Suppression without inhibition: how retinal computation contributes to saccadic suppression

Saad Idrees1,2,10, Matthias-Philipp Baumann1,3, Maria M. Korympidou1,2,4, Timm Schubert1,4, Alexandra Kling5, Katrin Franke4,6, Ziad M. Hafed1,3, Felix Franke7,8,9✉ & Thomas A. Münch1,4✉

Visual perception remains stable across saccadic eye movements, despite the concurrent strongly disruptive visual flow. This stability is partially associated with a reduction in visual sensitivity, known as saccadic suppression, which already starts in the retina with reduced ganglion cell sensitivity. However, the retinal circuit mechanisms giving rise to such suppression remain unknown. Here, we describe these mechanisms using electrophysiology in mouse, pig, and macaque retina, 2-photon calcium imaging, computational modeling, and human psychophysics. We find that sequential stimuli, like those that naturally occur during saccades, trigger three independent suppressive mechanisms in the retina. The main mechanism is triggered by contrast-reversing sequential stimuli and originates within the receptive field center of ganglion cells. It does not involve inhibition or other known suppressive mechanisms like saturation or adaptation. Instead, it relies on temporal filtering of the inherently slow response of cone photoreceptors coupled with downstream non-linearities. Two further mechanisms of suppression are present predominantly in ON ganglion cells and originate in the receptive field surround, highlighting another disparity between ON and OFF ganglion cells. The mechanisms uncovered here likely play a role in shaping the retinal output following eye movements and other natural viewing conditions where sequential stimulation is ubiquitous.
Vision appears as a continuous and coherent process. This is a striking achievement of the visual system, considering that the visual flow across the retina is not continuous, but governed by frequent and sudden changes, irregularities, and disruptions. As a consequence of this active vision, or the process of active exploration of the visual environment, the meaningful images falling onto the retina are only brief snapshots of the world, interrupted by blinks and rapid motion. The most prominent cause of such disruptions are eye movements. Saccades, for example, are critical for efficiently sampling the visual world, interrupted by blinks and rapid motion. Images falling onto the retina are only brief snapshots of the world, interrupted by blinks and rapid motion. The most prominent cause of such disruptions are eye movements. Saccades, for example, are critical for efficiently sampling the visual world, interrupted by blinks and rapid motion. In this study, we describe such a mechanism. We experimentally mimicked the visual flow resulting from saccades and recorded the neural activity of the output neurons of the retina (RGCs) from ex vivo retinas of mice, pigs, and macaque monkeys. We found that retinal saccadic suppression was the result of multiple mechanisms, the most significant of which was a specific visual processing motif within an RGC’s receptive field center. This motif, which we call dynamic reversal suppression, did not depend on any inhibitory signals; it resulted from temporal filtering of inherently slow cone photoreceptor responses coupled with nonlinearities in the downstream retina pathways. Two further components of suppression originated from beyond the RGC’s receptive field center, only one of them driven by GABAergic inhibition. Interestingly, these two additional components were observed primarily in ON RGCs, highlighting yet another disparity between ON and OFF type RGCs. Perhaps one of the most intriguing outcomes of this study, also consistent with observations of perception, is that the suppressive effects observed in RGCs were not exclusively triggered by saccades, but occurred for many scenarios involving sequential visual stimulation, which are ever-present during natural vision. Therefore, while the results described here are crucial for understanding the mechanisms of saccadic suppression, they also elucidate more general mechanisms of retinal signal processing across any time-varying visual input over short time scales (10–1000 ms).

**Results**

**Experimental approach.** We measured the modulation of retinal ganglion cell (RGC) output following saccade-like changes of the visual input with a variety of different light stimulation strategies (Fig. 1a and Supplementary Fig. 1). The basic experimental paradigm was similar to that described in1. Briefly, we recorded spiking activity of RGCs in isolated ex vivo mouse retinae using both high-density and low-density multielectrode arrays (MEAs). Each retina was exposed to a background texture having one of several possible spatial scales that defined its spatial spectrum (fine to coarse, “Methods”, Supplementary Fig. 2). We simulated saccade-like image displacements by rapidly translating the texture globally across the retina (“Methods”; Fig. 1a). Most RGCs responded robustly to such saccade-like texture displacements (see Fig. 1b for responses of example ON and OFF RGCs). At different times relative to the saccade-like texture displacements (saccades from now on), we presented a brief probe flash (Fig. 1c). We then analyzed how the response (spike rate of the RGC) to this probe flash was influenced by the preceding saccade, by comparing it to the response to the flash presented in isolation (baseline). To quantify RGC response modulation, we calculated a modulation index (“Methods”) which quantified how much a cell’s flash response was modulated by a temporally close saccade. We first isolated the flash-induced response component by subtracting the saccade-only response (e.g., Fig. 1b) from the response to the composite saccade-flash stimulus (e.g., Fig. 1c). Based on this flash-induced response component (Fig. 1d), we calculated the modulation index as \( (r_d - r_b)/(r_d + r_b) \). Here, \( r_d \) is the peak response to the probe flash presented with a delay \( d \) relative to saccade onset, and \( r_b \) is the baseline (peak response to the flash presented ~2 s after the saccade). This modulation index is negative when flash-induced responses are suppressed (Fig. 1d shows, on the horizontal dashed line, the example cells’ modulation indices for the responses at each flash time). In yet further recordings, we applied various manipulations to this base paradigm to probe for the mechanisms underlying modulation of RGC responses following saccades. To generalize our findings across other species, we also performed similar analyses of pig and macaque RGC data.

**Similarities and differences in retinal saccadic suppression across ON and OFF type RGCs.** Suppression was robust across
most RGCs that we recorded from, consistent with what we reported previously. Here, we more closely inspected functionally different RGCs. Specifically, throughout this study, we divided RGCs into ON and OFF types (i.e., into RGCs responding best to light increments or decrements, respectively; Methods). Unless otherwise noted, we always quantified the modulation index defined above for ON RGCs based on their responses to bright probe flashes and for OFF RGCs based on their responses to dark probe flashes (Fig. 1c, d). Flash responses following a saccade were suppressed in both ON and OFF RGCs, as seen in Fig. 1d for two example cells. Figure 1e shows the temporal profile of the mean population modulation index for ON and OFF cells, and Supplementary Fig. 3 the underlying population data. Suppression was consistently stronger for coarser background textures (Fig. 1e and Supplementary Fig. 3), for both ON and OFF RGCs. This is consistent with, where we showed that this
Fig. 1 Similarities and differences in retinal saccadic suppression across ON and OFF RGCs. a RGC action potentials were recorded from ex vivo retinae placed on multielectrode arrays. Saccades were mimicked by displacing a texture projected onto the retina (blue arrows in the left panel indicate texture-displacement paths). The texture remained static for 2 s and was then displaced over 100 ms (blue outlines) followed by a brief probe flash (here, a bright probe flash is depicted, orange outline). Each trial consisted of 39 successive saccade-flash sequences (Supplementary Fig. 1a). b, c Average activity (firing rate) of an example ON RGC (left column) and OFF RGC (right column) to 39 saccade sequences not followed by a probe flash (b), and to 39 saccade sequences followed by probe flashes at different delays after saccade onset (c). Blue window: timing of saccades; orange markers: timing of probe flashes. d Isolated flash-induced responses (firing rate) of the same RGCs obtained by subtracting responses to saccades-alone (b) from responses to saccades followed by probe flashes (c). Lines connecting the response peaks highlight the time courses of retinal saccadic suppression relative to baseline flash-induced responses. Numbers above each response peak represent the modulation index which quantifies how much the probe flash response is modulated by the preceding saccade (“Methods”, negative modulation indices correspond to suppressed flash-induced responses). e Population modulation index (mean ± s.e.m.) of ON (light gray) and OFF (dark gray) RGCs, for different background textures with different spatial scales (left to right: fine to coarse). The number of ON and OFF RGCs in the population varied between 68 and 574 for different flash times and textures (see Supplementary Fig. 3 for exact numbers and relevant statistics). Hash symbols: significant modulation difference between ON and OFF RGCs (P < 10^{-4}, two-tailed Wilcoxon rank-sum test).

dependency on the texture can be explained by the distinct statistics of luminance and contrast changes when coarse or fine textures move across the RGCs’ receptive fields. However, a striking difference existed in suppression recovery times: OFF RGCs on average recovered by ~350 ms after saccade onset, whereas ON RGCs fully recovered only by ~1 s. Similar results were obtained under scotopic conditions for coarse textures, while suppression for fine textures was very weak (Supplementary Fig. 4; all other mouse retina data were recorded at mesopic conditions). In general, the presence of post-saccadic suppression of probe flash responses in both ON and OFF type RGCs suggests a common mechanistic theme across these cell types. On the other hand, the different recovery times indicate either additional suppressive mechanisms in ON RGCs or additional recovery mechanisms in OFF RGCs.

Spatial origin of retinal saccadic suppression

Global component of suppression. To probe the mechanisms underlying suppression and its differences across ON and OFF type RGCs, we first examined the spatial origin of suppression. We hypothesized that suppression of flash responses was caused by circuits detecting rapid global shifts across the retina. Typically, these circuits include a lateral network of interneurons, communicating with RGCs even from beyond their classical center-surround receptive field (i.e., from their periphery, or far surround). To test whether suppression was caused by such circuits, we modified the spatial layout of the paradigm: we placed a square mask of 1000 × 1000 μm² (Fig. 2a, right) to restrict the saccades to the periphery of an RGC’s receptive field. Similar to the previous experiments, the probe flash was either a dark or bright flash presented over the entire retina, including the masked region. Figure 2b shows the mean population modulation indices of ON RGCs (top) and OFF RGCs (bottom) from these experiments (Supplementary Fig. 5 depicts the underlying population data and shows responses of representative ON and OFF RGCs from these experiments). In OFF RGCs, responses to full-field probe flashes were no longer suppressed when saccades were restricted to the periphery. The responses of ON RGCs, on the other hand, were still suppressed in this condition. The resulting suppression was however weaker and shorter-lived (recovered by 350 ms) than with full-field saccades. These observations (Fig. 2b and Supplementary Fig. 5c) were robust across ON and OFF RGCs whose receptive fields were completely contained within the masked region (Supplementary Fig. 5b).

We will refer to this component of suppression in ON RGCs, which originates from the periphery, as the global component from now on. Such spatially far-reaching inhibition is often mediated through GABAergic wide-field amacrine cells. We tested this hypothesis by blocking GABA_A receptors. Indeed, in the peripheral saccade condition, the modulation index for most ON RGCs was around 0 in the presence of the GABA_A receptor antagonist SR-95531 (Fig. 2b and Supplementary Fig. 5d). These results suggest that this short-lived global component of suppression is caused by inhibition via GABAergic amacrine cells, perhaps similar to the polyaxonal amacrine cells described previously. Thus, while suppression is indeed partially caused by circuits detecting global changes across the retina, those circuits seem to act predominantly on ON RGCs, and even there, they only account for a fraction of the total suppression observed with full-field saccades (without mask), which lasts longer. Other, probably more local sources of suppression must exist that account for most of the suppression in ON RGCs and all of the suppression in OFF RGCs.

Local components of suppression. To understand the more local components of suppression, we used different analyses and manipulations of the main experimental paradigm. As we will see below, the more local components can be subdivided into a central and a surround component. First, we eliminated the global component, by repeating our normal full-field saccade paradigm in the presence of GABA receptor blockers. The suppression profile of both ON and OFF RGCs was only weakly affected upon blocking GABA_A,C receptors (5 μM SR-95531 and 100 μM Picrotoxin; Fig. 2c and Supplementary Fig. 6a). Since the GABA-block eliminates the global component of suppression, the remaining more local components did not seem to rely on GABAergic inhibition. Also, this suggests that the local components dominate retinal saccadic suppression under full-field conditions. We then also blocked glycine receptors (1 μM of Strychnine; Fig. 2d and Supplementary Fig. 6b) to test if the local components of suppression were caused by local inhibition via glycinergic pathways. Here again, the suppression profiles of both ON and OFF RGCs were only weakly affected upon blocking glycine receptors in combination with blocking GABA_A,C receptors. Therefore, inhibitory synaptic interactions are not the major mechanism behind the local components, which dominate suppression of RGCs.

Next, we tested whether these local components originated from within the receptive field center. For this, we modified the spatial layout of our paradigm to exclude saccades from the very center of the receptive field. Simply reducing the size of our mask would have severely decreased the number of simultaneously recorded cells located inside the mask, and we therefore resorted to a different strategy: saccades and flashes were presented in small square regions spread across the retina, separated by gaps kept at mean luminance (checkerboard mask, Fig. 2e). In one condition (Fig. 2e, left), we presented saccades and flashes in all
Fig. 2 Spatial origins of retinal saccadic suppression. a Spatial layout of the visual stimulation paradigm used in experiments to probe the global component of suppression. Saccades were presented either full-field (left; same as in Fig. 1) or in the periphery (right), where a 1000 × 1000 μm² mask (intensity: mean luminance of texture) covered at least 2-σ of the 2D Gaussian fit to the RGC receptive fields (Supplementary Fig. 5b). b Population modulation index (mean ± s.e.m.) of ON (top) and OFF (bottom) RGCs for full-field saccades condition (thick gray lines, same as Fig. 1e rightmost panel; N = 68 to 574 RGCs (see Supplementary Fig. 3 for exact numbers)); periphery saccades condition (thin gray lines; N = 91 ON RGCs, N = 56 OFF RGCs); and periphery saccades condition in the presence of GABA(A) receptor blocker (5 μM SR-95531; green lines; N = 62 ON RGCs, N = 35 OFF RGCs). Blue window shows the timing of the saccade. In these experiments, we used a coarse background texture (300 μm spatial scale). Timing of probe flashes: 50 and 150 (only for full-field saccade), 177, 200, 350, 600, and 2100 ms (baseline) after saccade onset. c, d Population modulation index (mean ± s.e.m.) of ON (top) and OFF (bottom) RGCs for full-field saccades without any pharmacological agents (gray lines; N = 82 ON RGCs, N = 30 OFF RGCs) and with GABA(A,C) receptor blockers 5 μM SR-95531 + 100 μM Picrotoxin (c; green lines), and for a subset of RGCs where we additionally blocked glycine receptors using 1 μM Strychnine (d; green lines; N = 51 ON RGCs, N = 13 OFF RGCs). In these experiments, we used a coarse background texture (150 μm spatial scale). Probe flashes were presented at 117, 150, 200, 350, 600, 1100, and 2100 ms (baseline) after saccade onset. e, f Spatial layout of the visual stimulation paradigm used in experiments to probe local components of suppression. Saccades and flashes were presented in 100 × 100 μm² square regions, separated by 100 μm gaps with mean overall luminance. Left: Saccades and flashes were presented in all regions. Right: Saccades and flashes were presented in alternate regions; only cells with receptive fields (RFs) in the non-saccade regions (orange) were analyzed (black ellipse: 1-σ of the 2D Gaussian fit to an example RGC receptive field). Consequently, saccades were excluded from at most ~300 × 300 μm² of a cell’s RF center. In these experiments, we used a coarse background texture (150 μm spatial scale). f Population modulation index (mean ± s.e.m.) of ON (top; N = 32) and OFF (bottom; N = 38) RGCs for saccades and flashes in all regions (thick lines) or saccades excluded from RF center (thin lines). Red arrow indicates significant loss in suppression in ON RGCs for early flashes at 117 and 150 ms upon excluding saccades from RF center (P = 0.0016 and P = 0.002, respectively; two-tailed Wilcoxon rank-sum test). In all panels, hash symbols indicate statistically significant difference between groups (P < 0.01, two-tailed Wilcoxon rank-sum test).

regions of the checkerboard mask; in the other condition (Fig. 2e, right), saccades and flashes were presented in alternate regions. With this second arrangement, saccades were excluded from at most ~300 × 300 μm² of a cell’s receptive field center, even if that cell was perfectly centered on a non-saccade region. Flashes were presented in the set of regions that included the square region covering the receptive field center of the analyzed RGC (Supplementary Fig. 7a).

