (Fig. 7, *C* and *D*) clustered near the optimal ϵ_{max} value, i.e., the tuning curves were wider ($41.1 \pm 7.3^\circ$) than predicted from optimizing the average error but comparable to the prediction obtained by minimizing the maximum error over stimuli. This suggests that tuning curves are sufficiently wide so as to avoid large decoding errors for given stimulus directions.

Influence of Arrangement of Preferred Directions on Population Coding

As mentioned above, roughly speaking, the preferred directions of ON-OFF DSGCs are aligned with the four cardinal directions. But, more precisely, the preferred direction in each ON-OFF DSGC is slightly jittered away from "its" cardinal direction (Fig. 2, *D* and *E*). These considerations naturally beg for quantitative analyses in response to two questions: 1) How much more accurate is coding for preferred directions arranged along the four cardinal directions as opposed to those randomly arranged on the circle? 2) To what extent does the jitter in the cardinal direction affect the coding performance?

To answer these questions, we compared the coding performance for several arrangements of cells: 1) ON-OFF DSGCs with equidistant preferred angular directions (PDs) (Fig. 8*A*). We refer to this model as "Equidistant PD." 2) ON-OFF DSGCs with preferred directions along the cardinal directions

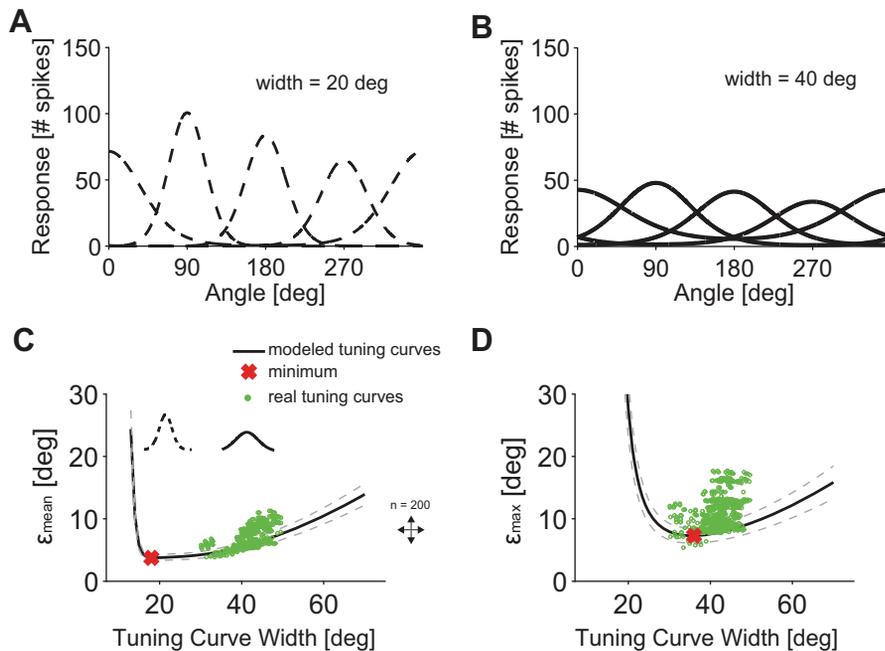


Fig. 7. Influence of tuning curve width on decoding precision. *A* and *B*: 2 examples (solid and dashed lines, von Mises function) of 4 simulated tuning curves with preferred directions separated by 90° for 2 different widths of 40° and 20° . The integral of the tuning curves was fixed. *C*: ϵ_{mean} , computed with model tuning curves, as a function of the tuning curve width (solid line, mean; dashed line, SD). Red cross, lowest ϵ_{mean} value; green dots, ϵ_{mean} values for measured tuning curves ($n = 600$ quadruplets). *D*: ϵ_{max} , computed with model tuning curves, as a function of tuning curve width (solid line, mean; dashed line, SD). Red cross, lowest ϵ_{max} value; green dots, ϵ_{max} values for measured tuning curves ($n = 600$, quadruplets).

(Fig. 8*B*). We refer to this model as “Cardinal PD.” 3) ON-OFF DSGCs with mean preferred directions along the cardinal directions but where individual preferred directions are jittered by a small increment, the distribution of which was derived from the data (Fig. 8*C*). We refer to this model as “Cardinal Jittered PD.” 4) ON-OFF DSGCs with preferred directions randomly distributed over the circle (Fig. 8*D*). We refer to this configuration as “Random PD.”

