A temporal basis for predicting the sensory consequences of motor commands in an electric fish

A Kennedy, GD Wayne, P Kaifosh, K Alvina, LF Abbott, and NB Sawtell

Supplementary Modeling

Contents

Stability Analysis

We follow the approach of Williams, Roberts, and Leen[4]. The MG cell spiking rate is a function of the MG cell voltage, $f(V)$, which we linearize around the steady state value V_0 (defined below). $V(t)$ and $r^i(t)$ are periodic on a fast time scale (the EOD response), denoted by t, whereas the w^i's change on the slower timescale of many EODs, denoted by τ.

We thus obtain the following dynamical system for $V(t)$ and w^i over many EODs:

$$\frac{dV(t)}{d\tau} = \sum_i (\mathcal{E} * r_i)(t) \frac{dw_i}{d\tau} \tag{1}$$

$$\begin{aligned} \frac{dw^i}{d\tau} &= -w_0^i \Delta^- \int [f(V(t)) - f(V_0)](\mathcal{L}_0 * r_i)(t) dt \\ &= -\left.\frac{df}{dV}\right|_{V_0} w_0^i \Delta^- \int [V(t) - V_0](\mathcal{L}_0 * r_i)(t) dt + \mathcal{O}((V(t) - V_0)^2), \end{aligned} \tag{2}$$

where $V_0 = f^{-1}\big(-\Delta^+/(\Delta^- \int \mathcal{L}_0(t)\, dt)\big)$ is the voltage at which the nonassociative and associative plasticities balance[4].

To confirm that our system will converge to form a stable negative image, we next generalized a result proved in [4] for the case of a delay line basis, that negative images are stable when $\mathcal{L}_0$ has the same shape as $\mathcal{E}$, to the case of an arbitrary GC basis.

Given a voltage perturbation $V(t)$ arising from sensory input to the MG cell, we defined $\tilde{V}(t) \equiv V(t) - V_0$, linearized around $V(t) = V_0$, and substituted equation (2) into (1) to obtain the

dynamics of the voltage perturbation from V_0.

$$\frac{d\tilde{V}(t)}{d\tau} = -\frac{df}{dV}\bigg|_{V_0} \sum_i w_0^i \Delta^-(\mathcal{E} * r_i)(t) \int \tilde{V}(s)(\mathcal{L}_0 * r_i)(s)ds \tag{3}$$

In the case that the learning rule matches the shape of the EPSP, i.e. $\mathcal{L}_0 = \mathcal{E}$, we have

$$\frac{d\tilde{V}(t)}{d\tau} = -\frac{df}{dV}\bigg|_{V_0} \sum_i w_0^i \Delta^-(\mathcal{E} * r_i)(t) \int \tilde{V}(s)(\mathcal{E} * r_i)(s)ds$$

We evaluated the stability by taking the inner product of a displacement from the equilibrium with the derivative.

$$\begin{aligned}\int \tilde{V}(t)\frac{d\tilde{V}(t)}{d\tau}dt &= -\frac{df}{dV}\bigg|_{V_0} \sum_i w_0^i \Delta^- \int \tilde{V}(t)(\mathcal{E} * r_i)(t)dt \int \tilde{V}(s)(\mathcal{E} * r_i)(s)ds \\ &= -\frac{df}{dV}\bigg|_{V_0} \sum_i w_0^i \Delta^- \left(\int \tilde{V}(t)(\mathcal{E} * r_i)(t)dt\right)^2 \leq 0 \end{aligned} \tag{4}$$

We thus see that any voltage perturbation within the subspace spanned by the EPSP-convolved firing rates will decay back to the equilibrium (note that Δ^- and all w^i are positive). Voltage perturbations outside of this subspace will be left unaltered by the learning rule.

Learning Dynamics Analysis

Discretizing t and s, Equation 3 may be written

$$\frac{d\tilde{V}(t)}{d\tau} = \sum_s \mathbf{A}(t,s)\tilde{V}(s) \tag{5}$$

where

$$\mathbf{A}(t,s) = -\frac{df}{dV}\bigg|_{V_0} \sum_i w_0^i \Delta^-(\mathcal{E} * r_i)(t)(\mathcal{L}_0 * r_i)(s)$$

The matrix $\mathbf{A}$ determines the dynamics by which the voltage perturbation $\tilde{V}$ decays due to learning. Eigenvectors of $\mathbf{A}$ reflect temporal patterns which can be cancelled by the GC basis $\{r_i\}$; we subsequently refer to these as eigenmodes. The eigenvalue corresponding to each eigenmode reflects the rate at which that pattern is cancelled by the basis, also referred to as its rate of decay.

Learning Rate Normalization

On a given trial, the sensory input to the MG cell is a noisy observation $\tilde{h}(t,\tau)$ of the true EOD-driven input $h(t)$:

$$\tilde{h}(t,\tau) = (\mathcal{E} * h)(t) + (\mathcal{E} * \eta)(t,\tau) - V_0,$$

where η is the sensory observation noise that reflects sensory signals and spiking noise uncorrelated with the EOD. Plugging into the linearized weight dynamics,

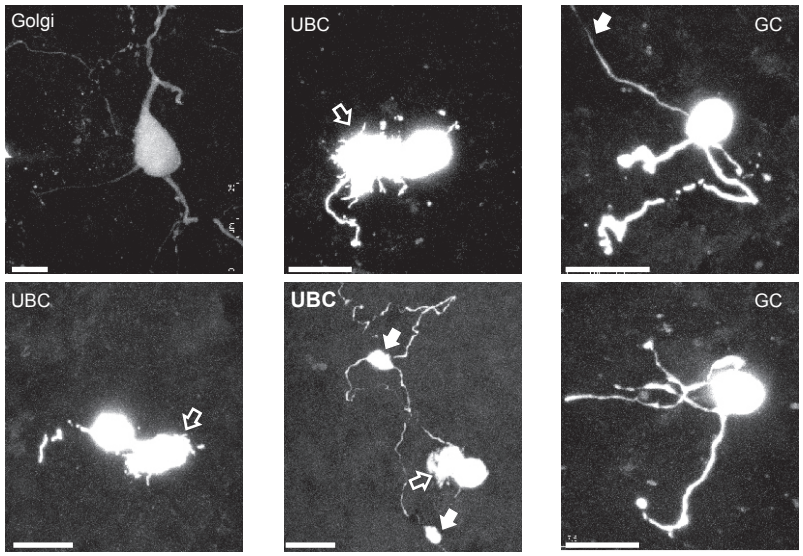
$$\frac{dw^i}{d\tau} = -\frac{df}{dV}\bigg|_{V_0} w_0^i \Delta^- \bigg(\int ((\mathcal{E} * h)(s) - V_0)(\mathcal{L}_0 * r_i)(s) ds + \int (\mathcal{E} * \eta)(s, \tau)(\mathcal{L}_0 * r_i)(s) ds \bigg)$$