Probe flash responses following saccades were suppressed in both ON and OFF RGCs when the saccade and flash were presented in all regions (Fig. 2f, thick lines; Supplementary Fig. 7c; see Supplementary Fig. 7b, e.g., cells), consistent with the suppression observed after full-field saccades (Fig. 1 and Supplementary Fig. 3). When saccades were excluded from the receptive field center, and were presented in alternate regions to the flash, the flash responses were no longer suppressed in OFF
RGCs (Fig. 2f, bottom, thin line), even though these cells showed spiking responses to saccades themselves (Supplementary Fig. 7b). In fact, flash responses were even enhanced. This suggests that the local component of suppression in OFF RGCs arises fully from within the receptive field center (central component). This highly localized origin of suppression in OFF RGCs was further confirmed by additional analysis of the large mask experiments (see Supplementary Fig. 7d, e). In ON RGCs, on the other hand, suppression persisted (Fig. 2f, top and Supplementary Fig. 7c), even though a loss in suppression was apparent for flashes presented immediately after the saccade, at 117 and 150 ms (marked with an arrow in Fig. 2f). This suggests that in ON RGCs, part of the early suppression originates from the central component. The leftover suppression during these early time points might be explained by the global component of suppression, described above (Fig. 2b and Supplementary Fig. 5c), which should also be triggered under this experimental setting. However, since the global component also recovers quickly (by 350 ms, Fig. 2b and Supplementary Fig. 5c), the persisting suppression at the later time points (350 ms and beyond) needs to originate from yet another source beyond the receptive field center. We call this the surround component, and it may originate from the ON RGCs’ immediate surround, which also experiences the saccade under this experimental setting. Therefore, in ON RGCs, the local component of suppression can be divided into a central and a surround spatial component.

**Summary of retinal saccadic suppression spatial origins.** In summary, our data suggest that retinal saccadic suppression is mediated by at least three components with distinct spatial origins and temporal properties (Fig. 3): a central, surround, and global component. Suppression in OFF RGCs is mediated exclusively by the central component, which originates from the cell’s receptive field center and is characterized by fast onset and fast recovery (by 350 ms after saccade onset). In ON RGCs, we most directly observed the global component (Fig. 2b and Supplementary Fig. 5a). It extends into the periphery and its timing is similar to the central component in OFF RGCs. Only this global component is affected by blocking GABA receptors (Fig. 2b and Supplementary Fig. 5). During full-field saccades, removing this component by blocking GABA receptors has little effect on the overall suppression (Fig. 2c, d and Supplementary Fig. 6), suggesting a more dominant role of the remaining components. The central component in ON RGCs can only be observed by the loss in suppression for early flashes when saccades are excluded from the receptive field center (marked with an arrow in Fig. 2f, top). Its full duration and time course are obscured by the concurrently acting global and surround components. However, given the identical pharmacological dependencies and spatial origins, it is plausible that the central component is symmetric across ON and OFF RGCs with a common underlying mechanism. Therefore, the longer suppression in ON RGCs can neither be attributed to the central nor global components. It likely originates from the immediate surround of the receptive field. This surround component is long-lasting (reverses by ~1 s) and has a slow onset (Fig. 2f and Supplementary Fig. 7).

**Suppression is triggered by the interaction between consecutive stimuli of opposite polarity.** We previously showed that retinal and perceptual saccadic suppression not only occur after texture displacements, but also after instantaneous texture jumps and structure-free uniform luminance steps. These observations suggested that saccadic suppression is the consequence of rather general mechanisms in which the response to a second stimulus (here: probe flash) gets suppressed by a previous visual transient (caused by saccades or luminance steps). In the following, we apply additional analysis to the luminance-step

---

**Fig. 3 Schematic summarizing the spatial origins of retinal saccadic suppression.** Saccadic suppression in OFF RGCs (right) is mediated primarily by the central component of suppression (blue) that originates from the cell’s receptive field center. ON RGCs (left) get suppressed from two additional components: First, the fast but short-lived global component (red), mediated by GABAergic inhibition, that originates from as far as the cell’s periphery. This global component has a similar temporal profile as the central component. However, it is weaker than the central component and acts in parallel to it, indicated by the red arrows parallel to blue arrows in the total suppression schematic. Second, the delayed but long-lasting surround component (green), which might originate from the cell’s immediate surround. The central component and surround component do not depend on classical GABAergic or glycineergic inhibitory pathways. The differences in the suppression recovery time in ON and OFF RGCs was mainly due to this surround component acting on ON RGCs. Inset shows the legend for arrow schematics. Length of the arrows represent suppression strength; spread of the arrows show the temporal profile of suppression.
Fig. 4 Suppression following luminance steps. **a** Population modulation index (mean ± s.e.m.) of ON (light gray, N = 259) and OFF (dark gray, N = 107) RGCs for probe flashes following luminance steps (blue line). Modulation index for each RGC was based on its average response to 56 or 156 luminance-step sequences (Supplementary Fig. 1b) spanning a contrast range of −0.5 to +0.5 Michelson contrast (“Methods”). Probe flashes were presented at 17 ms, 33, 50, 100, 250, 500, 1000, and 2000 (baseline) after luminance steps. Probe flash responses were suppressed in both ON and OFF RGCs, with similar time course and recovery as in the saccade paradigm with textures (Fig. 1e). Error bars are not visible due to small s.e.m. **b** Same as in (**a**), except that the modulation index for each RGC was separately based on average responses to probe flashes after positive-contrast luminance steps (left panel; +0.03 to +0.5 Michelson contrast), and after negative-contrast luminance steps (right panel; −0.03 to −0.5 Michelson contrast). Underlying population data are shown in Supplementary Fig. 8. **c** Same as in (**b**), for a subset of ON RGCs (N = 115) in control conditions (light-gray lines) and with GABAA,C and glycine receptors blocked (green lines; cocktail of 5 μM SR-95531, 100 μM Picrotoxin and 1 μM Strychnine). Hash symbol: significant difference between modulation of ON and OFF RGCs in **a** or between ON RGCs without and with pharmacological blockers in (**c**) (P < 10⁻⁴, two-tailed Wilcoxon rank-sum test).

Paradigm (Supplementary Fig. 1b) dataset of ref. 1, to investigate how the polarity and strength of the visual transients affect the suppression of ON and OFF RGCs.

Similar to all previous experiments, we analyzed the modulation index of ON and OFF RGCs separately, using bright probe flashes to analyze ON RGCs and dark probe flashes for OFF RGCs. Consistent with the suppression after texture displacements (Fig. 1e and Supplementary Fig. 3), responses to flashes after luminance steps were strongly suppressed in both ON and OFF RGCs, and ON RGC suppression outlasted suppression in OFF RGCs (Fig. 4a and Supplementary Fig. 8a). The two seemingly different experimental paradigms may therefore trigger similar mechanisms in the retina.

We hypothesized that the response to a luminance step might strongly activate RGCs, so that the response to a subsequent probe flash would drive the cells into adaptation or saturation, effectively resulting in suppressed flash responses. At least for the local components of suppression, this could be a viable mechanism as suppression is not caused by inhibitory synaptic interactions. If this was indeed the case, then positive-contrast luminance steps would suppress responses to bright flashes in ON RGCs, and negative-contrast luminance steps would suppress responses to dark flashes in OFF RGCs. To test this, we separately analyzed the effects of positive- and negative-contrast luminance steps on probe flash responses (Fig. 4b). Surprisingly, the resulting effects were contrary to our adaptation/saturation hypothesis: the responses of ON RGCs to bright probe flashes were only weakly suppressed after positive-contrast luminance steps (Fig. 4b, left), but strongly suppressed following negative-contrast luminance steps (Fig. 4b, right). Similarly, responses of OFF RGCs to dark probe flashes were weakly suppressed by negative-contrast luminance steps (Fig. 4b, right), but strongly suppressed by positive-contrast luminance steps (Fig. 4b, left). Supplementary Fig. 8 shows the underlying population data for these experiments, and Supplementary Figs. 9a, b shows the spiking response of a representative ON and OFF RGC, respectively. While ON RGCs did show a small component of suppression in support of our adaptation/saturation hypothesis (Fig. 4b left panel, see Supplementary Fig. 9a left column and Supplementary Fig. 10a for a detailed analysis), the dominant suppressive effect in both ON and OFF RGCs was caused by luminance steps with the opposite contrast as the subsequent flash.

Such crossover style of suppression would be consistent with mechanisms involving crossover inhibition via amacrine cells42, where activation of OFF pathways (here: by the negative-contrast luminance step) would inhibit responses in the ON pathway (here: to the bright probe flash) and vice versa. However, consistent with our earlier experiments (Fig. 2c, d), suppression in ON RGCs still persisted upon blocking GABAA,C and glycine receptors (5 μM SR-95531, 100 μM Picrotoxin, and 1 μM Strychnine) (Fig. 4c). We could not calculate a modulation index for OFF RGCs under these conditions because they did not respond to brief probe flashes in the presence of the pharmacological agents, and therefore the modulation index was mathematically undefined. However, in our texture displacement experiments, the same pharmacological agents (Fig. 2d) had no substantial effect on OFF RGC suppression. The crossover-style
suppression observed in Fig. 4b was therefore unlikely to be caused by classical crossover inhibition pathways involving amacrine cells and GABA_A,C or glycine receptors.

Central component of suppression results from cone response kinetics and nonlinearities in downstream retinal pathways. Taken together, our experiments so far suggest that suppression in OFF RGCs (1) is mediated solely by the central component of suppression that originates in the receptive field center (Fig. 2e and Supplementary Fig. 7b–e), (2) is predominantly triggered by the interaction between consecutive stimuli with opposite polarity (Fig. 4b and Supplementary Fig. 8b), and (3) is not caused by inhibitory amacrine cells (Fig. 2c, d and Supplementary Fig. 6). Similar conclusions can be drawn for the central component of suppression in ON RGCs. We term this suppressive retinal processing motif which does not rely on inhibition dynamic reversal suppression.

The highly localized origins of dynamic reversal suppression, and its lack of dependence on inhibition, restrict the possible cellular substrates for this motif to the feed-forward pathway in the retina, namely photoreceptors—bipolar cells—RGCs.

We wondered whether opposite-polarity stimulus–stimulus interactions could already modulate the responses of photoreceptors themselves. For this, we recorded the output of cones with an intensity-based glutamate-sensitive fluorescent reporter (iGluSnFR)33,43, predominantly expressed in horizontal cells postsynaptic to cone terminals (Methods). We presented a shortened version of the luminance-step paradigm in which a homogeneous background alternated between a brighter and darker gray value (Supplementary Fig. 1c) to induce positive-contrast (+0.4 Michelson contrast) and negative-contrast (~0.4 Michelson contrast) luminance steps. Dark or bright probe flashes (100 ms duration, ~0.33 or ~0.33 Michelson contrast, respectively) followed the luminance steps at different delays (50, 250, and 2000 ms), with the flash at 2000 ms serving as baseline.

The luminance steps caused sustained changes in the cones’ glutamate output (Fig. 5a). The transient responses to the probe flashes were superimposed on these glutamate modulations (Fig. 5b). This superposition was mostly linear and did not indicate nonlinear effects such as adaptation or saturation. Therefore, when we isolated the flash responses by subtracting the step responses, the resulting peak flash responses were only weakly affected by the preceding luminance step (Fig. 5c). Thus, at the level of the cone output (Fig. 5b), there was hardly any suppression when only considering the peak of the probe flash responses (Fig. 5c). How does the suppression observed at the level of RGC output arise from effectively linear cone responses?

### Fig. 5 Cone output in response to probe flashes following luminance steps.

#### a. Response to luminance steps alone

- **Stimulus:** Luminance step: Positive-contrast
- **Time:** 1 s
- **Cone output (a.u.):**
- **Negative-contrast**

#### b. Responses to luminance steps followed by probe flashes

- **Stimulus:** Positive-contrast
  - (Bright flashes)
  - Negative-contrast
  - (Bright flashes)
- **Time:** 1 s
- **Flash response peak**
- **Cone output (a.u.):**

#### c. Flash-induced responses

- **Time:** 1 s
- **Cone output (a.u.):**
The answer must lie in other properties of the cone response, such as its kinetics, which will be captured by downstream retinal pathways.