Arranging only four cells in a Random PD configuration was obviously disadvantageous and resulted in higher ϵ_{mean} and higher ϵ_{max} compared with Cardinal PD and Equidistant PD configurations (Fig. 8, *E* and *F*). However, for a larger number of cells, ϵ_{mean} converged to similar values for all possible angular arrangements of the preferred directions (Fig. 8*E*). Similarly, ϵ_{max} decreased when the number of cells increased, although the Equidistant PD configuration resulted in a ϵ_{max} $\sim 20\%$ lower than that of Cardinal PD and Random PD configurations (Fig. 8*F*).

Direction selectivity is established by spatial asymmetric GABAergic inhibition of ON-OFF DSGCs through starburst amacrine cells (Famiglietti 1992; Fried et al. 2002; Masland and Mills 1979; Tauchi and Masland 1984; Vaney and Pow 2000). Establishing more than four direction-selective cells would require a rearrangement of the inhibitory circuits and, in particular, a more precise spatial distribution of starburst amacrine cell dendrites (Briggman et al. 2011; Euler et al. 2002; Fried et al. 2002). Another physiological process, which is relevant for direction selectivity, is the spatially asymmetric cholinergic facilitation (Grzywacz and Amthor 1993). Facilitation significantly enhances the response of ON-OFF DSGCs along preferred direction, even outside cells’ receptive fields (Amthor et al. 1996). Therefore, asymmetric cholinergic facilitation is an important mechanism for the retinal direction-selective circuit, which is complementary to asymmetric GABAergic inhibition. Finally, Grzywacz et al. (1998) showed that the asymmetric cholinergic inputs can extend the range of stimuli that can elicit direction-selective responses.

Here we show that the retinal direction-selective system would not substantially benefit—in terms of decoding precision—from implementing a more complex cellular/molecular machinery featuring a direction-selective system with more than four preferred directions. Having more than four cells with different preferred directions would require a more precise wiring of GABAergic and cholinergic inputs.

Having cells with more than four preferred directions would decrease ϵ_{max} by $\sim 20\%$, which corresponds to $\sim 1^\circ$ improvement in decoding precision (Fig. 8*F*). Such improvement in decoding performance could be not behaviorally relevant enough to implement a more complex coordinate system that features more than four preferred directions.

Next, we asked to what extent jitter in the preferred directions of the DSGCs influences coding performance. We compared ϵ_{mean} and ϵ_{max} of ON-OFF DSGCs with preferred directions angularly spaced at 90° without jitter and with jitter (Fig. 8, *B* and *C*, and Fig. 8, *G* and *H*). We found that the direction-selective systems did not benefit on average (ϵ_{mean} , Fig. 8*G*) from perfectly aligning cells along the four cardinal directions. Only in the case of one quadruplet or two quadruplets of ON-OFF DSGCs (1 quadruplet contains 4 ON-OFF DSGCs with 4 different preferred directions) is a perfect alignment of the preferred directions along the four cardinal directions slightly advantageous, reducing ϵ_{max} by 5–15%. Again, these results indicate that the cellular or molecular burden needed for perfect alignment with the four cardinal directions may not be worth the marginal improvement in coding performance.

DISCUSSION

We recorded spiking activity from populations of ON-OFF DSGCs and analyzed their concerted activity for decoding the direction of motion of moving objects in the visual scene. Furthermore, we tested different decoding strategies (probabilistic decoders, population vector, optimal linear estimator; see