The first of these terms decays to zero provided $(\mathcal{E} * h)(t)$ is in the span of $\{(\mathcal{E} * r_i)(t)\}$. The second term does not decay, but rather introduces a small trial-to-trial fluctuation in the weights w^i due to interaction of $\eta(t, \tau)$ with the learning rule.

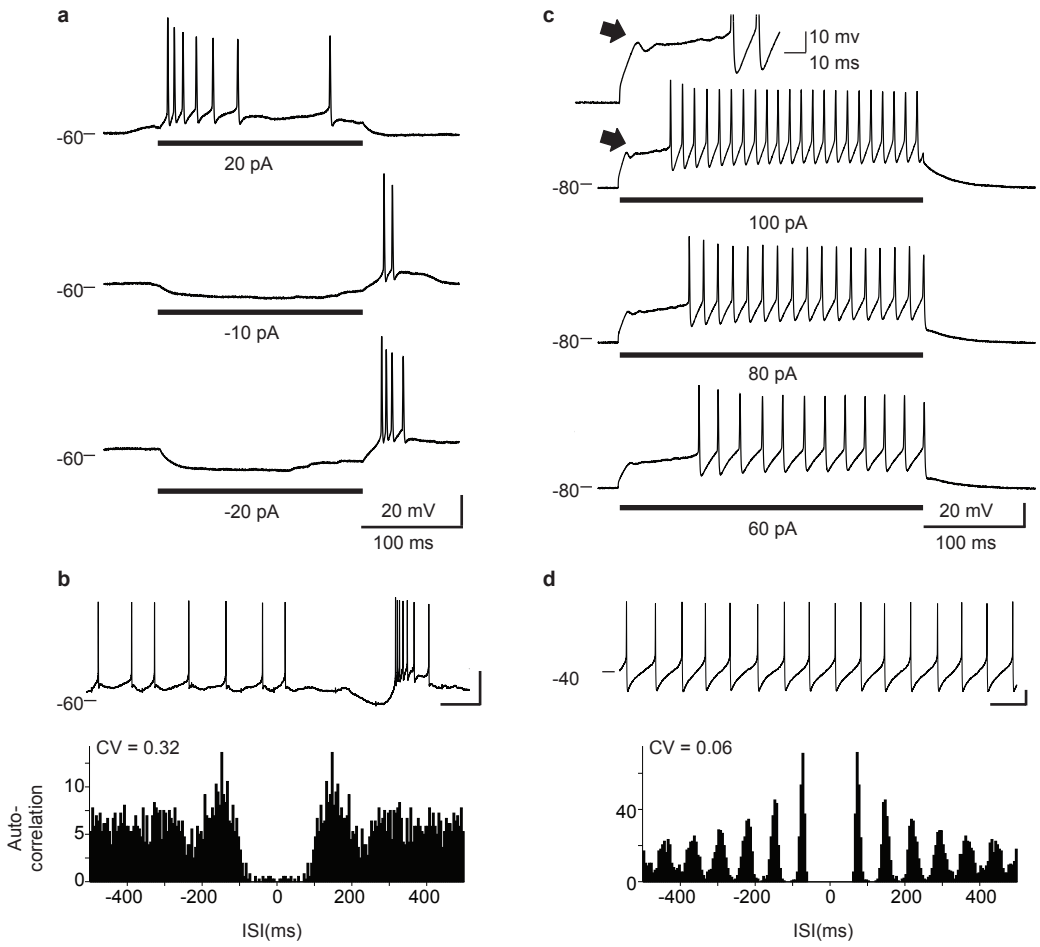
We next introduce a positive learning rate λ, substituting $\Delta^+ \to \lambda\Delta^+$ and $\Delta^- \to \lambda\Delta^-$ in the learning dynamics (note that introducing λ does not change the stability of learning or the value of V_0 provided λ is small). The weight dynamics become

$$\frac{dw^i}{d\tau} = -\lambda\frac{df}{dV}\bigg|_{V_0} w_0^i \Delta^- \bigg(\int ((\mathcal{E} * h)(s) - V_0)(\mathcal{L}_0 * r_i)(s) ds + \int (\mathcal{E} * \eta)(s, \tau)(\mathcal{L}_0 * r_i)(s) ds \bigg)$$