To demonstrate this, we used a published computational model\textsuperscript{15} that describes RGC spiking responses as a function of the light stimulus. In the model, a light stimulus is passed to model photoreceptors, feeding their output through a set of linear–nonlinear filters which reflect the processing by different bipolar pathways ("Methods"), and eventually converting these filter outputs into RGC spiking. Apart from the model component that captures the cone response and which is described by a differential equation, the model is a simple linear–nonlinear cascade model. Here, as the first step of analysis, we fitted the parameters of the cone component of the model to reflect our measured data of cone output. The model cone faithfully explained the observed cone responses (Supplementary Fig. 11) and gave us the opportunity to calculate cone responses to flashes at additional time points not measured in the experiments. This modeled cone output to step-flash combinations was fed into the model bipolar cells, finally yielding model RGC responses (Fig. 6).

In the model, different RGC types can be described by varying the bipolar cell filter properties. We first investigated transient model RGC responses (Fig. 6a–c) and calculated a modulation index (Fig. 6d) comparable with the modulation index of our real RGC data. As a control, we also passed the raw cone output data, instead of the fitted cone model, to the model bipolar cells and found qualitatively the same results (Supplementary Fig. 12).

The model’s ON and OFF RGCs (Fig. 6d) showed crossover-style suppression that was consistent with the suppression of real RGCs under similar luminance-step experiments (Fig. 4b): the model’s ON RGC showed strong suppression to the bright flashes presented immediately after the negative-contrast luminance step (Fig. 6d, bottom), while bright flashes after the positive-contrast luminance step were only weakly affected (Fig. 6d, top). Suppression recovered by 200 ms, consistent with the recovery time for the central component of suppression in real RGCs. Similarly, the model’s OFF RGC showed strong suppression of its response to dark flashes presented immediately after the positive-contrast luminance step (Fig. 6d, top); suppression was absent in OFF RGC when the dark flash was presented later or after a negative-contrast luminance step (Fig. 6d). The markers on the curves in Fig. 6d correspond to the time points when the flashes were presented to the cones in the experiments depicted in Fig. 5. Model RGC responses to step-flash combinations and flash-induced responses at these time points are shown in Fig. 6a–c. In short, the model could replicate dynamic reversal suppression observed in our real RGCs dataset (Fig. 4b).

What properties of the model led to the emergence of the suppressive effect in RGC responses, despite the mostly linear response superposition at the cone output? In the model, the bipolar cells have transient filter properties and are driven predominantly by the instantaneous rate of change in the cone output (i.e., its derivative, rather than by the absolute cone output; Fig. 6e). The response to a probe flash presented immediately after an opposite-contrast luminance step (50 ms) occurred during the initial phase (ramp) of the cone response to the luminance step (Fig. 5b, columns 2 and 3). This causes a much smaller rate of change in cone output than a flash presented after the cone response to a luminance step has already reached its steady-state value (compare flash response at different times in Fig. 5b). This smaller rate of change in cone output therefore drives the downstream bipolar cells only weakly, which then, together with the threshold nonlinearity, results in weak or even completely suppressed model RGC responses (Fig. 6c, columns 2 and 3). In other words, a non-preferred luminance step can hyperpolarize the retinal pathway for a brief duration. A subsequent flash of preferred contrast presented within that brief duration is therefore less effective because it first needs to reach the threshold. On the other hand, flashes presented during the steady-state phase of the luminance step response (250 ms and 2000 ms in Fig. 5b, columns 2 and 3), or flashes presented immediately after (50 ms) a same-contrast luminance step (Fig. 5b, columns 1 and 4) caused larger instantaneous changes in the cone output, and therefore resulted in relatively stronger spiking (250 ms and 2000 ms in Fig. 6c; 50 ms in Fig. 6c, columns 1 and 4). In conclusion, if two stimuli of opposite contrast occur closely together, an interplay of temporal filtering (emphasizing instantaneous changes) and nonlinear thresholding results in dynamic reversal suppression.

This would suggest that RGCs with different temporal properties (e.g., different transiency) may experience different degrees of dynamic reversal suppression. To test this, we reanalyzed our recorded RGC data of Fig. 4 to quantify suppression as a function of RGC response transiency. We found that suppression did indeed vary with RGC transiency: it was weaker for less transient RGCs, as indicated by the negative slope of the line fit to modulation indices in Fig. 6f; seen in the ON RGC suppression after negative-contrast steps (blue lines in row 2) and OFF RGC suppression after positive-contrast steps (red lines in row 1); columns represent different time points. To explain the origins of this effect, we resorted again to our computational model. Here, a simple change in the temporal filter kinetics could replicate the effect (Fig. 6g): making the filter less transient (Supplementary Fig. 13) led to weaker suppression in the model RGCs. Similarly, the effect could be replicated not by adjusting the linear filter properties, but by making the pathway’s thresholding nonlinearity (Eq. (11) in “Methods”, Supplementary Fig. 13) stronger (Fig. 6h), consistent with the stricter nonlinearities found in more transient RGC pathways\textsuperscript{46}.

The analyses in this section suggest that dynamic reversal suppression, the motif underlying the central component of retinal saccadic suppression, does not have a single site of origin. Instead, it appears to emerge from the temporal filtering of the relatively slow kinetics of cone responses (such that flash responses ride on the initial rising/falling phase of the cone’s step response), and the subsequent nonlinearities of downstream retinal pathways.

**Generalization to other species.** Retinal saccadic suppression, at least its central component, was triggered by stimulus–stimulus interactions (Figs. 4, 6 and Supplementary Fig. 8), governed by general retinal signal processing, without the need for any specialized saccadic suppression circuit. It is likely that such general processing is conserved across species. Indeed, we observed quantitatively similar retinal saccadic suppression in pig ON and OFF RGCs (Supplementary Fig. 14), including the dependency on background texture statistics. Interestingly, like in mouse, OFF RGCs in pig retina also recovered faster than ON RGCs, suggesting that the surround component of suppression was also present in pig ON RGCs. In an additional experiment, we also recorded the activity of RGCs from an ex vivo macaque retina while subjecting it to a shorter version of the luminance-step paradigm (Supplementary Fig. 1b) with fewer conditions than in the original paradigm. Our results (Fig. 7) indicate that macaque RGC responses to probe flashes, following luminance steps, were suppressed in a way similar to mouse retina. However, more data will be required to determine the population trend and for characterizing the dictionary of response modulations in macaque retina.
Downstream visual areas may modulate retinal saccadic suppression. Given the similarities we previously described between retinal and perceptual saccadic suppression, it was tempting to test whether the crossover style of suppression, observed in the retina (Fig. 4b), was also reflected in perception. We therefore conducted human psychophysics experiments where we asked human subjects (N = 5) to maintain saccade-free fixation, while we simply changed the luminance of the homogenous background to a brighter (0.3–0.56 Michelson contrast) or darker (−0.3 to −0.56 Michelson contrast) background (Fig. 8a; “Methods”). At random times relative to the luminance step, we presented a dark (−0.033 Michelson contrast) or bright probe.
Fig. 6 Model RGC responses to probe flashes following luminance steps. a, b Spiking response of model ON (columns 1–2) and OFF (columns 3–4) RGCs to luminance steps alone (a, blue) and to luminance steps followed by probe flashes (b, black) at 50, 250, and 2000 ms (different rows; analogous to real RGCs in Fig. 1b, c). Luminance steps are depicted by intensity bars in (a). First column in each cell type: responses following a positive-contrast luminance step; second column: responses following a negative-contrast luminance step. Vertical blue lines: timing of luminance step; orange bars: timing of probe flashes. Note the ON RGC and OFF RGC did not spike in response to negative-contrast (column 2) and positive-contrast (column 3) luminance steps respectively. c Flash-induced responses, after subtracting (a) from (b), overlaid to show the modulation of probe flash responses at different times (analogous to real RGCs in Fig. 1d). Lines connecting the response peaks highlights the time courses of suppression relative to baseline flash-induced responses (2000 ms). d Modulation indices for probe flashes in ON (light gray) and OFF model RGCs (dark gray), following positive-contrast (top panel) and negative-contrast (bottom panel) luminance steps. Modulation indices were calculated based on model responses to probe flashes presented at 10-ms intervals after luminance steps, and baseline as response to a flash at 2000 ms. Circle markers indicate modulation indices based on probe flashes at 50 and 250 ms shown in (b, c). Cyan and red arrows highlight the suppression of opposite-contrast flashes at 50 ms in ON and OFF RGCs, respectively. e Simplified schematic of the model (left), and stimulus and response traces at the different processing steps (right). Signal resulting from light stimulus (in this case positive-contrast luminance step followed by dark flashes) is passed through the cone model and the resulting output is filtered and thresholded. Extended schematic of the model showing different filter transparency and nonlinearity threshold is shown in Supplementary Fig. 13. f Modulation indices as a function of RGC transiency for real OFF RGCs (dark gray circles; N = 92; red line: linear regression fit) and ON RGCs (light-gray circles; N = 228; cyan line: linear regression fit). Individual panels correspond to different flash times after positive-contrast (top row) and negative-contrast (bottom row) luminance steps. These RGCs are a subset of the population data shown in Fig. 4b for which we could compute a transiency index (‘Methods’). Suppression after negative-contrast steps was weaker in less transient ON RGCs (bottom row, blue regression line has negative slope) and, after positive-contrast steps, suppression was weaker in less transient OFF RGCs (top row, red regression line has negative slope). Numbers in each panel indicate the slope of the fits and asterisk symbol indicates statistically significant slope (slope \( \neq 0 \), \( P < 0.01 \), two-tailed t-test). g Modulation index of model OFF (red) and ON (cyan) RGCs as a function of the pathway’s transiency, where transparency was varied by changing the transparency of the filter shown in (e). h Modulation index of model ON and OFF RGCs as in (g) but as a function of model’s nonlinearity threshold. Arrows in (g) highlight the same data as in (d). In (a–d), the filter transiency was set to 1 and the nonlinearity threshold to 0.1. In (g), nonlinearity threshold was set to 0.1. In (h), the filter transiency was set to 0. Supplementary Fig. 12 shows model RGC responses based on real cone data of Fig. 5 instead of model cone responses.

flash (+0.033 Michelson contrast), at one of four locations in the subjects’ field of view. At trial end, the subjects were asked to localize the probe flash.

Irrespective of the step → flash combination, subjects were strongly impaired in their ability to localize the probe flashes presented around the time of the luminance step (Fig. 8b). Most interesting in this context was the combination of negative-contrast luminance steps with dark probe flashes. In the mouse retina, even though few OFF RGCs did show weak suppression to this combination (Supplementary Fig. 8b, inverted histograms in row 2), this effect was virtually absent at the population level (Fig. 4b, right panel). In human perception, however, this combination led to strong suppression (Fig. 8b, right panel). We cannot exclude that stronger retinal suppression to this specific combination might be present under different light or stimulus conditions. Another possibility is that it might be more pronounced in the retina of humans and other primate species (Fig. 7). Nonetheless, visual mechanisms of suppression certainly exist in higher visual brain areas. Perceptual suppression after same-contrast stimulus combinations may arise from processing in these higher visual brain areas, which may modulate and complement retinal saccadic suppression to achieve robust effects at the perceptual level.

Apparent pre-saccadic suppression in the retina. Throughout this study we characterized suppression in RGCs of their responses to preferred contrast flashes (bright flashes for ON RGCs and dark flashes for OFF RGCs). Here, we analyze a subset of RGCs in our dataset that also showed responses to flashes of nonpreferred contrast (i.e., some ON RGCs showed responses to dark flashes, and some OFF RGCs to bright flashes). These flash responses were also strongly suppressed around the time of saccade-like image displacements (Fig. 9a, red lines), similar to suppression of preferred contrast flashes (Fig. 9a, green lines), with one key addition: nonpreferred contrast flashes that were presented before saccade onset were also suppressed. This pre-saccadic suppression in the retina is reminiscent of the pre-saccadic suppression observed perceptually and neurally in other areas of the brain. One explanation for this apparent pre-saccadic suppression in RGCs is that the responses to non-preferred flashes had higher latencies than responses to preferred flashes. Figure 9b shows responses of an example OFF RGC to preferred and nonpreferred flashes presented 84 ms before saccade onset; Fig. 9c summarizes response latencies for the population of ON and OFF RGCs. Even though the flash occurred before the saccade onset, its peak response did not occur until much later after the saccade offset (Fig. 9b, right; population data shown in Fig. 9c). Suppressive mechanisms triggered by saccade-like image displacements can therefore act on this delayed flash response to suppress it. Measuring modulation of responses as a function of flash time relative to saccade onset (Fig. 9a), like in most studies of saccadic suppression, rather than as a function of response time, then gives the impression of pre-saccadic suppression. We did not further explore if the suppression of responses to nonpreferred stimuli originated from the same central, surround, and peripheral components as identified above (Fig. 3). Nonetheless, it is indeed intriguing that a wide array of observations pertaining to saccadic suppression at the perceptual level, including pre-saccadic suppression, are also observed at the level of the retina.

Discussion

For most RGCs that we recorded, responses to brief probe flashes were strongly suppressed when presented after saccade-like texture displacements across the retina. Similar suppression occurred when texture displacements were replaced by sudden uniform changes in background luminance, suggesting that suppression was caused by rather generic mechanisms, triggered by visual transients across the retina, rather than specialized suppression circuits that react to image motion. We found that the suppression strength depended on four main factors: (1) strength of the visual transients, governed by the statistics of the background texture or the contrast of the luminance step; (2) elapsed time following the visual transient; (3) RGC polarity (ON vs. OFF RGCs); and (4) RGC response properties (RGC transiency). Stronger visual changes, elicited either by coarser textures or larger luminance-step contrasts, caused stronger suppression, peaking ~50 ms after the stimulus offset (Figs. 1e and 4). The
FIG. 7 Retinal saccadic suppression in Macaque RGCs. a Average activity of an example macaque ON RGC (top) and OFF RGC (bottom) to positive-contrast (left column) and negative-contrast (right column) luminance steps alone (blue traces) and luminance steps followed by probe flashes (black traces). ON RGCs were analyzed for bright probe flashes and OFF RGCs for dark probe flashes. Responses were averaged across the different positive-contrast (0.05–0.5 Michelson contrast, N = 10 sequences) and negative-contrast luminance steps (−0.05 to −0.5 Michelson contrast, N = 10 sequences). b Median modulation index (thick lines) of macaque ON (light gray) and OFF (dark gray) RGCs for probe flashes presented after positive-contrast (left panel; N = 13 ON RGCs, N = 1 OFF RGC) and negative-contrast luminance steps (right panel; N = 7 ON RGCs, N = 2 OFF RGC). Circles represent modulation indices of individual RGCs. This experiment was performed under scotopic light conditions.

recovery times depended on RGC polarity: OFF RGCs recovered by 250–350 ms whereas suppression in ON RGCs lasted for up to 1 s (Figs. 1e and 4). The suppression was stronger in more transient RGCs (Fig. 6f–h). Mechanistically, we identified at least three components of retinal saccadic suppression, with distinct spatial origins, which we defined as central, surround, and global components (Fig. 3). These components were mediated by different underlying mechanisms.