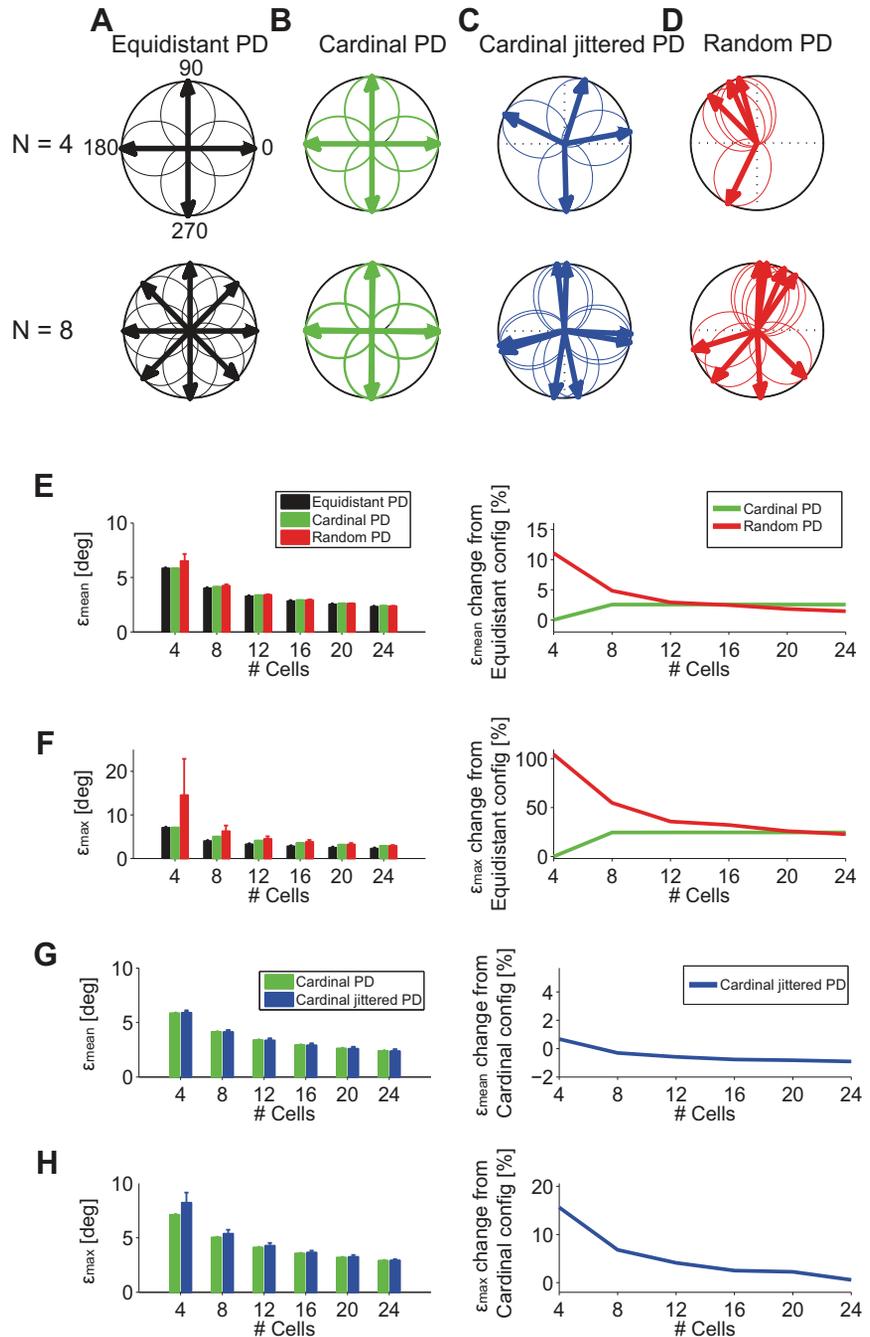


Fig. 8. Influence of angular arrangement of and jitter in preferred directions on decoding performance. *A*, *top*: 4 direction-selective cells with preferred directions equidistantly spaced by 90°. *Bottom*: 8 direction-selective cells equidistantly spaced by 45°, “Equidistant PD.” *B*, *top*: 4 direction-selective cells with preferred directions equidistantly spaced by 90° and aligned with the 4 cardinal directions. *Bottom*: 8 direction-selective cells equidistantly spaced by 90° and aligned with the 4 cardinal directions (2 per cardinal direction), “Cardinal PD.” *C*, *top*: 4 direction-selective cells with preferred directions jittered around 4 cardinal directions; jitter was obtained from recorded data (Fig. 2*E*). *Bottom*: 8 direction-selective cells with preferred directions jittered around 4 cardinal directions, “Cardinal Jittered PD.” *D*, *top*: 4 direction-selective cells randomly arranged in the angular space. *Bottom*: 8 direction-selective cells randomly arranged in the angular space, “Random PD.” *E*, *left*: ϵ_{mean} as a function of the number of cells and arrangement (Equidistant PD, Cardinal PD, and Random PD) of preferred directions. *Right*: ϵ_{mean} % change with respect to Equidistant PD configuration ($n = 200$ simulations for every group of cells). *F*, *left*: ϵ_{max} as a function of the number of cells and arrangement (Equidistant PD, Cardinal PD, and Random PD) of preferred directions. *Right*: ϵ_{max} % change with respect to Equidistant PD configuration. *G*, *left*: ϵ_{mean} as a function of the number of cells and arrangement (Cardinal PD and Cardinal Jittered PD) of preferred directions. *Right*: ϵ_{mean} % change with respect to “Cardinal PD” configuration. *H*, *left*: ϵ_{max} as a function of the number of cells and arrangement (Cardinal PD and Cardinal Jittered PD) of preferred directions. *Right*: ϵ_{max} % change with respect to “Cardinal PD” configuration.