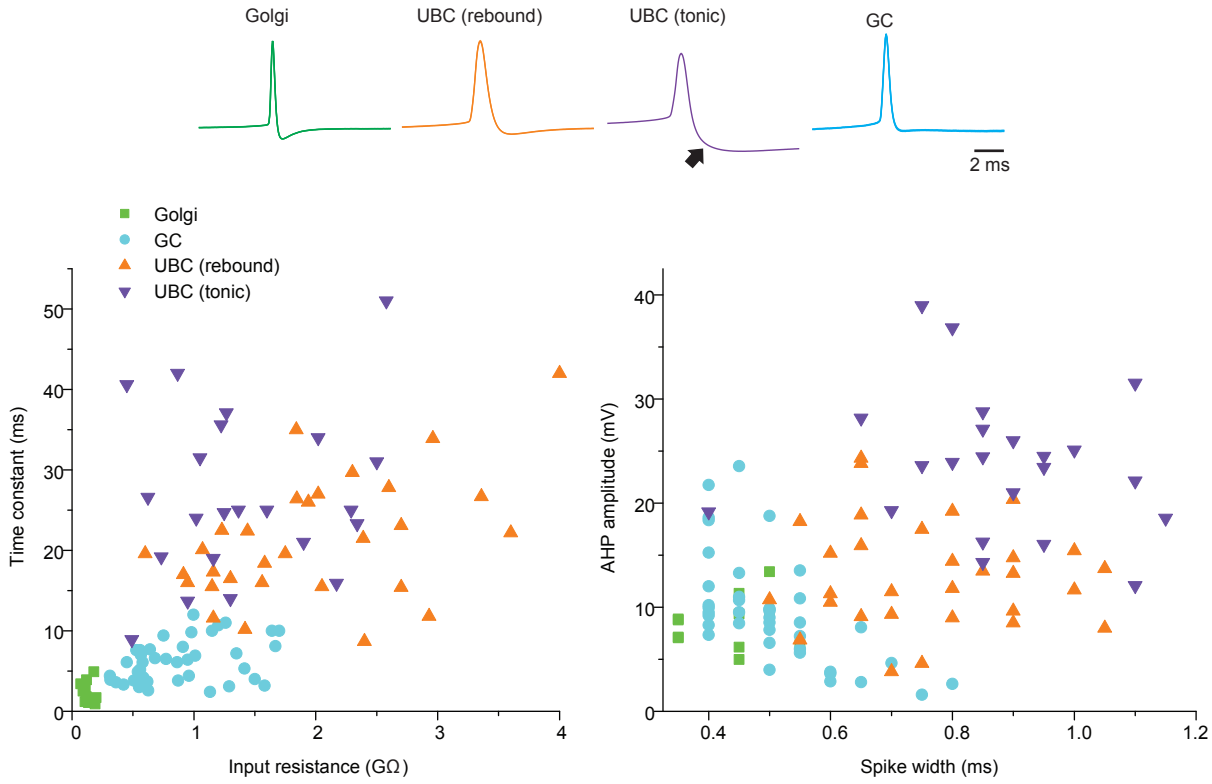
Smaller values of λ decrease the amplitude of the noise-induced weight fluctuations, but also increase the number of trials required for negative image formation. Thus given some assumption on the structure of η (such as its power spectrum), we may choose λ for a given basis such that the magnitude of noise-induced weight fluctuations is fixed.



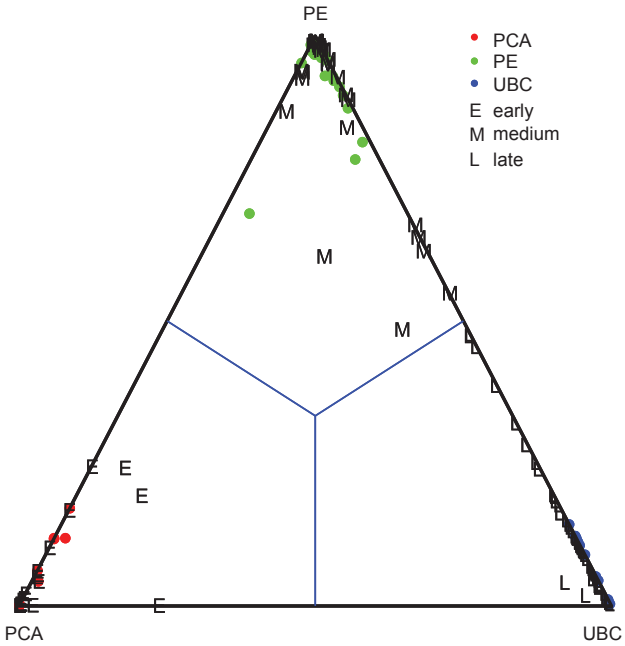
Supplementary Figure 1. Morphology of intracellularly recorded Golgi cells, UBCs, and granule cells. Maximum intensity projections of cells labeled with biocytin during in vivo whole-cell recording. Fixed sections were cut at 50 microns on a cryostat, reacted with a streptavidin-conjugated fluorescent dye, and imaged on a confocal microscope. All scale bars are 10 microns. Filled Golgi cells (n = 3) were multipolar and had cell bodies that were large relative to granule cells or UBCs. Filled UBCs (n = 11) had small ovoid cell bodies, a single distinctive filamentous brush-shaped process (open arrows) and in some cases an additional long, winding process with multiple swellings (filled arrows), similar in appearance to mossy fiber axons. Filled granule cells (GC) (n = 20) had small cell bodies and 2-5 short processes with claw-like endings. An additional thin process (filled arrow), presumably the axon, could often be traced to the ELL molecular layer.