The central component was the only source of suppression we could reliably find in OFF RGCs, and the dominant source in ON RGCs for time points immediately after a full-field saccade or luminance step. This component was short-lived (250–350 ms), originated from a cell’s receptive field center, and did not depend on inhibitory inputs. It was triggered by opposite-polarity stimulus–stimulus interactions, which naturally occur during saccades and other forms of sequential visual stimulation. This component of suppression resulted from temporal filtering of the relatively slow cone responses to two opposite-polarity consecutive stimuli and the subsequent thresholding nonlinearities. Such a mechanism, where the cone response itself remains linear, but nonetheless forms the basis for subsequent nonlinear response modulation, is clearly different from adaptation or desensitization mechanisms within the cones, which would evoke nonlinear responses of the cones themselves. We call this processing motif, triggered by sequential stimuli of opposite polarity, dynamic reversal suppression. Despite the simplicity of the underlying mechanism, this processing motif substantially shapes the input to downstream visual areas during conditions of natural vision.

The mechanism underlying dynamic reversal suppression also suggests that perceptual saccadic suppression is derived, at least in part, from the inherent response kinetics of photoreceptors, the very first cell in the visual processing cascade. In fact, this early implementation could also explain why we observed suppression in most RGCs we recorded from (Supplementary Figs. 3, 8, and 14), covering a wide spectrum of response properties and therefore presumably many RGC cell types (see Supplementary Fig. 6 of ref. 4). Still, our results suggest that the suppression initiated at the level of cone photoreceptors is translated differently by the different parallel pathways in the retina, leading to variability in response suppression at the ganglion cell level (such as the stronger suppression in more transient RGCs, Fig. 6f–h). Further cell type classification will be required to relate the degree of modulation with pathway specificity. The type and degree of modulation might also differ across species, even though we see qualitatively similar suppression in mouse, pig, and macaque RGCs.

ON RGCs, in addition to this central component, were suppressed by two more components. First, the global component is a fast but short-lived (~250–350 ms) component, caused by inhibition via GABAAergic wide-field amacrine cells, triggered by global image changes and carried to the RGC from as far as the cell’s periphery (Fig. 2b and Supplementary Fig. 5). This likely belongs to the same class of circuits that suppresses RGC responses to global motion. These circuits were previously suggested to suppress motion awareness during saccades, a phenomenon known as saccadic omission. As indicated by our results, such circuits also contribute towards suppressing RGC sensitivity even after the motion is completed (i.e., saccadic suppression). However, since their influence is masked by more local components of suppression during full-field saccades (Fig. 2c, d and Supplementary Fig. 6), they are unlikely to contribute much to perceptual saccadic suppression. The global suppressive component mediated by GABAAergic inhibition is nevertheless one mechanism to process global visual changes, in addition to several others. It may, for example, play a role in perceptual modulations during smooth pursuit eye movements. Here, the central component of suppression will not be triggered in RGCs whose receptive field centers are locked to the tracked object; but these RGCs will still be suppressed by the global component.

The second additional component is the surround component. It seems to act with a delay of ~200 ms and can last for up to 1000 ms. The spatial origins of this component were not investigated in this study, but our data indicate that it presumably arises from the immediate surround of a cell’s receptive field. Additionally, similar to the central component, it does not rely on GABAergic or glycine receptors (Fig. 2c, d and Supplementary
while Fig. 3 summarizes the three spatial components of sup-
ON bipolar cells but not in OFF bipolar cells46. In summary,
because response saturation has been observed predominantly in
component originates at the level of bipolar cells, especially
Fig. 10a), was found only in ON RGCs. It is possible that this
saturation-like mechanisms (Fig. 4b left and Supplementary
may not critically shape perception.

Interestingly, visual masking in cat LGN also lasts longer in ON
versus OFF cells52, which may be a consequence of the effects we
describe here in the retina. While this surround component plays
an important role in shaping RGC and downstream neural
responses following a visual transient, its contribution to per-
ceptual saccadic suppression can also be disputed. This is because,
during real saccades, eye-movement related signals (e.g., corollary
discharge) shorten the duration of suppression caused by visual
mechanisms4, such that the long-lasting surround component
may not critically shape perception.

Yet another additional component of suppression, based on
saturation-like mechanisms (Fig. 4b left and Supplementary
Fig. 10a), was found only in ON RGCs. It is possible that this
component originates at the level of bipolar cells, especially
because response saturation has been observed predominantly in
ON bipolar cells but not in OFF bipolar cells46. In summary,
while Fig. 3 summarizes the three spatial components of sup-
pression and their temporal properties, these components in turn
can have further sub-components.

Given its strength and time course, dynamic reversal sup-
pression is likely the most prominent suppressive retinal
component that contributes towards perceptual saccadic sup-
pression. Yet, irrespective of which components of retinal sac-
cadic suppression contribute towards perceptual saccadic suppression, our results show that retinal responses to stimuli
following visual transients are modulated concurrently by several
mechanisms (Fig. 3). Additional mechanisms might exist under
different stimulus conditions. From the perspective of retinal
visual feature processing, this would be consistent with how
multiple mechanisms concurrently process other visual features
in the retina, such as motion29.

The retinal suppression that we studied here likely contributes
to several other visual phenomena, such as visual masking53 or
neural adaptation with successive stimuli54. The similarities
between those phenomena and the suppression that we observed
in the retina suggest that the retina may be a common neural
substrate for these seemingly different types of perceptual phe-
nomena, unifying their underlying mechanisms.

It is remarkable that an elementary property of retinal sup-
pression, i.e., its dependence on the scene statistics (Fig. 1e), is
clearly preserved all the way to perception4. We also observed
pre-saccadic suppression in the retina, but only for responses to
flashes with nonpreferred contrast (Fig. 9a). This may, in addition
to other mechanisms55, contribute towards pre-saccadic sup-
pression observed in downstream visual areas10,48 and
perception4,17. Perhaps the apparent pre-saccadic suppression in

![Fig. 8 Perceptual suppression following luminance steps.](image-url)

**Fig. 8 Perceptual suppression following luminance steps.** A Example visual task trial. Subjects fixated a small spot on a uniform background with a set luminance (pre-step luminance) for a random duration (800-1700 ms). Background luminance then increased (like shown here) or decreased (positive or negative-contrast luminance step, respectively). At one of five times relative to the luminance step (−24, −12, 36, 72, or 108 ms), a luminance pedestal (probe flash, 147.8 × 147.8 min arc) was applied for -12 ms at one of four locations relative to the fixation spot: 7 deg above (shown here), below, to the right, or to the left. The probe flash was brighter (shown here) or darker than the current screen luminance. The background remained at the post-step luminance until the subject responded with the perceived location of the flash, plus an additional 500-1000 ms but without the fixation spot, allowing the subject to relax. The current luminance was the pre-step luminance of the consecutive trial. B Performance of human subjects (mean ± s.e.m., N = 5 subjects), to correctly localize a dark (dark gray) or a bright (light gray) probe flash presented at different times relative to positive-contrast (left panel) and negative-contrast (right panel) luminance steps (blue line). Each subject’s responses were averaged across the different positive-contrast (0.3–0.56 Michelson contrast) and negative-contrast (−0.3 to −0.56 Michelson contrast) luminance steps. Perceptual performance was reduced around the time of luminance steps, reflecting suppression, irrespective of the combination of luminance-step polarity and flash polarity. There were no statistically significant differences in suppression of dark and bright probe flashes (two-tailed Wilcoxon rank-sum test). Note that in the right panel, the suppression profile for bright probe flashes almost completely overlaps the suppression profile of dark probe flashes.
the retina may help elucidate the neural loci and mechanisms of backwards visual masking, a phenomenon observed in downstream visual areas and perceptually, where response to a stimulus is suppressed by a subsequent stimulus. We did not investigate further whether the suppression of nonpreferred contrast flashes also originates from the components identified in this study. A necessary prerequisite will be to understand the pathways mediating the delayed responses in the case of nonpreferred contrast flashes. Such pathways remain largely unexplored.

Not all properties of retinal suppression were preserved in perception. In the retinal output of mice, only successive stimuli of opposite contrasts triggered suppression (crossover-style suppression) (Fig. 4b), while in human perception, all contrast combinations led to strong suppressive effects (Fig. 8). The more far-reaching perceptual suppression may be the result of additional processing beyond the retina. Alternatively, our results may not capture the full array of retinal processing. For one, our stimulus conditions in mouse retina experiments may not have been comprehensive enough to capture all retinal suppressive...
Fig. 9 Suppression of nonpreferred contrast flashes. a Population modulation index (mean ± s.e.m.) of ON (left) and OFF (right) RGCs for preferred contrast flashes (green lines) and nonpreferred contrast flashes (red lines) presented before and after 150-μm texture-scale displacements (blue bar). For ON RGCs, preferred contrast flashes are bright and nonpreferred contrast flashes are dark. Vice versa for OFF RGCs. Probe flashes were presented at −117, −84, −67, −50, 50, 117, 200, 350, 600, 1100, and 2000 ms (baseline). The number of RGCs in the population varied at each time point; ON RGCs: 18, 18, 17, 17, 0, 17, 17, 17, 36, 17, 29; OFF RGCs: 35, 35, 31, 35, 21, 82, 83, 80, 119, 84, 58. These RGCs are a subset of the population data shown in Fig. 1e (column 3) that also showed robust responses to nonpreferred contrast flashes. Error bars are not visible in the right column due to small s.e.m. b Flash-induced response of an example OFF RGC (after subtracting saccade response which is not shown) to preferred contrast (dark) flash (left) and nonpreferred contrast (bright) flash (right) presented 84 ms before saccade onset (indicated by orange markers). Dashed line indicates the baseline (response to flash at 2000 ms after saccade onset) below which a flash response’s modulation index is negative. Labels “1” and “2” in brown circles refer to the corresponding points in panel d. Response to the preferred contrast flash (left) is not suppressed (peak response close to baseline). Response to nonpreferred contrast flash (right) occurs with a higher latency and is nonpreferred relative to the baseline. e Distribution of response latencies for preferred (green bars) and nonpreferred (red bars) flashes for ON RGCs (left, preferred (mean ± s.e.m.): 117 ± 3 ms, nonpreferred: 347 ± 15 ms) and OFF RGCs (right, preferred: 126 ± 4 ms, nonpreferred: 287 ± 8.2 ms). In both RGCs, nonpreferred contrast flash responses had longer latencies than the preferred contrast flashes. d Modulation indices of ON RGCs (columns 1-3) and OFF RGCs (columns 4-6), at different flash times as a function of time of flash response peak relative to saccade onset. Columns: flashes presented at different times before saccade onset. Rows: preferred contrast probe flashes (top) and nonpreferred contrast probe flashes (bottom). Dashed lines correspond to zero modulation on the y axis and saccade onset on x axis. Modulation indices of nonpreferred contrast flashes (red circles) were below 0 even though the flashes were presented before saccade onset (indicated by times above row 1). However, for these nonpreferred flashes, their peak responses occurred long after the saccade onset, as compared to preferred contrast flashes. The longer response latency of nonpreferred flashes therefore gives the impression of pre-saccadic suppression, when suppression is measured as time of flash presentation (like in a). Black circles in OFF RGC plot for flash at −84 ms are the modulation indices for the two example RGC responses shown in (b).

Effects. Further, the retinal output of humans and other primate species (Fig. 7) might differ from mouse retina in this respect.

Retinal suppression is only one way that the retina alters its output during dynamic vision. Other forms certainly co-exist, such as brief changes in RGC polarity following peripheral shifts or sensitization of some RGC types following a change in background luminance. These and several other studies, together with ours, demonstrate the complex image processing capabilities in the retina to facilitate downstream visual processing for the ultimate service of perception during natural vision. Looking forward, the detailed characterization of retinal output provided here paves the way to investigate the visual features that the retina encodes during dynamic vision. Moreover, it also paves the way to investigate the interactions between retinal and extraretinal (visual and nonvisual) mechanisms of saccadic suppression, to further our understanding of how the visual system maintains stability in the face of constant disruptions.

Methods

Experimental model and subject details

Animals. We performed electrophysiological experiments on ex vivo mouse, pig, and macaque retinas; and imaging experiments on ex vivo mouse retina. Mouse and pig ex vivo retina experiments were performed in Tübingen, in accordance with German and European regulations, and animal experiments were approved by the Regierungspräsidium Tübingen. Macaque ex vivo retina experiment was performed at Stanford University. Eyes were removed from a terminally anesthetized macaque chimpus monkey used by other laboratories in the course of their experiments, in accordance with the Institutional Animal Care and Use Committee guidelines of Stanford University. For mouse retina electrophysiology, we used 47 retinae from 15 male and 30 female pigs after they had been sacrificed during independent studies at the Department of Experimental Surgery at the Medical Faculty of the University of Tübingen. Pigs were anesthetized using atropine, azaperone, benzodiazepine (midazolam), and ketamine, and then sacrificed with embutramide (T61). Before embutramide administration, heparin was injected.

One experiment was conducted with a retina extracted from a macaque chimpus monkey.

Humans. Human psychophysics experiments were performed in Tübingen. Human subjects provided written, informed consent, and they were paid 10 Euros per session of 60 min each, for three sessions. Human experiments were approved by ethics committees at the Medical Faculty of Tübingen University, and they were in accordance with the Declaration of Helsinki. In total, we collected data from five subjects (24–29 years old; one female).