MATERIALS AND METHODS) and reported how a change in stimulus parameters (velocity, object size, direction of motion) influenced the decoding of visual stimuli. In addition, based on real data and a theoretical framework, we analyzed how the observed tuning curve widths and the distribution of the preferred directions along the four cardinal directions influenced the coding performance.

Populations of functionally identified retinal ON- and OFF-cells have been recorded from the macaque retina, and it was shown that recording of their concerted activity allows for extraction of more information from the visual scene (Pillow et al. 2008). Furthermore, several studies have used physiologically identified orientation- and direction-selective neurons in the primary visual cortex for decoding directions of moving

objects (Britten et al. 1992; Graf et al. 2011; Montijn et al. 2014).

Another study focused on stimulus parameter separability by recording rabbit ON-OFF DSGCs with patch clamp (Nowak et al. 2011). It was reported that direction, speed, and luminance of the stimulus could be separated by quantifying cell responses in terms of spike count and by using an encoding model based on multiplication of three independent functions. Furthermore, groups of retinal ON-OFF DSGCs have been recorded from, and the obtained spike trains have been analyzed for stimulus-dependent correlation between cells with same, opposite, and orthogonal preferred directions (Amthor et al. 2005). In our study, we simultaneously recorded from populations of ~ 12 ON-OFF DSGCs that featured all four

different preferred directions, of which we used 4 cells with different preferred directions for studying stimulus-dependent decoding of motion directions.

In the context of data on primate primary visual cortex, a linear decoder and a ML decoder yielded comparable performance in angular motion estimation (Graf et al. 2011). We found that in the case of the retina, however, the coding precision was approximately two times higher for a Bayesian decoder compared with a linear decoder, such as population vectors. However, optimizing the parameters in a linear decoder can approximate the Bayesian results quite closely. In a recent study in the mouse (Montijn et al. 2014), it was shown that a probabilistic decoder significantly outperformed linear methods, such as population vectors, when decoding motion directions by visual cortical neurons. This is consistent with our result (Fig. 3, *D–G*) and indicates that the neural code in front-end sensory areas may not be optimized for a simple linear readout scheme. In addition, in the work of Montijn et al. (2014) it was shown that increasing the population size (1 to 80 cells) for the probabilistic decoder quickly saturated the decoding performance after ~35 neurons with a ~60–70% improvement in decoding precision. Similarly, we found that increasing the population size (from 4 to 24 cells) for the probabilistic decoder improved the decoding precision by ~50% (data not shown).

The most informative region of a bell-shaped direction-tuning curve depends on the respective noise levels (Butts and Goldman 2006). In the presence of small noise high-slope regions are the most informative, whereas in the presence of large noise high-firing rate regions (tuning curve peaks) are the most informative (Butts and Goldman 2006). For groups of four direction-tuning curves with preferred directions that are angularly spaced at 90°, it was shown that the most informative regions change as a function of noise levels (Butts and Goldman 2006). In the work of Butts and Goldman, for noise levels reported as “3x” with Fano factor values similar to ours, the most informative tuning curve regions were located at ~45° from the cells’ preferred directions (Butts and Goldman 2006). Similarly, we report that in the retinal direction-selective system the lowest decoding errors occurred at ~45° from the cells’ preferred directions (Fig. 4*I*, bar width = 0.5 mm). However, the result from Butts and Goldman was derived for the cricket cercal sensory system (Miller et al. 1991), where the Fano factor increased with increasing cell activity, whereas we found an inverse relationship in the rabbit retina: the Fano factor was lowest at the peak of the tuning functions (Fig. 3*B*).