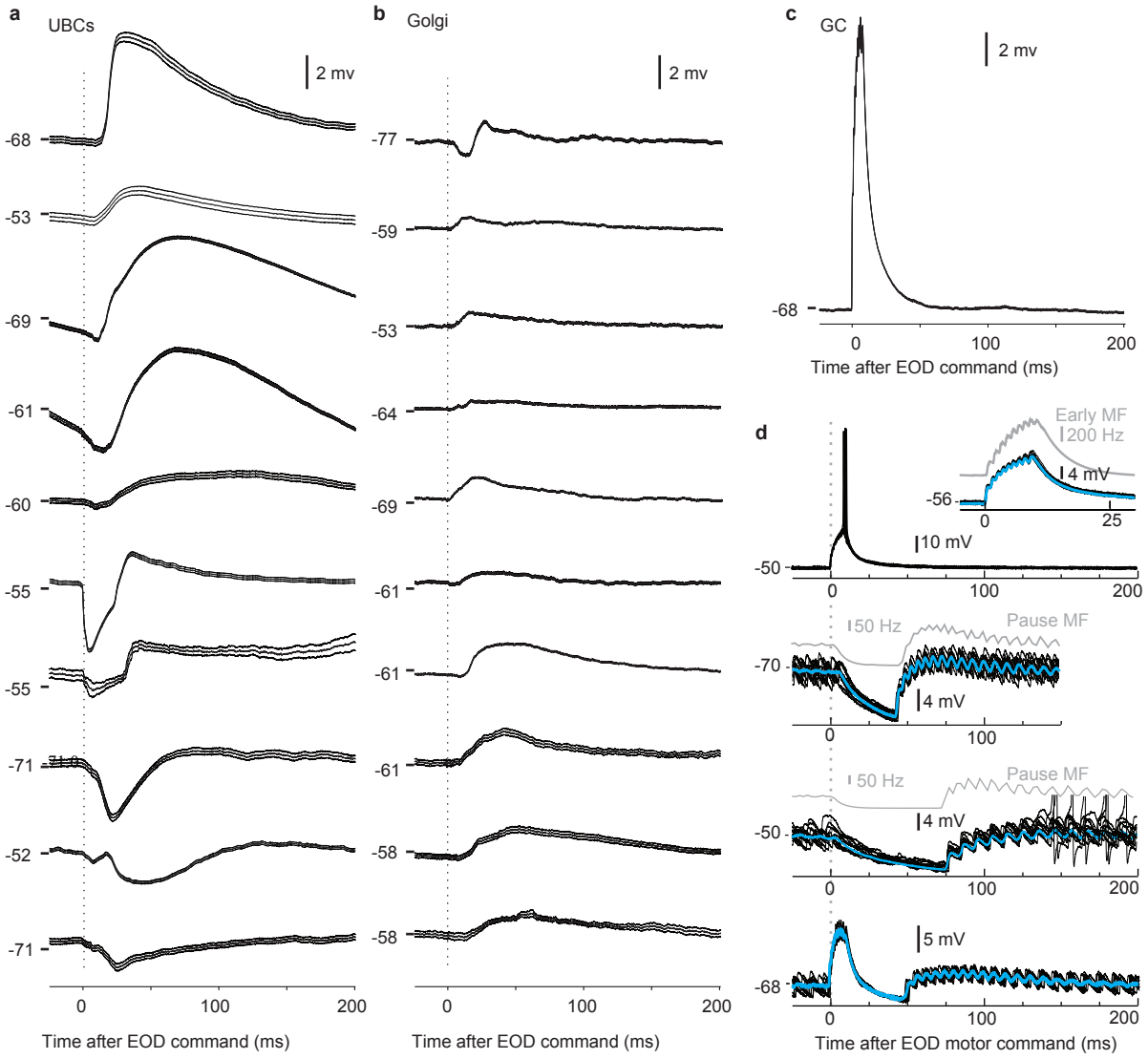
Supplementary Figure 2. UBCs exhibit characteristic electrophysiological properties. UBCs could be distinguished from Golgi cells and granule cells based on characteristic responses to intracellular current injections and, for a subset of UBCs, highly regular tonic firing. **a,** A subset of UBCs (30/54) exhibited prominent rebound firing in response to hyperpolarizing current injections. Such prominent rebound firing was never observed in granule cells or Golgi cells. Depolarizing current injections often evoked bursts of spikes followed by a decrease in firing rate (**top trace**). Five UBCs of this type were identified morphologically. **b,** Spontaneous firing in UBCs that exhibited prominent rebound tended to be irregular. Scale bar: 20 mV and 200 ms. **c,** Other UBCs (22 of 54) lacked prominent rebound firing in response to hyperpolarizing current injection (not shown). In response to depolarizing current injections such cells exhibited a delayed onset to spike firing accompanied by a distinctive "hump" and "sag" pattern during the delay (**arrow** in **c**). This pattern of response to current injections was never observed in Golgi or granule cell recordings. **d,** UBCs that lacked rebound and displayed a delayed onset to firing also exhibited extremely regular tonic firing (average CV = 0.139; n = 22). Similar regular tonic firing patterns were seen in pause mossy fibers recorded extracellularly in EGp. Such highly-regular firing was never observed in granule cells or Golgi cells. Four UBCs of this type were morphologically identified. Scale bar: 10 mV and 100 ms. Two additional UBCs were identified based on morphology alone.



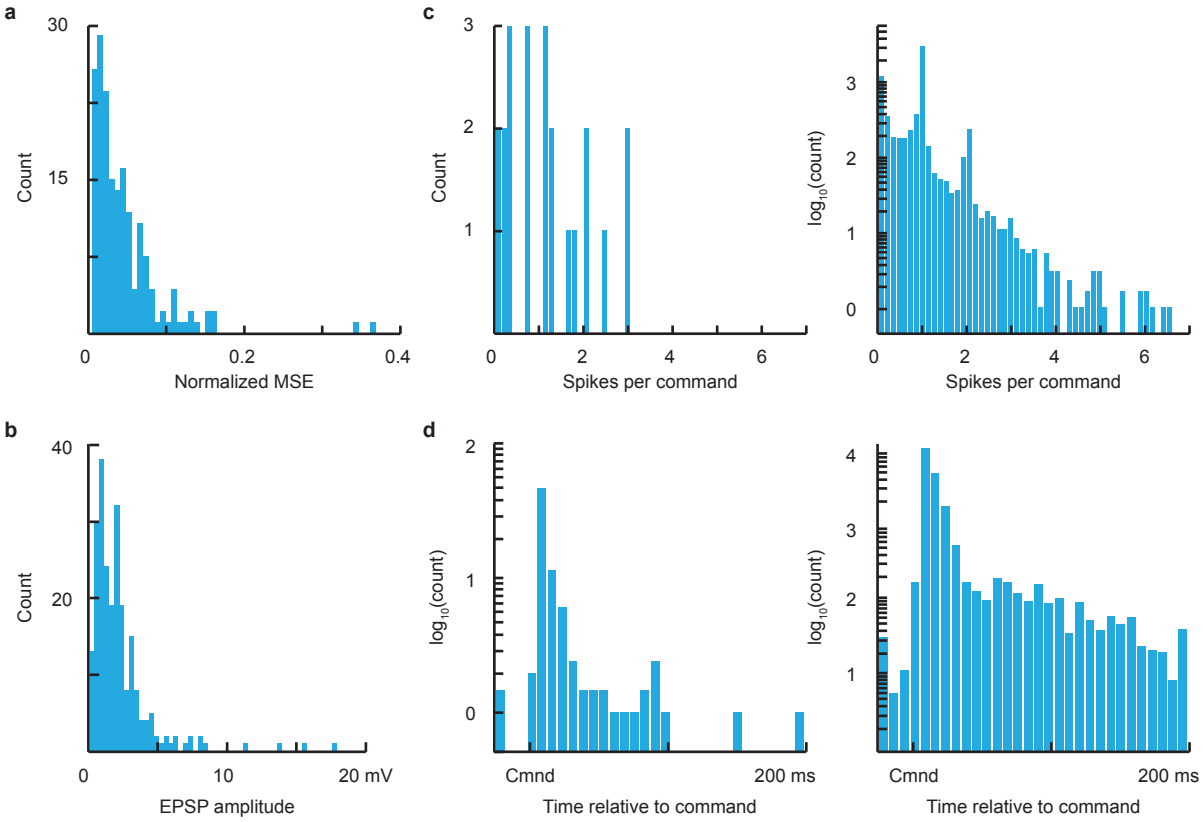
Supplementary Figure 3. Basic electrophysiological properties of Golgi cells, UBCs, and granule cells. Basic electrophysiological properties were analyzed for all intracellularly recorded Golgi cells (n = 11), UBCs (n = 54), and a subset of recorded granule cells (GC) (n = 40). Input resistance and membrane time constants were measured at the resting membrane potential in response to 10-50 pA hyperpolarizing current pulses. Spike width was taken as the width at half-height, with threshold defined as the point at which the second derivative of the voltage waveform reaches 50% of its maximum. Afterhyperpolarization (AHP) amplitude was defined as the distance between threshold and the most hyperpolarized potential following the action potential peak. Golgi cells exhibited much lower input resistance and a shorter membrane time constants than UBCs or granule cells (left, $P < 0.0001$; two-tailed Student's t-test), allowing them to be identified unambiguously based on these characteristics alone. UBCs exhibited substantial variation in their intrinsic properties, but on average they had longer time constants, broader spikes, and larger AHPs than Golgi cells or granule cells. Consistent with the possibility that UBCs comprise distinct functional classes, UBCs with regular tonic firing exhibited action potentials with significantly larger AHP amplitudes than UBCs exhibiting rebound firing (**arrow, top traces,** $P < 0.0001$; two-tailed Student's t-test).