Experimental setup: mouse and pig retina electrophysiology. Mice were dark-adapted for 4–16 h before experiments. We then sacrificed them under dim red light, removed the eyes, and placed eyecups in Ringer solution (in mM: 110 NaCl, 2.5 KCl, 10 CaCl₂, 1.6 MgCl₂, 10 glucose, and 22 NaHCO₃) bubbled with 5% CO₂ and 95% O₂. We removed the retina from the pigment epithelium and sclera while in Ringer solution.

Pigs were dark-adapted for 15–20 min before sacrifice. Immediately after veterinary-confirmed sacrifice, the eyes were enucleated under dim red light, and the cornea, lens, and vitreous were removed. Eyecups were kept in CO₂-independent culture medium (Gibco) and protected from light. We transported eyecups to our laboratory and cut pieces from mid-peripheral or peripheral retinas. Only those retinas which showed ganglion cell responses to light stimuli were used in our experiments.

We recorded mouse and pig retinal ganglion cell (RGC) activity using either low- or high-density multi-electrode arrays (MEAs). The low-density setup consisted of a perforated 60-electrode MEA (60pMEA200/30ir-Ti-gt, Multichannel Systems (MCS), Reutlingen, Germany) having a square grid arrangement and 200 μm inter-electrode distance. We whole mounted an isolated retina on nitrocellulose filter (Millipore) with a central 2 × 2 mm hole. The mount was placed with the RGC side down into the recording chamber, and good electrode contact was achieved by negative pressure through the MEA perfusion. We superfused the tissue with Ringer solution at 30–34 °C during recordings, and we recorded extracellular activity at 25 kHz using a USB-MEA-system (USB-MEA1006, Multichannel Systems) or a memory-card based system (MEA1060, Multichannel Systems). Data was acquired using MC Rack version 4.6.2 (Multichannel Systems). A detailed step-by-step approach is provided in ref. 59.

The high-density MEA setup consisted of either a HiDens CMOS MEA60 (developed by the lab of Andreas Hierlemann, Basel, Switzerland) or a MaxOne system (Multichannel Systems). The HiDens CMOS MEA featured 11,011 metal electrodes with inter-electrode (center-to-center) spacing of 18 μm placed in a honeycomb pattern over an area of 2 × 1.75 mm. Any combination of 126 electrodes could be selected for simultaneous recording. The MaxOne MEA featured 26,400 metal electrodes with center-to-center spacing of 17.5 μm in a grid-like arrangement over an area of 3.85 × 2.1 mm. In this system, up to 1024 electrodes could be selected for simultaneous recordings. For each experiment, a piece of isolated retina covering almost the entire electrode array was cut and placed RGC-side down in the recording chamber. We achieved good electrode contact by applying pressure on the photoreceptor side of the retina by carefully lowering a transparent permeable membrane (Corning Transwell polyester membrane, 10 μm thick, 0.4-μm pore diameter) with the aid of a micromanipulator. The membrane was drilled with 200 μm holes, with center–center distance of 400 μm, to improve access of the Ringer solution to the retina. We recorded extracellular activity at 20 kHz using FPGA signal processing hardware. In the case of the HiDens CMOS MEA, data were acquired using custom data acquisition software, called MEA 1x Scope (developed by the lab of Andreas Hierlemann, Basel, Switzerland). In the case of the MaxOne MEA, data were...
acquired using MaxLab software provided by Maxwell Biosystems, Basel, Switzerland.

In total, we performed 59 recordings, 47 from mouse and 12 from pig retinas. In total, 24 of the 59 recordings were done using low-density MEAs. Once a basic experimental protocol was established, we shifted to HiDens CMOS MEA providing much higher throughput. 12 experiments were done using this setup. We tended to be weaker in the central retina, most scan fields were acquired using MaxLab software provided by Maxwell Biosystems, Basel, Switzerland. 12 experiments were done using this setup. We updated to the MaxOne MEA for even higher throughput and did 23 recordings using this setup of the data collected. In the 27 experiments, we used 5× (from mouse and 12 from pig retinae), was also used in our previous study 9. Here, we show further in-depth analysis of that data.

We presented light stimuli to the retinal piece that was placed on the MEA using a DLP projector running at 60 Hz (Acer K11 for low-density MEA experiments and Lightcrafter 4500 from EKB Technologies Ltd. with internal red, green and blue light-emitting diodes, for high-density MEA experiments). 60 Hz is above the flicker fusion frequency of both mouse and pig retinae; therefore, the framerate of these projectors was adequate for our purposes. The Acer K11 projector had a resolution of 800 × 600 pixels covering 3.225 mm on the retinal surface. Light on the retina was perfused with oxygenated Ames solution containing (in mM): 125 NaCl, 2.5 KCl, 2 CaCl₂, 1 MgCl₂, 1.25 NaH₂PO₄, 26 NaHCO₃, 20 glucose, and 0.5 14C-glutamine (pH 7.4). We then moved the tissue to the recording chamber where it was immersed in the recording chamber the retina was perfused with oxygenated Ames solution (ACSF) solution containing (in mM): 125 NaCl, 2.5 KCl, 2 CaCl₂, 1 MgCl₂, 2.5 NaH₂PO₄, 26 NaHCO₃, 20 glucose, and 0.5 14C-glutamine (pH 7.4) before the recording chamber was filled with the Ringer solution in a 1:1000 ratio. Wash-in was performed for at least 15 s before the laser scanning started and before stimuli were presented.

Experimental setup: macaque retina electrophysiology. In one experiment, we recorded the activity of macaque retinal ganglion cells. For this experiment, we used a high-density MEA 62,63. Following enucleation, the anterior portion of the eye and vitreous were removed. The eye was stored in a dark container in oxygenated Ames solution (Sigma, St. Louis, MO) at 33 °C, pH 7.4. Under infrared illumination, a small piece of retina (~1 × 1 mm), from a retinal region with eccentricity around 12 mm (4.0 – 17 mm temporal equivalent eccentricity; Chichilnisky and Kalmar 65), was dissected and placed ganglion cell side down on a MEA for recording. The retina pigment epithelium remained attached during the recording. A retinal detachment was performed using a dissecting microscope to planar large-scale MEA 63,65 with a hexagonal outline of 519 electrodes at 30 μm pitch was used. Recorded voltages were band-pass filtered, amplified, and digitized at 20 kHz using custom electronics 65. The detailed spike sorting process is described in ref. 66.

Visual stimulation was performed with the optically reduced image of a gamma-corrected OLED microdisplay (eMagin) refreshing at 60.35 Hz focused on the photoreceptor outer segments. The visual stimulus was delivered through the mostly-transparent electrode array. The power of each display primary was measured at the preparation with a calibrated photodetector (UDT Instruments). At the mean flicker fusion frequency of individual analyses from the individual setups revealed no effects of recording conditions in the different setups. For experiments of Supplementary Fig. 4, we also recorded at scotopic light levels where a stimulus intensity of 200 – 800 μm, extending 3.072 × 1.92 mm on the retinal surface. Light from the projector was focused through the microscope objective (high-density MEAs, illumination from above). In each case, the light path contained a shutter and two motorized filter wheels with a set of neutral density (ND) filters (Thorlabs NE101B-A to NE501B-A), having optical densities from 1 (ND1) to 5 (ND5). The filters allowed us to adjust the absolute light level of the stimulus.

We measured the spectral intensity profile (in μW cm⁻² nm⁻¹) of our light stimuli with a calibrated USB2000 + spectrophotometer (Ocean Optics) and confirmed the physical intensity into a biological equivalent of photoosmoticmosm of blood vessels and any damaged cells in the retina. We aimed to cover a range of 3 log units (i.e., 1000-fold difference between black and white pixels, over an 8-bit range). We linearized the photoreceptors using a condenser (low-density MEA recordings, illumination from below) or a x5 objective (high-density MEA, illumination from above). In each case, the light path contained a shutter and two motorized filter wheels with a set of neutral density (ND) filters (Thorlabs NE101B-A to NE501B-A), having optical densities from 1 (ND1) to 5 (ND5). The filters allowed us to adjust the absolute light level of the stimulus.

In one experiment, we recorded the activity of macaque retinal ganglion cells. For this experiment, we used a high-density MEA 62,63. Following enucleation, the anterior portion of the eye and vitreous were removed. The eye was stored in a dark container in oxygenated Ames solution (Sigma, St. Louis, MO) at 33 °C, pH 7.4. Under infrared illumination, a small piece of retina (~1 × 1 mm), from a retinal region with eccentricity around 12 mm (4.0 – 17 mm temporal equivalent eccentricity; Chichilnisky and Kalmar 65), was dissected and placed ganglion cell side down on a MEA for recording. The retina pigment epithelium remained attached during the recording. A retinal detachment was performed using a dissecting microscope to planar large-scale MEA 63,65 with a hexagonal outline of 519 electrodes at 30 μm pitch was used. Recorded voltages were band-pass filtered, amplified, and digitized at 20 kHz using custom electronics 65. The detailed spike sorting process is described in ref. 66.

Visual stimulation was performed with the optically reduced image of a gamma-corrected OLED microdisplay (eMagin) refreshing at 60.35 Hz focused on the photoreceptor outer segments. The visual stimulus was delivered through the mostly-transparent electrode array. The power of each display primary was measured at the preparation with a calibrated photodetector (UDT Instruments). At the mean flicker fusion frequency of individual analyses from the individual setups revealed no effects of recording conditions in the different setups. For experiments of Supplementary Fig. 4, we also recorded at scotopic light levels where a stimulus intensity of 200 – 800 μm, extending 3.072 × 1.92 mm on the retinal surface. Light from the projector was focused through the microscope objective (high-density MEAs, illumination from above). In each case, the light path contained a shutter and two motorized filter wheels with a set of neutral density (ND) filters (Thorlabs NE101B-A to NE501B-A), having optical densities from 1 (ND1) to 5 (ND5). The filters allowed us to adjust the absolute light level of the stimulus.

In one experiment, we recorded the activity of macaque retinal ganglion cells. For this experiment, we used a high-density MEA 62,63. Following enucleation, the anterior portion of the eye and vitreous were removed. The eye was stored in a dark container in oxygenated Ames solution (Sigma, St. Louis, MO) at 33 °C, pH 7.4. Under infrared illumination, a small piece of retina (~1 × 1 mm), from a retinal region with eccentricity around 12 mm (4.0 – 17 mm temporal equivalent eccentricity; Chichilnisky and Kalmar 65), was dissected and placed ganglion cell side down on a MEA for recording. The retina pigment epithelium remained attached during the recording. A retinal detachment was performed using a dissecting microscope to planar large-scale MEA 63,65 with a hexagonal outline of 519 electrodes at 30 μm pitch was used. Recorded voltages were band-pass filtered, amplified, and digitized at 20 kHz using custom electronics 65. The detailed spike sorting process is described in ref. 66.

Visual stimulation was performed with the optically reduced image of a gamma-corrected OLED microdisplay (eMagin) refreshing at 60.35 Hz focused on the photoreceptor outer segments. The visual stimulus was delivered through the mostly-transparent electrode array. The power of each display primary was measured at the preparation with a calibrated photodetector (UDT Instruments). At the mean flicker fusion frequency of individual analyses from the individual setups revealed no effects of recording conditions in the different setups. For experiments of Supplementary Fig. 4, we also recorded at scotopic light levels where a stimulus intensity of 200 – 800 μm, extending 3.072 × 1.92 mm on the retinal surface. Light from the projector was focused through the microscope objective (high-density MEAs, illumination from above). In each case, the light path contained a shutter and two motorized filter wheels with a set of neutral density (ND) filters (Thorlabs NE101B-A to NE501B-A), having optical densities from 1 (ND1) to 5 (ND5). The filters allowed us to adjust the absolute light level of the stimulus.

Experimental setup: cone photoreceptor imaging. To record the output of cone photoreceptors in the mouse retina, we measured the glutamate release using an intensity-based glutamate-sensitive fluorescent reporter, iGluSnFR 43 expressed in horizontal cell processes postsynaptic to cone terminals, using a viral approach. We recorded the cone output from 4 retinae obtained from two C57BL/6 male mice, 9–10 weeks old. Below, we reproduce the methods, previously described in ref. 44. We injected the mice with 2–1 h before the experiments. If not stated otherwise, all experiments were done using this setup. We then adapted the mice for 2–1 h before the experiments. They were then anesthetized using isoflurane (Baxter) and sacrificed by cervical dislocation. The eyes were enucleated and hemisected in carbogenated (95% O2 and 5% CO2) Ringer solution containing NaCl, KCl, MgCl2, CaCl2, glucose, and HEPES (pH 7.4). We then moved the tissue to the recording chamber and perfused with oxygenated Ringer solution in a 1:1000 ratio. Wash-in was performed for 20 min.

Visual stimuli for the saccade paradigm used in retina electrophysiology experiments, we used two broad visual stimulation paradigms: a saccade (texture displacement) paradigm (Supplementary Fig. 1a), and a luminance-step paradigm (Supplementary Fig. 1b), described in detail below. In different experiments, we used different spatial and/or pharmacological manipulations of these two paradigms.

Background textures. We created background textures (Supplementary Fig. 2a) by convolving a random binary (i.e., white or black) pixel image with a two-dimensional Gaussian blur filter. The filter width was chosen to match the size of the display pixels, which was approximately 6.1° of visual angle.
The parameter \( \sigma \) of the kernel influenced the amount of blurring. This resulted in textures having effectively low-pass spectral content (Supplementary Fig. 2b) with a cutoff frequency depending on \( \sigma \). For easier interpretation, we define the spectral content of the textures by a spatial scale. Intuitively, the spatial scale approximates the size of the smallest dark and bright image blobs of the texture (Supplementary Fig. 2a). Quantitatively, the spatial scale is defined as the 2\( \sigma \) parameter of the Gaussian blurring filter. We generated textures with four different spatial scales: 25, 50, 150, and 300 \( \mu \)m, that resulted in dark and bright image blobs approximating a range of receptive field sizes between bipolar cells (texture with spatial scale 25 \( \mu \)m, see ref. 73) and RGCs (textures with spatial scale 150 and 300 \( \mu \)m). In other words, coarser textures matched the resolution of RGCs, and finer textures matched the resolution of one processing stage earlier, the retinal bipolar cells. Calculating power spectra for the textures (Supplementary Fig. 2b) confirmed that the spatial scale and hence the cutoff frequencies were consistent with this design aim. In different experiments, we used textures of all or a subset of the different spatial scales.