In other studies it was shown that heterogeneity in neural responses could be advantageous for decoding (Chelaru and Dragoi 2008; Tripathy et al. 2013). For example, neural population diversity reduced the negative effects of noise correlations on the decoding performance (Shamir 2014). The preferred directions of ON-OFF DSGCs are jittered and therefore are heterogeneous (Fig. 2, *D* and *E*) (Oyster and Barlow 1967). In our simulations with different angular arrangements of the ON-OFF DSGCs’ preferred directions (Fig. 8), we could not find a substantial positive effect by preferred direction heterogeneity across ON-OFF DSGCs pointing to the same cardinal direction.

In mammals, ~20 retinal ganglion cell types convey information about the visual scene to the brain (Levick 1965, 1967; Rodieck and Brening 1983). Direction-selective neurons, with

properties similar to those of retinal ON-OFF DSGCs, have been found in the rabbit lateral geniculate nucleus (LGN) (Levick et al. 1969). It was reported that LGN tuning curves were more sharply tuned than retinal ON-OFF DSGCs tuning curves. If LGN cells recruit information from several ON-OFF DSGCs, they could exhibit sharp tuning curves as a result of computational strategies similar to the decoders used in the present study (Dayan and Abbott 2001; Georgopoulos et al. 1986; Salinas and Abbott 1994).

In the last decade, genetically identified retinal ganglion cells have been molecularly characterized and with them their target areas in the brain. (Dhande et al. 2013; Huberman et al. 2008, 2009; Kay et al. 2011; Kim et al. 2008). Recently, it has been shown that retinal direction-selective neurons in the mouse transmit, through lateral geniculate neurons, information about motion directions to cortical neurons in the primary visual cortex (Cruz-Martin et al. 2014). Moreover, it has been demonstrated that dendritic spines of individual neurons in the mouse visual cortex can be tuned to diverse preferred orientations and that the tuning of a given neuron (i.e., its preferred tuning direction as defined by its spiking output) can be predicted by the prevalent tuning of its dendrites when all the dendritic spine signals are averaged (Chen et al. 2013). Thus it is possible that the inputs of retinal ON-OFF DSGCs are combined in the visual cortex to decode the direction of a moving object. In future investigations of sensory input coding, it will be advantageous to record the simultaneous activity of genetically defined sensory neurons, as reported in this work, and to trace their brain targets in order to possibly understand how the activity of multiple cells is combined to perform the respective computations.

ACKNOWLEDGMENTS

The authors acknowledge Marcel Stimberg (École Normale Supérieure, Paris, France) for the code to do fits with von Mises functions. The Circular Toolbox was provided by Philipp Behrens (Bernstein Center for Computational Neuroscience, Tübingen, Germany). We acknowledge Thomas Lee Russell and Ian Lloyd Jones at ETH Zurich for technical support with animals.

GRANTS

This work was supported by the Advanced ERC Grant “NeuroCMOS” under contract number AdG 267351, the SNSF Sinergia Project CRSII3_141801, and the CNRS through UMR 8550. M. Fiscella acknowledges individual support through Swiss SystemsX interdisciplinary PhD Grant No. 2009_031.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: M.F., F.F., and K.F. conception and design of research; M.F. performed experiments; M.F. and F.F. analyzed data; M.F., F.F., K.F., J.M., and R.A.d.S. interpreted results of experiments; M.F. prepared figures; M.F., F.F., and R.A.d.S. drafted manuscript; M.F., F.F., R.A.d.S., and A.H. edited and revised manuscript; M.F., F.F., B.R., R.A.d.S., and A.H. approved final version of manuscript.

REFERENCES

Amthor FR, Grzywacz NM, Merwine DK. Extra-receptive-field motion facilitation in on-off directionally selective ganglion cells of the rabbit retina. *Vis Neurosci* 13: 303–309, 1996.