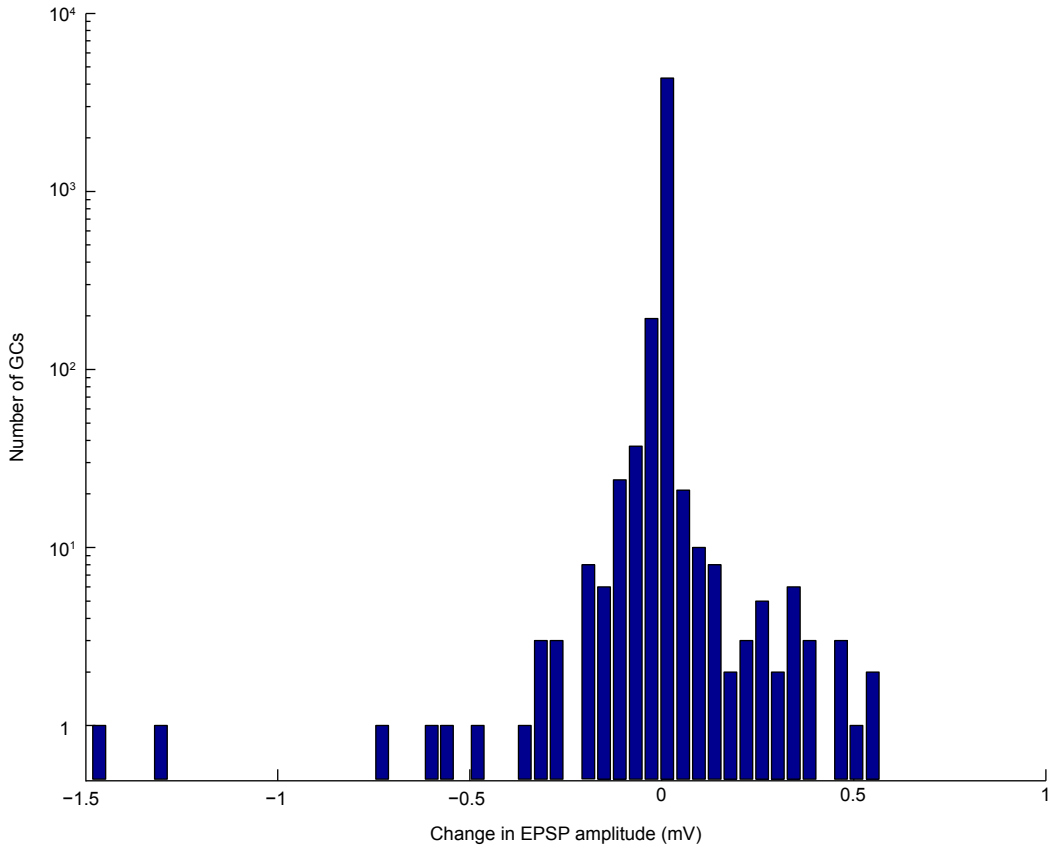
Supplementary Figure 4. Objective classification of extracellular mossy fiber recordings. Putative mossy fibers recorded extracellularly in EGp were subjected to a classifier analysis by multinomial logistic regression to determine the similarity of their corollary discharge responses to those recorded from identified sources. The sources were neurons recorded from two midbrain nuclei (PCA and PE) known from anatomical studies to provide mossy fiber input to EGp and intracellular recordings from UBCs, which are thought to give rise to a system of intrinsic mossy fibers. Three features of the corollary discharge spiking responses of the identified neurons were used as input to the classifier (see Methods). The classifier was trained to predict the probabilities of the class-labels PCA, PE, and UBC from these features. EGp mossy fiber spike trains were then converted to this feature representation and classified by the previously-trained classifier. EGp mossy fibers classified as most similar to PCA, PE, and UBC were called early, medium, and late mossy fibers, respectively. The blue lines represent the discrimination boundaries of the classifier. Pause mossy fibers were treated separately, as they were unique among EGp mossy fibers in exhibiting both a corollary discharge response (i.e. a pause) and regular, tonic firing (>4 Hz) independent of the EOD command. UBCs were the only other identified neural elements that shared these properties.