We normalized the pixel intensities in the textures to have uniform variations in luminance around a given mean. We used pixel intensities (from our 8-bit resolution scale) ranging from 0 to 60 around a mean of 30, or ranging from 30 to 90 around a mean of 60 (see sub-section "Saccades and probe flashes" for when each paradigm was used).

Saccades and probe flashes. To simulate saccades in our ex vivo retina electrophysiology experiments, we displaced the texture across the retina in 6 display frames (corresponding to a velocity range of 300–9300 \( \mu \)m s\(^{-1}\) on the retinal surface). In visual degrees, this corresponds to a velocity range of 100–300 deg s\(^{-1}\) and displacement range of 10–30 deg in mice, well in the range of observed mouse saccade amplitudes\(^8\). Similar to primates, mice also have oculomotor behavior, even under cortical control\(^9\). For example, they make miniature, saccade-like rapid eye movements per minute when their head is fixed\(^14\) (humans make several saccades per second). We used the same retinal displacement range of 310 \( \mu \)m to 930 \( \mu \)m for pig retinas. To the best of our knowledge, pig oculomotor behavior has not been documented in the literature. However, with their larger eyeball sizes, our translations of the retinal image would correspond to saccades (e.g., small saccades in humans and monkeys), which are also associated with saccadic suppression. Moreover, retinal saccadic suppression is not critically dependent on the details of movement kinematics, as it is triggered by visual transients (Fig. 4, also see Figs. 4, 5 in ref. 4).

Each trial consisted of successive sequences (Fig. 1a and Supplementary Fig. 1a) that combined a saccade with a probe flash, as follows: there was first a pre-saccade fixation of 2 s, where the texture remained static over the retina, then a 100 ms saccade, followed by post-saccade fixation where the texture again remained static over the retina but now with a shifted texture. At a certain time from saccade onset (delay \( d \), range: 30 ms to 2100 ms), we presented a probe flash (see below). Following the post-saccade fixation, the texture remained as-is the texture remained in the same position for another 2 s before the next saccade of the successive sequence occurred. The post-probe flash fixation of one sequence was therefore also the pre-saccade fixation of the next sequence. This way the texture remained visible during the entire trial, being translated during saccades of the successive sequences. In a single trial, 20 to 30 sequences occurred. The direction and amplitude of the saccade were pseudo-randomly determined by the range of allowed saccade amplitudes and directions. The texture always landed at unique locations within a trial. The end result was that, within a single trial, RGCs experienced a wide spectrum of saccade amplitudes, directions, and contrasts across these 29 saccades. As such, by analyzing the average effects of the 39 saccades on RGC responses to probe flashes, we captured a wide range of saccade-induced kinematics and luminance changes over the RGC receptive fields. In most cases, the probe flash had a duration of 2 frames (\( \sim \)33 ms). We used 1 frame (\( \sim \)16 ms) in a subset of earlier experiments (mouse: 709 of 1616 cells; pig: 116 of 228 cells). Results were pooled across these paradigms as the effects were indistinguishable.

The probe flash was a full-screen positive (bright) or negative (dark) stimulus transient.

Bright or dark probe flashes could happen in two different ways across our experiments. The results were indistinguishable between the two ways, so we pooled them throughout. Briefly, in one way, the probe flash replaced the texture with a homogeneous bright (pixel intensity of 60 in our 8-bit projectors) or dark (pixel intensity of 0) full-screen (in these experiments, the textures themselves had intensities ranging from 0 to 60 pixel intensity; see "Background textures" above). This way, the flash contrast from the underlying background luminance could be varied across a bright full-screen (on a dark pixel background space) to a bright flash on a bright background, or a dark flash on a dark background (on a bright pixel background space). In the second manipulation, the bright and dark flashes were simply luminance increments or decrements (by pixel values of 30 on our 8-bit projectors) over the existing textures. This way, spatial contrast relationships in the background textures were maintained. In the above experiments, the textures themselves had a range of 30–90 pixel intensities and a mean pixel value of 60 (on our 8-bit projectors). Out of the 1616 RGCs that we analyzed for saccadic suppression across all experiments where texture displacements were used as saccades (irrespective of the spatial or pharmacological manipulations), 1120 RGCs experienced such probe flashes, whereas the rest (487 RGCs) experienced the homogenous probe flash.

Full-field saccades. In the full-field saccades experiments, saccades and probe flashes occurred over the entire retina. This was the main experimental paradigm that we used for characterizing how saccades modulate RGC responses to probe flashes. This paradigm was also used as a control in experiments in which we applied different spatial or pharmacological manipulations of this paradigm to probe for spatial origins and mechanisms of saccadic suppression. Further, results from this paradigm served as a baseline standard across different experimental rigs. This paradigm was used in a total of 32 retinal recordings (32 retinas from 30 mice). In different recordings, we used a subset of the textures scales and probe flash delays. This explains the different values of \( N \) seen for different conditions in, for example, Fig. 1e and Supplementary Fig. 3. However, to ensure comparison, some conditions were always overlapped across different recordings. This paradigm was also used in 12 recordings with retinæ from 6 pigs.

Periphery saccades (global component of suppression). In this manipulation, we restricted saccades to the RGC's receptive fields periphery (i.e., its far surround). This manipulation was used to investigate the global component of suppression (Fig. 2a, b and Supplementary Fig. 5). We performed 13 recordings (13 retinæ from 13 mice) with this paradigm, always with the high-density multielectrode array system (MaxOne by MaxWell), as it provided a large electrode array \( \sim \)2 × 4 mm large, allowing the periphery stimulus to be placed on the retina to be placed on the retina. This paradigm was typically either a high-density block of electrodes (inter-electrode spacing: \( \sim \)17.5 \( \mu \)m) or a block with one-electrode spacing (inter-electrode spacing: \( \sim \)35 \( \mu \)m). The recording region was selected close to the center of the electrode array. We centered a large square mask (1000 × 1000 \( \mu \)m\(^2\)) over the recording region to restrict the texture and saccade presentation to the periphery of RGC receptive fields (Supplementary Fig. 5b). The mask had a homogenous intensity corre-
we used probe flash duration of ~17 ms, and a coarse texture background of spatial scale of 150 μm.

**Visual stimuli for luminance-step paradigm used in retina electrophysiology (Figs. 4 and 7 and Supplementary Figs. 8-10).** In this paradigm (Supplementary Fig. 1b), we used no textures at all. The screen was always a homogenous gray field, and the visual event of a saccade was replaced by an instantaneous step to a different gray background, independently of luminance intensities between −90 (on our 8-bit projector). The instantaneous step in intensity caused either a positive-contrast luminance step (in the range of +0.03 to +0.50 Michelson contrast) or a negative-contrast luminance step (−0.03 to −0.50 Michelson contrast). This paradigm was used to characterize the stimulus-stimulus interactions that ultimately make up retinal saccadic suppression (4, 7 and Supplementary Figs. 8-10). We performed a total of four recordings (four retinae from four mice) with this paradigm, always with the high-density MaxOne MEA rig. A trial consisted of either 56 or 156 successive sequences (Supplementary Fig. 1b) that each combined a luminance step with a probe flash, as follows: there was first a pre-step fixation of 2 s where the retina was exposed to a fixed gray level (analogous to pre-saccade fixation in texture displacements), then an instantaneous switch to post-step fixation (analogous to post-saccade). At a certain time from the luminance step (delay: 17, 33, 50, 100, 250, 500, 1000, or 2000 ms), we presented a 2-frame (~33 ms) dark (−0.33 Michelson contrast) or bright (+0.33 Michelson contrast) probe flash. Some sequences combined suppression flash, the next luminance step then happened 4 s after the previous one. In a given experiment, we had 17 trials representing the 17 conditions: 8 flash delays × 2 probe flash polarities +1 condition with no probe flash. Similar to the saccadic paradigm, we always interleaved all conditions; that is, in any one of the 17 necessary trials, each luminance step could look like either a bright or a dark probe, to both ON and OFF steps, or no probe flash. Moreover, we repeated the 17 trials at least four times. A shorter version of this paradigm was used in our macaque retina recording (Fig. 7). Here, a trial consisted of 20 successive sequences. The 20 luminance steps induced contrasts in the range −0.5 to +0.5. Flashes of −33 ms were presented with a delay of 17, 50, 100, 250, 500, and 2000 ms after the luminance step.

**Visual stimuli for RGC characterization batch used in all retina electrophysiology experiments.** We used other stimuli unrelated to the main experiments to help us characterize RGC properties (e.g., response polarity, latency, transience, and spatial receptive fields). These stimuli had the same mean intensities and intensity ranges as the textures or luminance steps used in each experiment. Below, we describe these stimuli for the condition in which the texture intensities ranged from 0 to 60 pixel intensity (represented as grayscale RGB values in the units of our 8-bit projector). In experiments in which the textures ranged in intensity from 30 to 90, or the luminance-step experiment, all intensities reported below were shifted upward by 30. (1) Full-field contrast steps. ON steps: stepping from 0 to 30 (+1 Michelson contrast) and from 30 to 60 (+0.33) for 2 s. OFF steps: stepping from 60 to 30 (~0.33) and from 30 to 0 (~1) for 2 s. (2) Full-field Gaussian flicker, 1 min. Screen brightness was updated every frame and was drawn from a Gaussian distribution with mean 30 and standard deviation 9. This stimulus was used to calculate linear filters representing the temporal receptive fields of RGCs through reverse correlation (spike-triggered averaging of the stimulus history). (3) Binary checkerboard flicker, 10–15 min. The screen was divided into a checkerboard pattern; each checker either covered an area of 55 × 55 μm, 60 × 60 μm, or 65 × 65 μm depending on the recording rig. The intensity of each checker was updated independently from the other checkers and randomly switched between 10 and 50 or 0 and 120. This stimulus also allowed us to calculate linear filters representing the spatial receptive fields of RGCs.

**Visual stimuli for cone photoreceptors imaging (Figs. 5 and 6).** For cone imaging experiments, we used a minimalistic version of the luminance-step paradigm used in retinal electrophysiology. A homogeneous background alternated between a darker (pixel intensity 50 on 8-bit projector) and brighter gray value (pixel intensity 120 on 8-bit projector); the transitions between these two background values represented positive and negative contrast of 0.4 Michelson contrast. At various times after the luminance step (50, 250, and 2000 ms) we presented a probe flash (100 ms duration, +0.33 or −0.33 Michelson contrast). The probe flash at 2000 ms served as the baseline. The next background transition always happened 2 s after the preceding probe flash. The combination of the two luminance steps and the two probe flash polarities yielded a total of four combinations: negative-contrast luminance step followed by dark flash; negative-contrast luminance step followed by bright flash; positive-contrast luminance step followed by dark flash; and positive-contrast luminance step followed by a bright flash. A single trial (Supplementary Fig. 1c) was composed of the four step-flash combinations occurring three times (for the three delays with which the flash was presented); and the negative- and positive-contrast luminance step without a flash. Within a trial, these combinations were randomly ordered. A trial consisted of three times, the luminance steps and the flashes were presented within a 700-μm disc region centered over the scan field.

**Visual stimuli for human psychophysics (Fig. 8).** In the human psychophysics experiment (Fig. 8), we mimicked the retinal luminance-step experiments of Fig. 4. The paradigm in Fig. 8a was similar to the one described in the central fixation spot over a gray background that remained there for the entire duration of a trial. The background had one of 8 luminances (22.4, 30.24, 38.08, 45.92, 53.76, 61.6, 69.44, 77.48 cd m−2). After a random initial fixation duration (800–1700 ms after fixation spot onset), the luminance of the background was changed suddenly (in one display frame update) to one of the remaining 7 luminances, inducing a positive-contrast luminance step or a negative-contrast luminance step. In our analysis, we used the luminance steps that induced contrasts in the range +0.3 to +0.56 Michelson contrast and −0.3 to −0.56 Michelson contrast. A total of 16 different possible combinations (each step: −24, −12, 36, 72, or 108 ms), a luminance pedestal (probe flash) was applied briefly for one display frame (~12 ms) at one of four locations relative to display center (7 deg above, below, to the right of, or to the left of center). Note that because the display was rasterized that (rasterized, that is, drawn by the computer graphics board from the top left corner in rows of pixels), the exact flash time and duration depended on the location of the flash on the display (but in a manner like other psychophysical experiments studying the same phenomenon, and also in a manner that is unlikely to affect our results). The luminance pedestal consisted of a square of 147.8 × 147.8 min arc, in which we added or subtracted a value to represent bright and dark probe flashes. We ensured that the contrast of the flash (relative to the currently displayed background luminance) was always the same across all trials: +0.033 for a bright flash, and −0.033 for a dark flash. Following each trial, the fixation spot was removed from the background to allow the subjects to relax. This inter-trial period lasted for 500–1000 ms. The next trial happened consecutively, in a way that the current luminance of the background was used as the previous luminance. Subjects were instructed to fixate through the entire trial period and simply reported the locations of the brief flashes. Each subject performed three sessions, with 1120 trials per session.

**Data analysis for retina electrophysiology**

**MEA recordings preprocessing.** Low-density MEA recordings were high-pass filtered at a 500 Hz cutoff frequency using a tenth-order Butterworth filter. We extracted spike waveforms and times using thresholding, and we semi-manually sorted spikes using custom software. For high-density MEA recordings, we performed spike sorting by an offline automatic algorithm76 and assessed the sorted units using a custom-developed tool, the UnitBrowser77. We judged the quality of all units using inter-spike intervals and spike shape variation. Low-quality units, such as ones with high inter-spike intervals, missing spikes, or contamination, were discarded. After removing these units, all based on spiking criteria. In total, we extracted 3510 high-quality units after the spike sorting (referred to as RGCs from now on), from recordings of mouse retina. From pig retina recordings, we extracted 376 RGCs and from macaque retina, we extracted 57 RGCs after the spike sorting. However, as we mention below, only a subset of these could be analyzed for saccadic suppression.