- Amthor FR, Oyster CW, Takahashi ES.** Morphology of on-off direction-selective ganglion cells in the rabbit retina. *Brain Res* 298: 187–190, 1984.
- Amthor FR, Takahashi ES, Oyster CW.** Morphologies of rabbit retinal ganglion cells with complex receptive fields. *J Comp Neurol* 280: 97–121, 1989.
- Amthor FR, Tootle JS, Grzywacz NM.** Stimulus-dependent correlated firing in directionally selective retinal ganglion cells. *Vis Neurosci* 22: 769–787, 2005.
- Ariel M, Daw NW.** Pharmacological analysis of directionally sensitive rabbit retinal ganglion cells. *J Physiol* 324: 161–185, 1982.
- Averbeck BB, Latham PE, Pouget A.** Neural correlations, population coding and computation. *Nat Rev Neurosci* 7: 358–366, 2006.
- Barlow HB, Hill RM.** Selective sensitivity to direction of movement in ganglion cells of the rabbit retina. *Science* 139: 412–414, 1963.
- Barlow HB, Levick WR.** The mechanism of directionally selective units in rabbit's retina. *J Physiol* 178: 477–504, 1965.
- Berry MJ 2nd, Brivanlou IH, Jordan TA, Meister M.** Anticipation of moving stimuli by the retina. *Nature* 398: 334–338, 1999.
- Briggman KL, Helmstaedter M, Denk W.** Wiring specificity in the direction-selectivity circuit of the retina. *Nature* 471: 183–188, 2011.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA.** The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.
- Butts DA, Goldman MS.** Tuning curves, neuronal variability, and sensory coding. *PLoS Biol* 4: e92, 2006.
- Chelaru MI, Dragoi V.** Efficient coding in heterogeneous neuronal populations. *Proc Natl Acad Sci USA* 105: 16344–16349, 2008.
- Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS.** Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499: 295–300, 2013.
- Cohen MR, Kohn A.** Measuring and interpreting neuronal correlations. *Nat Neurosci* 14: 811–819, 2011.
- Cruz-Martin A, El-Danaf RN, Osakada F, Sriram B, Dhande OS, Nguyen PL, Callaway EM, Ghosh A, Huberman AD.** A dedicated circuit links direction-selective retinal ganglion cells to the primary visual cortex. *Nature* 507: 358–361, 2014.
- Cruz-Martin A, El-Danaf RN, Osakada F, Nguyen PL, Callaway EM, Ghosh A, Huberman AD.** A “labeled line” linking direction selective circuits in retina to superficial layers of primary visual cortex. *Neuroscience Meeting Planner* 2013: 639.07/EE6, 2013.
- Dayan P, Abbott LF.** *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press, 2001.
- Dhande OS, Estevez ME, Quattrochi LE, El-Danaf RN, Nguyen PL, Berson DM, Huberman AD.** Genetic dissection of retinal inputs to brainstem nuclei controlling image stabilization. *J Neurosci* 33: 17797–17813, 2013.
- Elstrott J, Anishchenko A, Greschner M, Sher A, Litke AM, Chichilnisky EJ, Feller MB.** Direction selectivity in the retina is established independent of visual experience and cholinergic retinal waves. *Neuron* 58: 499–506, 2008.
- Euler T, Detwiler PB, Denk W.** Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature* 418: 845–852, 2002.
- Famiglietti EV.** Dendritic co-stratification of ON and ON-OFF directionally selective ganglion cells with starburst amacrine cells in rabbit retina. *J Comp Neurol* 324: 322–335, 1992.
- Fiscella M, Farrow K, Jones IL, Jackel D, Muller J, Frey U, Bakkum DJ, Hantz P, Roska B, Hierlemann A.** Recording from defined populations of retinal ganglion cells using a high-density CMOS-integrated microelectrode array with real-time switchable electrode selection. *J Neurosci Methods* 211: 103–113, 2012.
- Frechette ES, Sher A, Grivich MI, Petrusca D, Litke AM, Chichilnisky EJ.** Fidelity of the ensemble code for visual motion in primate retina. *J Neurophysiol* 94: 119–135, 2005.
- Frey U, Egert U, Heer F, Hafizovic S, Hierlemann A.** Microelectronic system for high-resolution mapping of extracellular electric fields applied to brain slices. *Biosens Bioelectron* 24: 2191–2198, 2009.
- Fried SI, Munch TA, Werblin FS.** Mechanisms and circuitry underlying directional selectivity in the retina. *Nature* 420: 411–414, 2002.
- Georgopoulos AP, Schwartz AB, Kettner RE.** Neuronal population coding of movement direction. *Science* 233: 1416–1419, 1986.
- Graf AB, Kohn A, Jazayeri M, Movshon JA.** Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14: 239–245, 2011.
- Grzywacz NM, Amthor FR.** Facilitation in ON-OFF directionally selective ganglion cells of the rabbit retina. *J Neurophysiol* 69: 2188–2199, 1993.