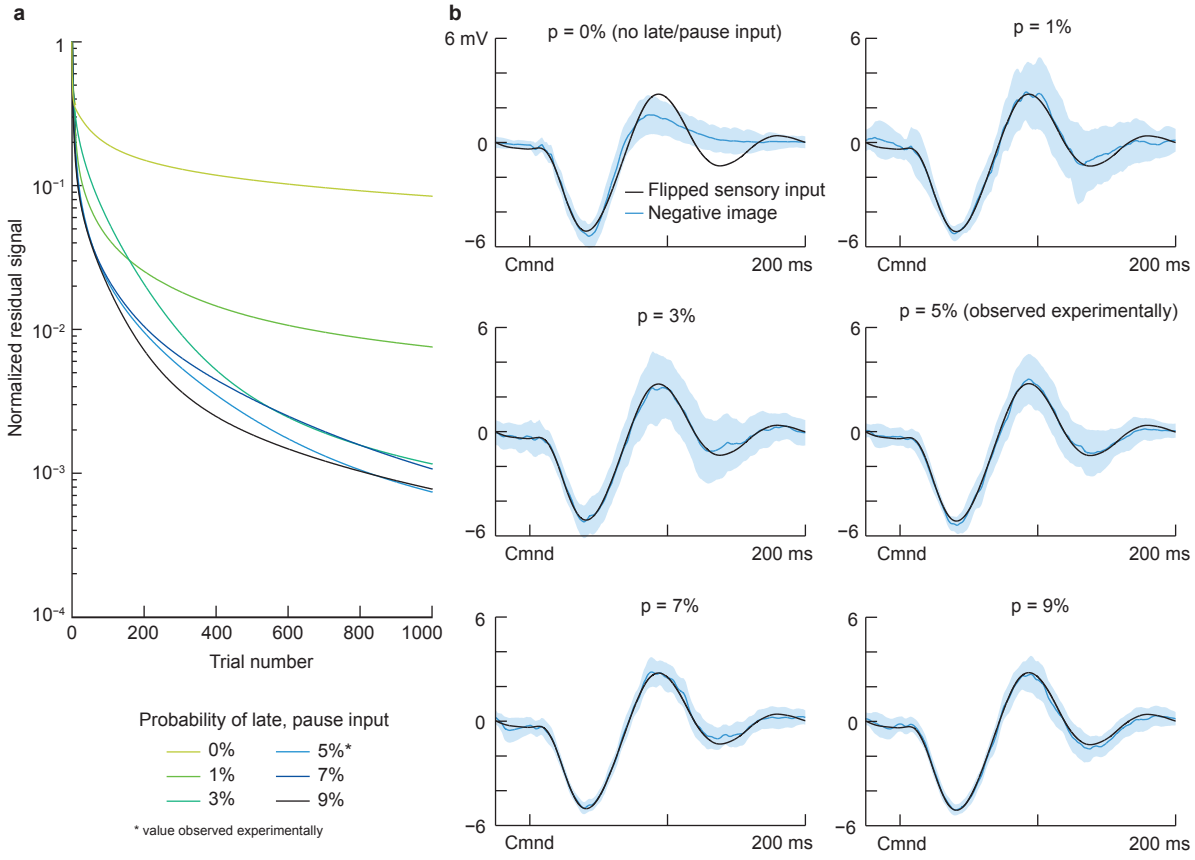
Supplementary Figure 5. Subthreshold corollary discharge responses in UBCs, Golgi cells, and granule cells. Average and SEM of subthreshold corollary discharge responses for ten representative UBCs (**a**) and Golgi cells (**b**). For comparison, a single representative granule cell (GC) response is shown in **c** on the same scale. In some cases a small amount of hyperpolarizing current was used to prevent spiking. UBCs often exhibited depolarizations that were much more prolonged than those seen in granule cells or Golgi cells (top 5 traces in **a**). Hyperpolarization time-locked to the EOD command (bottom 5 traces in **a**), suggestive of synaptic inhibition, was also much more commonly observed in UBCs than in Golgi cells or granule cells. **d**, Temporal features of granule cell corollary discharge responses closely resemble firing patterns of individual mossy fibers (MFs) and UBCs. Each panel shows the responses of a granule cell to ten consecutive EOD commands overlaid (black traces) along with the average taken from a larger number of commands (blue).The gray traces above are average command-evoked firing rate for selected mosy fibers recorded extracellularly in EGp smoothed with an exponential filter with width (10 ms) similar to the average granule cell membrane time constant. From top to bottom: granule cell with a subthreshold response that resembles temporal firing patterns characteristic of early mossy fibers, as seen in the inset. Timing of inflections on the rising phase of the granule cell EPSP closely match the timing of mossy fiber spikes. Middle, granule cells with subthreshold responses resembling temporal firing patterns characteristic of pause mossy fibers, spikes are truncated. Time-locking of mossy fiber spikes after the pause matches the periodicity of granule cell membrane potential changes. Bottom, granule cell with subthreshold response that appears to reflect a combination of an early response and a pause response. Individual mossy fibers or UBCs with firing patterns resembling such mixed responses were never observed, suggesting that this granule cell integrates two different corollary discharge inputs.