**RGCs characterization using receptive fields, ON−OFF index, transiency index.** We first characterized the properties of RGCs. We calculated linear filters in response to full-field Gaussian flicker and binary checkerboard flicker by summing the 500-ms stimulus history before each spike. The linear filters allowed for determining cell type. More specifically, the amplitude of the first peak of the filter was used. If the peak was positively deflected, the cell was categorized as an ON cell; if negatively deflected, the cell was an OFF cell. ON cells were later always analyzed with respect to their responses to bright probe flashes, and OFF cells were analyzed with dark probe flashes. We determined the spatial receptive fields of RGCs by calculating the linear filters for each region (checker) defined by the binary checkerboard flickering stimulus. The modulation strength of each linear filter, measured as the standard deviation (s.d.) along the 500 ms temporal kernel, is an estimate of how strongly that region drives ganglion cell responses. We fitted the resulting 2D-map of s.d. values with a two-dimensional Gaussian and took the 2-σ ellipse (long axis) as the receptive field diameter. For all other figures and analyses, we converted spike times to estimates of firing rate by convolving these times with a Gaussian of σ = 10 ms standard deviation and amplitude 0.25 s−1 μm−2.

For each RGC, we used responses to full-field contrast steps to calculate an ON−OFF index, a transiency index, and a response latency index. These indices were used to characterize the properties of RGCs that is that is unique to our analyses. The ON−OFF index was calculated by dividing the difference between ON and OFF step peak response by their sum. The resulting index values ranged between −1 (OFF) and +1 (ON) and were then scaled to span between 0 (OFF) and +1 (ON). The transiency index was defined as the ratio of the response area within the first 1700 ms after the total response time. For each individual unit, we calculated the transiency index values for a range of pixel threshold values. The threshold value was defined as the value from stimulus onset to 90% of peak response. This value was normalized to the maximum response latency in our dataset to create the response latency index.

**Modulation index.** To quantify retinal saccadic suppression, we first determined a baseline response, defined as the response to a probe flash ~2 s after texture displacement onset or 2 s after luminance step (delay between 1967 and 2100 ms, depending on the specific flash times used in a specific experiment). This baseline...
response was compared to responses of the same cell to the same flash when it occurred at an earlier time (i.e., closer in time to the saccade). Usually, the saccade-like texture displacement caused their own response/saccade-alone response, e.g., Fig. 1b), and the responses to the flashes were superimposed on these saccade-responses. We therefore first isolated the component of the responses caused by the flashes by subtracting the saccade-alone responses from the composite saccade and flash responses. We refer to this isolated component as the flash-induced component of the response component (see above for how we isolated this from the composite saccade + flash response) and \( r_p \) is the peak firing rate for the flash-induced component of the response (see above for how we isolated this from the composite saccade + flash response) and \( r_p \) is the peak firing rate for the baseline flash response (i.e., the same flash but occurring ~2 s away from any saccade; see above). Here, peak firing rate was taken as the maximum firing rate within 300 ms time window after the flash onset of the averaged response from all repetitions of a given condition (delay \( d \) or baseline) for a given RGC, across all saccades.

The modulation index for a given flash of preferred contrast (bright flash in ON RGCs and dark flash in OFF RGCs) delay \( d \) after saccade onset was calculated as \( (r_p - r)/r \). Briefly, in a given RGC, we first extracted its baseline flash response and used that as a template, including the timing of this response relative to flash onset. Then, for each flash delay \( d \) after saccade onset, we subtracted the saccade-only response from the response to the saccade-flash combination to isolate the flash-induced component of the response, which we then compared to the template. We quantified the strength of the flash-induced response component (relative to baseline) by making a least-squared-error linear fit of the template to that response component as follows:

\[
\text{response} \sim s \times \text{template} \tag{2}
\]

The linear term \( s \) of this fit was then interpreted as the relative response strength, with a value < 1 indicating a weaker response relative to the baseline flash. The modulation index was then calculated as \( (s - 1)/\sigma \), where \( s \) is the relative response strength of flash delay \( d \) from saccade onset. For preferred contrast flashes, this method and the one described in the previous paragraph produced similar modulation index values.

To quantify the modulation at a population level, we averaged the modulation indices of the individual RGCs in that population. For some analyses, we also calculated modulation indices of RGCs for each of the 39 individual saccades using the same procedure.

In some cells, individual saccades from the sequence of 39 were discarded. For example, imagine that “saccade No 3” gets discarded. This would happen when the baseline response strength (response to the probe flash with 2 s delay) after saccade No 3 is weak (specifically: peak amplitude less than 60% of the median of all 39 baseline response strengths). We did this to ensure that our modulation indices were not marred by a denominator approaching zero (e.g., if both flash and baseline responses were weak). We did, however, re-include some sequences. For example, if the probe flash with delay 100 ms after saccade No 3 triggered a strong response (specifically: peak amplitude larger than the median baseline response peak across the 39 saccades), then saccade No 3 would be re-included for the condition “100 ms delay.” This was done in order to re-include sequences (if discarded by the first step) for which the baseline flash response was weak but a flash after saccades nonetheless gave a robust response. For example, this could happen if a cell did not respond to the baseline flash but the saccade enhanced the response to a flash following it.

The modulation index was based on responses to the brief probe flashes, it could only be computed for RGCs that did respond to these brief flash stimuli. In our analysis, we included all such RGCs. Of the spike sorted RGCs across all paradigms, we included: 2002 of 3510 in mice (47 retinae); 228 of 376 in pigs. Full-field saccades. We analyzed 1010 mouse RGCs (633 ON; 377 OFF) and 228 pig RGCs (197 ON; 31 OFF) for saccade suppression using the full-field version of the saccade paradigm. A subset of data from 688 of the 1010 mouse RGCs and a subset of data from all the 228 pig RGCs was presented previously. Here, we perform further analyses on the complete datasets from the RGCs recorded previously, in addition to analyzing the newly recorded RGCs. For each RGC, we calculated the modulation index for a full-field probe flash presented at different times from saccade onset.

A subset of these RGCs were also tested for saccadic suppression while blocking GABAergic and glycinergic inhibition. For 82 ON and 30 OFF RGCs, we had a direct current given RGC, without GABAA, C blockers (5 μM SR-95531 + 100 μM Picrotoxin) (Fig. 2c and Supplementary Fig. 6a). For 51 ON and 13 OFF RGCs, we had a direct comparison with and without GABAA, C blockers in addition to glycine blocker (1 μMstrychnine; Fig. 2d and Supplementary Fig. 6b).

In yet another subset of RGCs (72 ON; 49 OFF), we had an isolated saccadic suppression at the scopic and mesopic levels. Periphery saccades (global component of suppression). For each recorded RGC, we first isolated the flash-induced response component (using 5 μM SR-95531) on the modulation of probe flash responses in RGCs with masking factors >2, i.e., with receptive field centers and immediate surround effectively masked. This included 62 ON RGCs and 35 OFF RGCs with robust responses to brief probe flashes with the pharmacological agent; only these RGCs were analyzed for saccadic suppression under the presence of SR-95531 (Fig. 2b). In a subset of these RGCs (29 ON; 25 OFF), we had a direct comparison of modulation indices with the control condition, i.e., periphery saccades in the absence of SR-95531 (Supplementary Fig. 5d).

Checkerboard mask regions saccades. In experiments where we investigated the local component of suppression (Fig. 2e, f and Supplementary Fig. 7b, c), we had three stimulus presentation settings (Supplementary Fig. 7a): (1) saccade and flash presented in all squares regions; (2) saccades and flashes in alternate regions of a checkerboard; and (3) saccade and flash presented in the same region (Supplementary Fig. 7b). For each recorded RGC, we calculated the modulation index for the following two scenarios: saccades and flashes in all regions (Supplementary Fig. 7a presentation setting 1); and saccades excluded from the receptive field center (Supplementary Fig. 7a presentation setting 1 or 2 depending on the RGCs spatial position relative to the presentation area). For this, we first calculated the spatial receptive field center (see heading “RGCs character-ization”); receptive field center was defined as the 1-ellipse of the 2D Gaussian fit) for each RGC. We then calculated the Euclidean distance between the receptive field center and the closest flash region for both presentation settings 2 and 3. For further analysis, we used the presentation setting with the shortest distance between a flash region and the receptive field center. In some cases, the closest flash region was not perfectly centered over the receptive field center and will be excluded from a region of at most 300 μm diameter centered over the receptive field center. In most cases, the RGCs were indeed well centered within a flash region since the center of each region coincided with electrodes in the low-density MEA. Nonetheless, for each RGC we calculated the intersection between its receptive field area and saccade regions in pixels (where each pixel corresponded to 3.75 μm on the retinal surface). RGCs for which more than 15% of their receptive field area intersected with the saccade regions were excluded from further analysis. In the end, a total of 51 RGCs (32 ON; 38 OFF) from 4 retinae recordings were used for further analysis of saccadic suppression.

Luminance-step paradigm. To quantify retinal saccadic suppression with the luminance-step paradigm (Figs. 4 and 7), we used the same analyses and statistical procedures to those described above for the saccade (texture displacement) para- digm. The only difference was that instead of 39 successive sequences in a trial, we now had either 56 or 156 successive sequences (or 20 in case of macaque retina
Supplementary Fig. 10 shows the association index for each RGC and for which we could also compute a transiency index. The relation between RGC quantification of saccadic suppression in the presence of pharmacological blockers. We, therefore, convolved this output vector with a moving average filter, capturing the peak for the 100 ms duration probe flash, as shown in Fig. 5b.

The “stitched” signal was then cut into snippets that captured the relevant responses to our stimulus (e.g., step followed by probe flash). For this, we used stimulus trigger signals that marked the presentation of each luminance step. Each snippet was then baseline-adjusted by subtracting the average response over 800 ms prior to the luminance step. This way the output signal was 0 prior to a luminance step (Fig. 5). The output was averaged across the three repetitions of the trial. The normalized and averaged snippets represented the cone response to a particular stimulus sequence (e.g., Fig. 5a) and were used to fit parameters for model cones as described in the next section.

**Data analysis for cone photoreceptors.** We analyzed eye movements in all trials and detected saccades using established methods.25 We excluded from analysis trials in which a saccade or microsaccade happened anywhere in the interval from 200 ms before to 50 ms after a probe flash. At each flash time, we calculated the proportion of correct trials to obtain time courses of this perceptual measure. We obtained time course curves for each subject individually and then averaged it across trials and different contrasts of the luminance steps. Reduced proportion of correct trials at any flash time indicated perceptual saccadic suppression. A subset of data from four of the five subjects was used in our previous study4. Here we perform further analyses of the complete dataset.

We applied a two-tailed Wilcoxon rank-sum test to determine if the suppression after luminance steps differed across bright and dark probe flashes.

### Computational model of retinal ganglion cells.

To describe cone responses (Fig. 5) and RGC responses (Fig. 6) to our luminance-step paradigm, we used a phenomenological model of the retina, previously published in ref. 15.

The original model related light intensity to retinal ganglion spiking activity by three layers of processing: first, the light stimulus was passed to model cone photoreceptors. Their activity was modulated by negative feedback from model horizontal cells. Second, the output of the model cones was passed to six inner retina pathways describing retinal processing by three different ON and three different OFF bipolar cells (fast, intermediate and slow pathways). Third, the output from the model pathways were then fed into model RGCs to yield RGC spiking activity. This cascade modeled RGC spiking in response to the light stimulus passed to model cones.

The cone responses were described as

\[
\mathbf{r(t)} = \mathbf{a}_c(t) - \mathbf{b}(t)
\]

where

\[
\mathbf{r(t)} = \mathbf{V(t)} - \mathbf{V}_{dark} \tag{4}
\]

\[
\mathbf{V(t)} - \mathbf{V}_{dark} = \text{instantaneous and dark membrane potentials of the cone, respectively, } \mathbf{b}(t) = \text{the feedback signal from the horizontal cell, and } \mathbf{a}_c \text{ and } \mathbf{b} \text{ were numerical factors. The time-varying functions } \gamma(t) \text{ and } \zeta(t) \text{ were related to light input through linear convolution, as}
\]

\[
\gamma(t) = \int_{t-\tau}^{t} K(t-\tau) \iota(\tau) d\tau \tag{5}
\]

\[
\zeta(t) = \int_{t-\tau}^{\infty} K(t-\tau) \iota(\tau) d\tau \tag{6}
\]

where \(\iota(t)\) was the incident light intensity (or, more precisely, \(R^{\nu-1}\)). The kernels describing the cone response were given by

\[
K(t) = \frac{t - \tau}{\tau^\nu} \tag{7}
\]

and

\[
\mathbf{K}(t) = \mathbf{e}^{-t - \tau} \tag{8}
\]

where \(\gamma(t)\) was larger than \(\tau_t\), and \(0 \leq \tau_s \leq 1\) ensured proper normalization. Note that \(d\tau K(t-\tau) = 1\) for all filters. The response of the horizontal cell was described
by

\[ h(t) = a_t \int_{-\infty}^{t} K_{t}(t'-t) \alpha(t') dt' \]  

(9)

with

\[ K_{t}(t) = \frac{t e^{-\frac{t}{\tau_1}}} {\tau_1} \]  

(10)

Here, instead of the cone model parameters used in the published model, we refitted the parameters of the model cone to reflect our measured data of cone output (Fig. 5b) which yielded faithful fits (Supplementary Fig. 11) to our experimentally measured cone responses.

All the parameters of the outer retina component (Eqs. (3) and (10)) of our circuit model were fit once to cone responses (Fig. 5b and Supplementary Fig. 11) and then kept unchanged for all simulations reported in Fig. 6. The fitted values are given as follows (original values are reported in brackets):

- \( \alpha_{c} = -3.342 \times 10^{-5} \times -9.602 \times 10^{-4} \),
- \( \beta_{c} = 1.273 \times 10^{-4} \times -1.148 \times 10^{-3} \),
- \( \gamma = 0.842 \times 0.763 \),
- \( \alpha_{k} = 0.016 \times 0.177 \),
- \( t_{p} = 48.98 \mathrm{ms} (50.64) \),
- \( t_{r} = 200 \mathrm{ms} (576.9) \),
- \( t_{n} = 1.232 \times 10^{3} \mathrm{ms} (371) \).