- Grzywacz NM, Amthor FR, Merwine DK.** Necessity of acetylcholine for retinal directionally selective responses to drifting gratings in rabbit. *J Physiol* 512: 575–581, 1998.
- Hill DN, Mehta SB, Kleinfeld D.** Quality metrics to accompany spike sorting of extracellular signals. *J Neurosci* 31: 8699–8705, 2011.
- Hubel DH, Wiesel TN.** Receptive fields of single neurones in the cat's striate cortex (republication of *J Physiol* 148: 574–591, 1959). *J Physiol* 587: 2721–2732, 2009.
- Huberman AD, Manu M, Koch SM, Susman MW, Lutz AB, Ullian EM, Baccus SA, Barres BA.** Architecture and activity-mediated refinement of axonal projections from a mosaic of genetically identified retinal ganglion cells. *Neuron* 59: 425–438, 2008.
- Huberman AD, Wei W, Elstrott J, Stafford BK, Feller MB, Barres BA.** Genetic identification of an On-Off direction-selective retinal ganglion cell subtype reveals a layer-specific subcortical map of posterior motion. *Neuron* 62: 327–334, 2009.
- Jensen RJ, DeVoe RD.** Comparisons of directionally selective with other ganglion cells of the turtle retina: intracellular recording and staining. *J Comp Neurol* 217: 271–287, 1983.
- Jeon CJ, Kong JH, Strettoi E, Rockhill R, Stasheff SF, Masland RH.** Pattern of synaptic excitation and inhibition upon direction-selective retinal ganglion cells. *J Comp Neurol* 449: 195–205, 2002.
- Kay JN, De la Huerta I, Kim IJ, Zhang Y, Yamagata M, Chu MW, Meister M, Sanes JR.** Retinal ganglion cells with distinct directional preferences differ in molecular identity, structure, and central projections. *J Neurosci* 31: 7753–7762, 2011.
- Kim IJ, Zhang Y, Yamagata M, Meister M, Sanes JR.** Molecular identification of a retinal cell type that responds to upward motion. *Nature* 452: 478–482, 2008.
- Kühn N, Gollisch T.** Motion encoding in the salamander retina. *European Retina Meeting, Alicante, Spain*, 2013.
- Levick WR.** Receptive fields of rabbit retinal ganglion cells. *Am J Optom Arch Am Acad Optom* 42: 337–343, 1965.
- Levick WR.** Receptive fields and trigger features of ganglion cells in the visual streak of the rabbits retina. *J Physiol* 188: 285–307, 1967.
- Levick WR, Oyster CW, Takahashi E.** Rabbit lateral geniculate nucleus: sharpener of directional information. *Science* 165: 712–714, 1969.
- Maisak MS, Haag J, Ammer G, Serbe E, Meier M, Leonhardt A, Schilling T, Bahl A, Rubin GM, Nern A, Dickson BJ, Reiff DF, Hopp E, Borst A.** A directional tuning map of *Drosophila* elementary motion detectors. *Nature* 500: 212–216, 2013.
- Marshall JH, Kaye AP, Nauhaus I, Callaway EM.** Anterior-posterior direction opponency in the superficial mouse lateral geniculate nucleus. *Neuron* 76: 713–720, 2012.
- Masland RH, Mills JW.** Autoradiographic identification of acetylcholine in the rabbit retina. *J Cell Biol* 83: 159–178, 1979.
- Miller JP, Jacobs GA, Theunissen FE.** Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary interneurons. *J Neurophysiol* 66: 1680–1689, 1991.
- Montijn JS, Vinck M, Pennartz CM.** Population coding in mouse visual cortex: response reliability and dissociability of stimulus tuning and noise correlation. *Front Comput Neurosci* 8: 58, 2014.
- Nikolaou N, Lowe AS, Walker AS, Abbas F, Hunter PR, Thompson ID, Meyer MP.** Parametric functional maps of visual inputs to the tectum. *Neuron* 76: 317–324, 2012.
- Nowak P, Dobbins AC, Gawne TJ, Grzywacz NM, Amthor FR.** Separability of stimulus parameter encoding by on-off directionally selective rabbit retinal ganglion cells. *J Neurophysiol* 105: 2083–2099, 2011.
- Oesch N, Euler T, Taylor WR.** Direction-selective dendritic action potentials in rabbit retina. *Neuron* 47: 739–750, 2005.
- Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC.** Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433: 597–603, 2005.
- Oyster CW, Amthor FR, Takahashi ES.** Dendritic architecture of ON-OFF direction-selective ganglion cells in the rabbit retina. *Vision Res* 33: 579–608, 1993.
- Oyster CW, Barlow HB.** Direction-selective units in rabbit retina: distribution of preferred directions. *Science* 155: 841–842, 1967.
- Oyster CW, Takahashi E, Collewijn H.** Direction-selective retinal ganglion cells and control of optokinetic nystagmus in the rabbit. *Vision Res* 12: 183–193, 1972.

- Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP.** Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454: 995–999, 2008.
- Piscopo DM, El-Danaf RN, Huberman AD, Niell CM.** Diverse visual features encoded in mouse lateral geniculate nucleus. *J Neurosci* 33: 4642–4656, 2013.
- Pouget A, Dayan P, Zemel R.** Information processing with population codes. *Nat Rev Neurosci* 1: 125–132, 2000.
- Pu ML, Amthor FR.** Dendritic morphologies of retinal ganglion cells projecting to the lateral geniculate nucleus in the rabbit. *J Comp Neurol* 302: 675–693, 1990.
- Rodieck RW, Brening RK.** Retinal ganglion cells: properties, types, genera, pathways and trans-species comparisons. *Brain Behav Evol* 23: 121–164, 1983.
- Salinas E, Abbott LF.** Vector reconstruction from firing rates. *J Comput Neurosci* 1: 89–107, 1994.
- Schwartz G, Taylor S, Fisher C, Harris R, Berry MJ 2nd.** Synchronized firing among retinal ganglion cells signals motion reversal. *Neuron* 55: 958–969, 2007.
- Seung HS, Sompolinsky H.** Simple models for reading neuronal population codes. *Proc Natl Acad Sci USA* 90: 10749–10753, 1993.
- Shamir M.** Emerging principles of population coding: in search for the neural code. *Curr Opin Neurobiol* 25C: 140–148, 2014.
- Stanley GB, Jin J, Wang Y, Desbordes G, Wang Q, Black MJ, Alonso JM.** Visual orientation and directional selectivity through thalamic synchrony. *J Neurosci* 32: 9073–9088, 2012.
- Swindale NV.** Orientation tuning curves: empirical description and estimation of parameters. *Biol Cybern* 78: 45–56, 1998.
- Tauchi M, Masland RH.** The shape and arrangement of the cholinergic neurons in the rabbit retina. *Proc R Soc Lond B Biol Sci* 223: 101–119, 1984.
- Taylor WR, Vaney DI.** Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *J Neurosci* 22: 7712–7720, 2002.
- Theunissen FE, Miller JP.** Representation of sensory information in the cricket cercal sensory system. II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *J Neurophysiol* 66: 1690–1703, 1991.
- Tripathy SJ, Padmanabhan K, Gerkin RC, Urban NN.** Intermediate intrinsic diversity enhances neural population coding. *Proc Natl Acad Sci USA* 110: 8248–8253, 2013.
- Vaney DI, Pow DV.** The dendritic architecture of the cholinergic plexus in the rabbit retina: selective labeling by glycine accumulation in the presence of sarcosine. *J Comp Neurol* 421: 1–13, 2000.
- Weng S, Sun W, He S.** Identification of ON-OFF direction-selective ganglion cells in the mouse retina. *J Physiol* 562: 915–923, 2005.
- Wyatt HJ, Daw NW.** Directionally sensitive ganglion cells in the rabbit retina: specificity for stimulus direction, size, and speed. *J Neurophysiol* 38: 613–626, 1975.
- Yang G, Masland RH.** Direct visualization of the dendritic and receptive fields of directionally selective retinal ganglion cells. *Science* 258: 1949–1952, 1992.
- Yang G, Masland RH.** Receptive fields and dendritic structure of directionally selective retinal ganglion cells. *J Neurosci* 14: 5267–5280, 1994.
- Yonehara K, Farrow K, Ghanem A, Hillier D, Balint K, Teixeira M, Jüttner J, Noda M, Neve RL, Conzelmann KK, Roska B.** The first stage of cardinal direction selectivity is localized to the dendrites of retinal ganglion cells. *Neuron* 79: 1078–1085, 2013.
- Yoshida K, Watanabe D, Ishikane H, Tachibana M, Pastan I, Nakanishi S.** A key role of starburst amacrine cells in originating retinal directional selectivity and optokinetic eye movement. *Neuron* 30: 771–780, 2001.