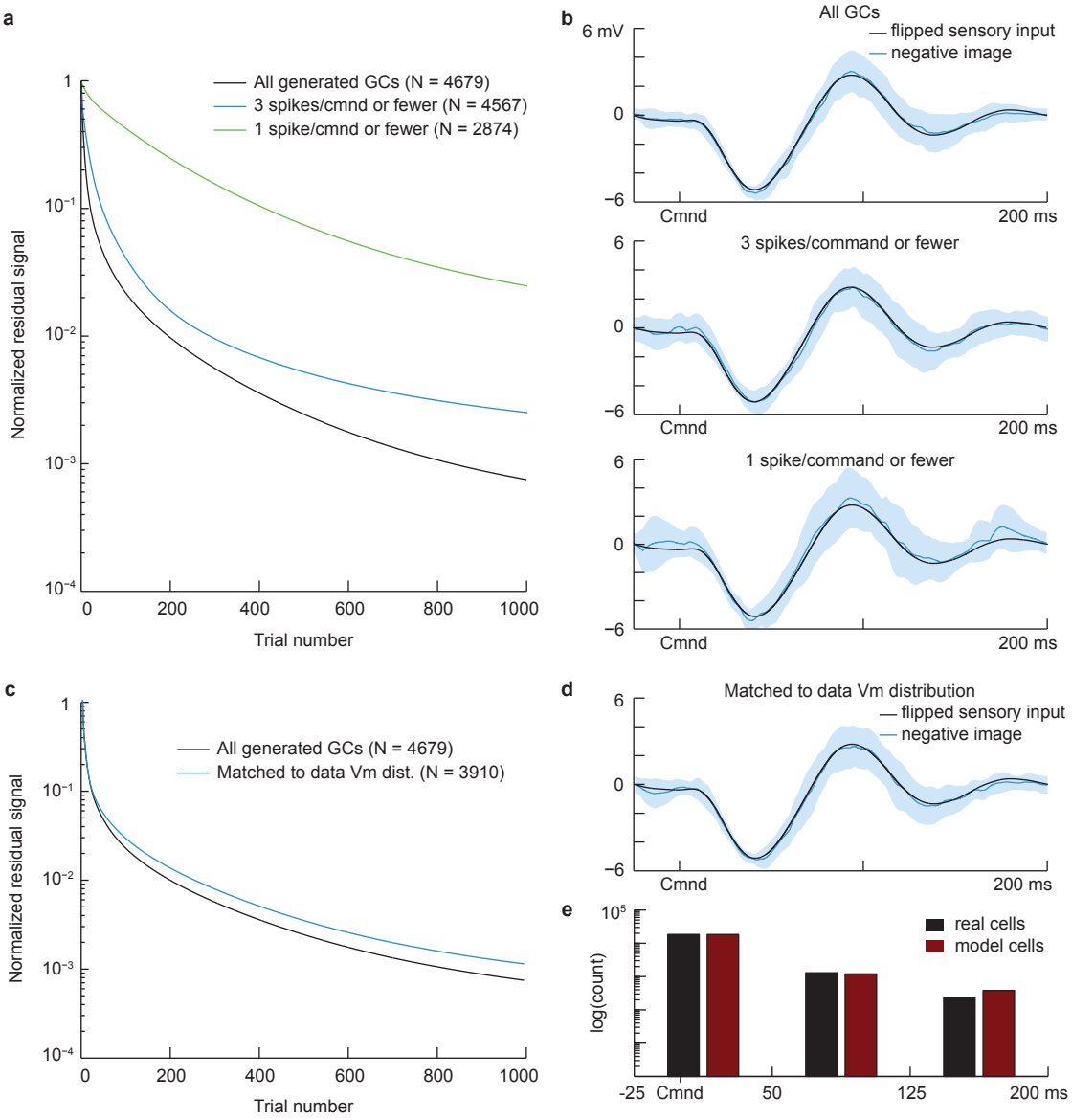
Supplementary Figure 6. Properties of granule cell model fits and generated granule cells. **a,** Normalized MSE of model fits to granule cell responses; average MSE was 4.61%. **b,** Histogram of EPSP amplitudes in fit mossy fiber - granule cell synapses. This distribution is in line with the range of EPSP amplitudes observed in recorded granule cells. **c,** average number of spikes fired per EOD command in experimentally recorded granule cells (left), and in 20,000 generated model granule cells (right); note log scale of y-axis on right. As in recorded granule cells, the model cells fired a small number of spikes per command and none of the model cells were tonically active. A comparison of the two samples using a Kolmogorov-Smirnov test (granule cells firing zero spikes per command were included in the analysis but not shown in the display) failed to reject the null hypothesis that the two samples were drawn from the same distribution. **d,** Histogram of the time relative to the command of maximum depolarization for recorded (left) versus model (right) granule cells. The similarity of these distributions again confirms that the model granule cells we generated are very similar to the granule cells recorded experimentally.



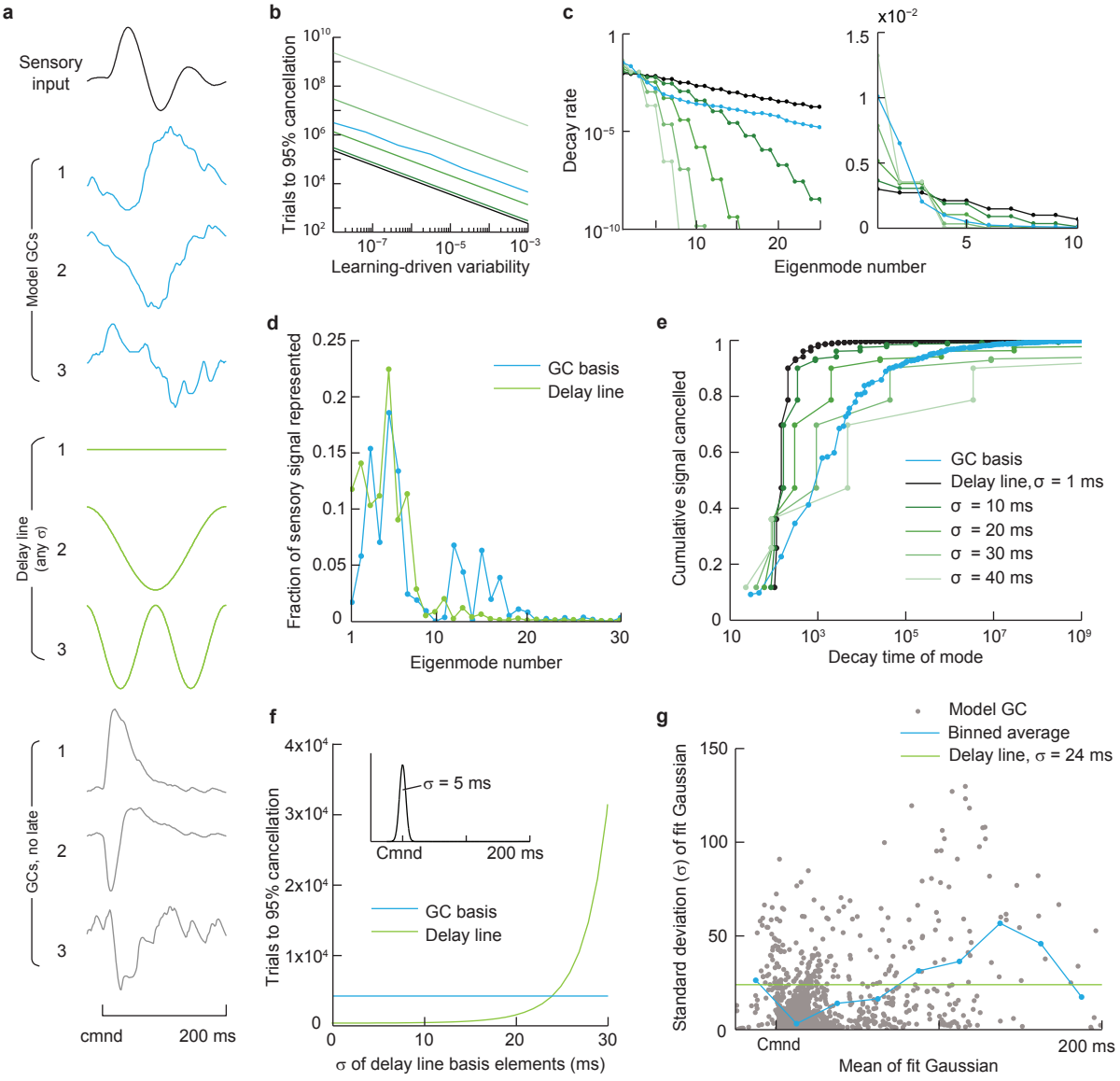
Supplementary figure 7. Learning-evoked changes in EPSP amplitude. Histogram of changes in EPSP amplitude from model granule cells (GCs) to the model medium ganglion cell, following cancellation of self-generated sensory input (initial amplitude 8 mV).