The small \( \alpha_{k} \) suggest that in our cone recordings, horizontal cell feedback had only a minor effect.

In ref. 43, three different retinal pathways were modeled according to

\[ b_{x},(t) = \left[ -1 \left( \int_{-\infty}^{t} K_{x}(t'-t) \alpha(t')dt' - \theta_{x} \right) \right] \]  

(11)

where \( p, q, 2, 3 \) labeled the pathway based on its response properties (1 = fast, 2 = intermediate, 3 = slow), \( k = 0 \) for OFF pathways and \( k = 1 \) for ON pathways.

\[ |x| = \{ 0, x > 0 \} x \geq 0 \]  

(12)

was a thresholding nonlinearity, and \( \theta_{x} \) acted as a threshold.

The main difference between the pathways was the temporal characteristics of the filters \( K_{x} \). In this study, we wanted to smoothly vary the transparency of the model ganglion cells. To this end, we based our bipolar pathway on the fast bipolar pathway (\( p = 1 \)) and modified its temporal characteristics to make it less transient. \( K_{t} \) represented a high-pass filter which took the derivative of the cone potential on the order of 1 ms. We obtained \( K_{t} \) by convolving the high-pass filter of the form

\[ G_{t}(t) = \sin \left( \frac{\pi t}{\mu} \right) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{t^2}{2\sigma^2}} \]  

(13)

with an exponential function.

\[ K_{t}(t) = \int_{-\infty}^{t} \left( e^{-\frac{t'}{\tau_2}} \right) G_{t}(t') dt' \]  

(14)

Higher values of the time constant, \( \tau_2 \), of the exponential function decreased the pathway transience. We set \( \tau_2 = 0.5 \mathrm{ms} \) as the default transient pathway (Fig. 6a–d). The filter transience values shown in Fig. 6g were obtained by normalizing 15 ms > \( \tau_2 \) > 0.5 ms in the range 0 to 1, with 0 being less transient. In the original model, \( K_{t}(t) = G(t) \).

The threshold \( \theta_{c} \) was set to \( 1 - 10^{-4} \) – 0.1, except in Fig. 6h, where we varied its value between 0 and 1, while keeping \( t_{r} \) fixed at a value of 15 ms (0 or in the normalized scale). This change in nonlinearity was another way to change response properties of the pathway. Note that the threshold values are not on a normalized scale.

\[ b_{x},(t) \]  

(11)

the output of fast inner retina models was used as the input to the model RGCs used in this study. The spiking rate of the model RGC was obtained as the thresholded input to a temporally coarse version of the input’s derivative,

\[ R_{t}(t) = \left[ 1 - \alpha_{1} b_{1}(t) + \sigma \right] \int_{-\infty}^{t} K_{t}(t'-t) b_{1}(t') dt' - \theta \]  

(15)

where \( K(t) \) was a biphasic filter similar in its form to \( G(t) \). The threshold, \( \theta \), was a multiple of the peak response to any given input. We used the same parameters for the inner retina component as the ones described in the published model.

(i) Transient OFF. \( R_{t}(t) \) : \( a = 0, \theta = 0 \);

(ii) Transient ON. \( R_{t}(t) \) : \( a = 0, \theta = 0 \);

All simulations were computed with a 1 ms sampling interval.

Using the above cascade, we calculated model RGCs’s spike rate in response to the recorded cone output when subjected to “luminance steps alone” and “luminance steps followed by probe flash” stimuli (Fig. 6a, b). Similar to real RGC analysis, we calculated the “flash-induced responses” (Fig. 6c) by subtracting “response to luminance steps alone” (Fig. 6a) from “response to luminance steps followed by probe flash” (Fig. 6b). We then calculated the modulation index also in the same way as for the real RGCs: \( \tau_{z} = \tau_{r} = \tau_{n} + \tau_{n} \) where \( t_{r} \) was the peak spiking rate of the flash-induced response for the flash presented with delay \( d \) from the luminance step, and \( t_{n} \) was the peak firing rate for the baseline flash-induced response (flash at 2000 ms).

As a control, we replaced the model cone responses with the experimentally acquired cone responses (Fig. 5) thus forming a hybrid model. Before passing the cone response to the downstream model pathways, we passed it through a low-pass filter to further smooth the fluctuations at the stitching boundaries in order to avoid discontinuities in the calculation of its temporal derivative. For this smoothing, we convolved the cone output with a moving average filter of size 40 ms. Here, the model ON RGC responses (Supplementary Fig. 12a, b, columns 1–2) were calculated using the cone responses to bright probe flashes (Fig. 5a, b, columns 1–2), and OFF RGC responses (Supplementary Fig. 12a, b, columns 3–4) were calculated using cone responses to dark probe flashes (Fig. 5a, b, columns 3–4). The resulting hybrid model RGC responses (Supplementary Fig. 12) were consistent with the pure model responses shown in Fig. 6.

**Statistics and reproducibility.** We applied statistical tests both at the level of individual RGCs and at the level of the population to determine whether the response to a probe flash following a saccade or a luminance step was significantly modulated. At the individual cell level, we determined whether the modulation index for a probe flash presented at a certain delay was significantly different from 0 (i.e., “Is the response of this cell modulated by the ‘saccade’?”). For this, we performed a one-tailed sign test of the null hypothesis that the 39 individual modulation indices (or its subset in case weak sequences were discarded, as described above) came from a distribution with zero median. The alternative hypothesis was that the median was below (for negative modulation index) or above (for positive modulation index) zero. The modulation index was considered significant (i.e., the flash response was modulated by the saccade) at \( P < 0.05 \).

However, we did not consider cells significantly modulated if the test had a power (1-\( \beta \)) smaller than 0.8, which could happen if they previously had to exclude too many sequences (\( N \leq 39 \)). At the population level, we determined whether the retinal output as a whole was modulated by saccades. For this, we performed a two-tailed Wilcoxon signed-rank test of the null hypothesis that the median of the distribution of modulation indices did not differ from 0. Lastly, we tested whether the population modulation index was significantly different across populations of ON and OFF RGCs or across different paradigms. For this, we performed a two-tailed Wilcoxon rank-sum test of the null hypothesis that the median of the distribution of modulation indices did not differ across the two populations being tested.

Since the modulation index was based on responses to the brief probe flashes, it could only be computed for RGCs that did respond to these brief flash stimuli. In our analysis, we included all such RGCs. Of the spike sorted RGCs across all paradigms, we included: 2002 of 3510 in mice (47 retinae); 228 of 376 in pigs (pieces from 12 retinae); and 15 of 57 in macaque (1 retina).

In human psychophysics experiments, we applied a two-tailed Wilcoxon rank-sum test to determine if the suppression after luminance steps differed across bright and dark probe flashes (\( N = 5 \) subjects).

All data analyses were performed in MATLAB (The MathWorks Inc).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The source data underlying the figures in this manuscript are available in the public repository [https://doi.org/10.5281/ZENODO.6562460](https://doi.org/10.5281/ZENODO.6562460) (ref. 79). All other data are available upon reasonable request.

**Code availability**

The code and source data underlying the figures in this manuscript are available in the public repository [https://github.com/saadidrees/saccadic-suppression](https://github.com/saadidrees/saccadic-suppression) with identifier [https://doi.org/10.5281/ZENODO.6562460](https://doi.org/10.5281/ZENODO.6562460) (ref. 79). All other code used for analysis and computational modeling is available upon reasonable request.

Received: 30 January 2022; Accepted: 23 May 2022; Published online: 12 July 2022

**References**

1. Yarbus, A. L. *Eye Movements and Vision* (Springer US, 1967).

2. Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M. & Velichkovsky, B. M. Yarbus, eye movements, and vision. *Ipercognition* 1, 7–27 (2010).

3. Schütz, A., Braun, D. & Gegenfurtner, K. Eye movements and perception: a selective review. *J. Vis.* 11, 1–30 (2011).
Wurtz, R. H. Neuronal mechanisms of visual stability. *Nat. Commun.* 11, 1977 (2020).

Zuber, B. L. & Stark, L. Saccadic suppression: elevation of visual threshold associated with saccadic eye movements. *Exp. Neurol.* 16, 65–79 (1966).

Zakut, A. P. & Meister, M. Segregation of object and background motion in the retina. *Nature* 423, 401–408 (2003).

Zwass, V. & Adel, W. R. Retinal ganglion cells of the cat transfer information on saccadic eye movement and quick target motion. *Brain Res. Biot.* 70, 340–345 (1974).

Passaglia, C. L., Freeman, D. K. & Troy, J. B. Effects of remote stimulation on the modulated activity of cat retinal ganglion cells. *J. Neurosci.* 29, 2467–2476 (2009).

Roska, B. & Werblin, F. Rapid global shifts in natural scenes block spiking in specific ganglion cell types. *Nat. Neurosci.* 6, 600–608 (2003).

Lin, B. & Masland, R. H. Populations of wide-field amacrine cells in the mouse retina. *J. Comp. Neurol.* 499, 797–809 (2006).

Krekelberg, B., P. B. Baccus, S. A. & Meister, M. Segregation of object and background motion in the retina. *Nature* 423, 401–408 (2003).

Zwass, V. & Adel, W. R. Retinal ganglion cells of the cat transfer information on saccadic eye movement and quick target motion. *Brain Res. Biot.* 70, 340–345 (1974).

Passaglia, C. L., Freeman, D. K. & Troy, J. B. Effects of remote stimulation on the modulated activity of cat retinal ganglion cells. *J. Neurosci.* 29, 2467–2476 (2009).

Roska, B. & Werblin, F. Rapid global shifts in natural scenes block spiking in specific ganglion cell types. *Nat. Neurosci.* 6, 600–608 (2003).

Lin, B. & Masland, R. H. Populations of wide-field amacrine cells in the mouse retina. *J. Comp. Neurol.* 499, 797–809 (2006).

Krekelberg, B., P. B. Baccus, S. A. & Meister, M. Segregation of object and background motion in the retina. *Nature* 423, 401–408 (2003).

Zwass, V. & Adel, W. R. Retinal ganglion cells of the cat transfer information on saccadic eye movement and quick target motion. *Brain Res. Biot.* 70, 340–345 (1974).
Acknowledgements

We thank Elisabeth Gustafsson for providing technical support; Andreas Hierlemann for providing the HiDens CMOS MEA system and helping establish our high-density MEA recordings; Roland Diggelmann for helping in setting up the pipeline (including providing code) for automatic spike sorting of high-density MEA recordings; Martin Schenk at the Department of Experimental Surgery, Tübingen, for providing us domestic pig eyes. Our work was supported by funds of the Deutsche Forschungsgemeinschaft (DFG) to the Werner Reichardt Center for Integrative Neuroscience (EXC 307) and to TAM (MU3792/1-1 and MU3792/3-1). T.A.M. received support from the Tistou and Charlotte Kerstan Foundation. S.I., T.A.M., and Z.M.H. were also supported by an intra-mural funding program (Project 2013-05) of the Werner Reichardt Center for Integrative Neuroscience. F.F. was supported by a Swiss National Science Foundation Ambizione grant (PZ00P3_167989) and Swiss National Science Foundation Eccellenza grant (PCEFP3_187001). M.P.B. and Z.M.H. were further funded by the SBF 1233 on Robust Vision (DFG, project number 276693517). K.F. was supported by the Bundesministerium für Bildung und Forschung (BMBF, 01GQ1002) and the Max Planck Gesellschaft (MPG, M.F.E.A.K.Y.BE0004).

Author contributions

T.A.M. conceptualized the study; S.I. and T.A.M. designed the overall study; F.F. developed and supervised computational modeling; S.I., M.P.B., A.K., K.F., Z.M.H., F.F., and T.A.M. designed experiments; S.I. performed ex vivo mouse and pig retina experiments; A.K. performed ex vivo macaque retina experiment; M.P.B. performed human psychophysics experiments; K.F. performed viral injections; M.K. performed cone imaging experiments; S.I. and F.F. implemented computational model; S.I., M.P.B., M.K., T.S., A.K., K.F., Z.M.H., F.F., and T.A.M. analyzed and interpreted the data and wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03526-2.

Correspondence and requests for materials should be addressed to Felix Franke or Thomas A. Münch.

Peer review information This manuscript has been previously reviewed at another Nature Portfolio journal. This document only contains reviewer comments and rebuttal letters for versions considered at Communications Biology. Communications Biology thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: George Inglis. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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

© The Author(s) 2022

72. Schwartz, G. W. et al. The spatial structure of a nonlinear receptive field. Nat. Neurosci. 15, 1572–1580 (2012).
73. Zhang, Y., Kim, I.-J., Sanes, J. R. & Meister, M. The most numerous ganglion cell type of the mouse retina is a selective feature detector. Proc. Natl Acad. Sci. USA 109, E2391–E2398 (2012).
74. Sakatani, T. & Isa, T. Quantitative analysis of spontaneous saccade-like rapid eye movements in C57BL/6 mice. Neurosci. Res. 58, 324–331 (2007).
75. Itokazu, T. et al. Streamlined sensory motor communication through cortical reciprocal connectivity in a visually guided eye movement task. Nat. Commun. 9, 338 (2018).
76. Diggelmann, R., Fiscella, M., Hierlemann, A. & Franke, F. Automatic spike sorting for high-density microelectrode arrays. J. Neurophysiol. 120, 3155–3171 (2018).
77. Idrees, S., Franke, F., Diggelmann, R., Hierlemann, A. & Münch, T. A. UnitBrowser—a tool to evaluate and post-process units sorted by automatic spike sorting algorithms. Front. Neurosci. 10, (2016).
78. Tikidji-Hamburyan, A. et al. Rods progressively escape saturation to drive visual responses in daylight conditions. Nat. Commun. 8, 1–17 (2017).
79. Idrees, S. et al. [Dataset] Suppression without inhibition: how retinal computation contributes to saccadic suppression. https://doi.org/10.5281/ZENODO.6562460 (2022).