Supplementary Figure 8. Effects of varying the fraction of UBC inputs on negative images and sensory cancellation. UBCs exhibit temporally diverse and delayed responses, termed late and pause, that play an important role in cancelling the later components of the sensory response evoked by the fish's own EOD. We estimated the probability of late and pause inputs from our model fits of granule cell responses, and found the probability of both late-to-granule cell and pause-to-granule cell synapses (p_l and p_p) to be around 5%. To ensure that sensory cancellation and negative image formation did not depend critically on this exact value, we varied this value between 1 and 9% and studied model performance. **a,** The rate at which the granule cell basis was able to cancel self-generated input due to the EOD was comparable for p_l and p_p = 3, 5, 7, and 9%. The rate of cancellation was slightly lower for $p_l = p_p$ = 1%, but still substantially higher than that of a granule cell basis lacking late and pause input entirely, indicating that only a small fraction of granule cells need receive late/pause input to form a sufficient basis. **b,** We also verified that altering the fraction of late/pause inputs to granule cells did not substantially alter the trial-to-trial variability of negative images. Plots show trial-averaged negative images, with shaded blue regions indicating +/- 1 standard deviation.



Supplementary Figure 9. Effects of granule cell response properties on negative image formation. Though the distributions of spikes per EOD command and time of peak membrane potential (Vm) in recorded and model granule cells (GCs) are not statistically different (Suppl. Fig. 6), we wanted to confirm that negative image formation was not sensitive to any differences between real and model cells. First, a small number of model granule cells averaged more than three spikes per command, a firing rate which was not observed in the experimentally recorded granule cells; we removed these more active granule cells from our simulations to ensure that our results still held. Second, only a small fraction (1.18%) of recorded granule cells had peak Vms occurring later than 125 ms relative to command time. We divided the command interval into thirds and measured the fraction of recorded cells peaking in each third, and removed granule cells from our model population to match these proportions. **a,** The rate of signal cancellation of the original granule cell population is not substantially different from that of granule cell populations with highly active (>3 spikes per command) model cells removed. The cancellation rate drops for a more restricted population (cells with <=1 spikes per command), though this may be due to the smaller number of cells remaining. **b**, While the variability of the negative images increased when highly-active cells were removed, the overall shape of the negative image is not affected. **c,** The learning rate of the granule cell population matched to the observed peak Vm distribution was not substantially different from that of the original granule cell population. **d,** The shape and variability of the negative image formed by the modified granule cell population was also unaffected. **e,** Proportion of granule cells with Vm peak in each third of the command interval (20,000 model cells, 170 real cells. Real cell proportions were scaled to sum to 20,000.)



Supplementary Figure 10: Analysis of learning dynamics and comparison to delay line basis. We characterize learning in terms of the eigenmodes and eigenvalues of the learning dynamics, as outlined in the Supplementary Methods. **a,** First three eigenmodes of the three bases considered in this study: the model granule bell (GC) basis (blue), a delay line of uniform Gaussian bumps with with standard deviation σ (green), and the model basis without UBC input (gray). The eigenmodes of a uniform delay line are the Fourier basis, regardless of the shape of individual basis functions. **b,** Decreasing the learning rate λ increases the number of trials needed for sensory input cancellation, but decreases the amplitude of negative image variability arising from learning-induced weight fluctuations (see Supplementary Modeling; we assume η takes the form of Gaussian white noise.) Regardless of λ, the relative rates of cancellation of the different bases remain the same; we therefore set the learning rate of each basis such that the negative image variability was 10^{-5} (noting that other values would yield the same results.) See panel **e** for legend. **c,** Decay rate of the eigenmodes of the granule cell and delay line bases, on log (**left**) and normal (**right**) scales. Increasing σ decreases the number of modes of the delay line basis which can be canceled on ethologically relevant timescales. See panel **e** for legend. **d,** Projection of the sensory input to be cancelled onto the eigenmodes (sorted by rate of decay); in both bases the first twenty (or fewer) modes are sufficient to capture almost all of the sensory input. Note that the shape of the delay line eigenmodes does not depend on σ. **e**, Combining panels **c** and **d**, we plot the decay rate of each eigenmode *n* against the fraction of the sensory signal canceled by modes 1 to *n*, providing an estimate of the learning rate of each basis. **f**, Time to 95% signal cancellation as a function of s. The delay line outperforms the granule cells provided that σ is sufficiently small. Inset shows an example delay line basis function with σ = 5 ms, for reference. **g,** We fit a Gaussian to the trial-averaged response of each model cell, revealing a drop in temporal precision of model cell responses at long delays to the EOD command. The green line shows a standard deviation of 24 ms, which in a delay line learned at a rate comparable to the model cell basis.