Sensory and Motor Systems

The Neural Code for Motor Control in the Cerebellum and Oculomotor Brainstem†‡

Kris S. Chaisanguanthum,[1]* Mati Joshua,[2]* Javier F. Medina,[3] ⬤William Bialek,[4] and ⬤Stephen G. Lisberger[2]

[1]Sloan-Swartz Center for Theoretical Neurobiology and Center for Integrative Neuroscience, Department of Physiology, University of California, San Francisco, San Francisco, California 94143, [2]Howard Hughes Medical Institute and Department of Neurobiology, Duke University School of Medicine, Durham, North Carolina 27710, [3]Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, and [4]Joseph Henry Laboratories of Physics and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544

Abstract

Spike trains are rich in information that can be extracted to guide behaviors at millisecond time resolution or across longer time intervals. In sensory systems, the information usually is defined with respect to the stimulus. Especially in motor systems, however, it is equally critical to understand how spike trains predict behavior. Thus, our goal was to compare systematically spike trains in the oculomotor system with eye movement behavior on single movements. We analyzed the discharge of Purkinje cells in the floccular complex of the cerebellum, floccular target neurons in the brainstem, other vestibular neurons, and abducens neurons. We find that an extra spike in a brief analysis window predicts a substantial fraction of the trial-by-trial variation in the initiation of smooth pursuit eye movements. For Purkinje cells, a single extra spike in a 40 ms analysis window predicts, on average, 0.5 SDs of the variation in behavior. An optimal linear estimator predicts behavioral variation slightly better than do spike counts in brief windows. Simulations reveal that the ability of single spikes to predict a fraction of behavior also emerges from model spike trains that have the same statistics as the real spike trains, as long as they are driven by shared sensory inputs. We think that the shared sensory estimates in their inputs create correlations in neural spiking across time and across each population. As a result, one or a small number of spikes in a brief time interval can predict a substantial fraction of behavioral variation.

Key words: abducen; floccular complex; population coding; rate code; smooth pursuit eye movements; temporal code

---

Significance Statement

Our paper evaluates the nature of the neural code under conditions where we can understand the details of trial-by-trial neural variation and its relationship to behavior. We demonstrate that a single extra spike in the activity of a neuron can predict impending motor behavior. This occurs in a system where motor variation is driven primarily by correlated noise in sensory processing, so that neurons in the motor pathways and the behavior itself share a common noise source. As a result, correlations occur between neurons with similar response properties. Further, in the motor system, the common noise source is much stronger than any local noise and dominates the trial-by-trial fluctuations in the spike trains. As a result, single spikes are meaningful.

Introduction

Neurons code information with action potentials, or "spikes". Certain features of the pattern or number of spikes across a population of neurons determine perceptual or motor behavior. Given that a collection of spikes across neurons and/or time drives behavior, it follows that single spikes in single neurons contribute to behavior. The only question is the size of the contribution of a single extra spike in a single neuron, and whether we can measure the contribution. By analogy to how thinking has evolved about the relationship between single neurons and behavior, it seems plausible that we can.

Long ago, we accepted that a single neuron has an effect on behavior. Because large numbers of neurons drive behavior in the mammalian brain, however, we have assumed that the impact of a single neuron would be too small to measure. Now we know that noise in a sensory-motor system can originate in sensory processing and be distributed in parallel to many neurons in the downstream circuits (Osborne et al., 2005). As a result, spike trains are correlated across neurons within the sensory-motor pathways and fluctuations in the firing of a single neuron can predict fluctuations in other neurons, as well as perceptual or motor behavior (Britten et al, 1996; Shadlen et al., 1996; Medina and Lisberger, 2007; Schoppik et al., 2008; Hohl et al., 2013; Joshua and Lisberger, 2014). It follows that information may be contained in the presence or absence of a single spike at a given time in one neuron (Bialek et al., 1991) or in the fine-grained pattern of spikes and silences across small populations of similar neurons (Osborne et al., 2008).

Several prior examples confirm the possible importance of one or a few spikes in single neurons. In the human somatosensory system, perception correlates with single spikes in sensory afferent neurons (Valbo, 1995). In the visual system, arrival at the retina of a very small number of photons can be perceived (Barlow 1956) and single photons cause measurable responses in retinal ganglion cells (Barlow et al., 1971). We can infer, therefore, that one or a few spikes leaving the retina can correlate with perception. Finally, a series of experiments and analyses on the motion sensitive neurons of the fly visual system attest to the sensory information that is carried by single spikes even when inputs have the complex, dynamic character found in the natural setting (de Ruyter van Steveninck and Bialek, 1988; Bialek and Zee, 1990; Bialek et al., 1991). Thus, it seems worthwhile to investigate whether single spikes can predict something about the impending behavior in a working motor system.

Here, we work with published neural recordings from a small, manageable neural system, namely the smooth pursuit eye movements of monkeys. The anatomy of the essential circuit and the transformations of neural signals at different levels of the circuit are already understood (Lisberger, 2010). The sensory inputs for pursuit encode visual motion and are represented in extrastriate area MT (Newsome et al., 1985). Correlated noise in the population of MT neurons (Huang and Lisberger, 2009) leads to errors in sensory estimates of target speed and direction (Osborne et al., 2005). These errors propagate through the motor system to cause variation in eye speed and direction at the initiation of pursuit. The relevant groups of neurons in the pursuit motor system include Purkinje cells in the cerebellar floccular complex, floccular target neurons (FTNs) in the vestibular nucleus, other neurons in the vestibular nucleus that are not FTNs, and motoneurons in the abducens nucleus. In the present paper, we analyze the spike trains in published recordings from these neurons to ask whether or not a single spike can make meaningful predictions about the impending movement, and why (or why not).

Our main finding is that a single extra spike can be quite informative about the impending behavioral output from the system. We find the same effects in simulated spike trains that have the same probability of spiking as do the real neurons and that mimic the firing statistics for each of the classes of neurons we have studied. The informative nature of a single spike results from the fact that the dominant source of variation in the sensory-motor behavior is the noise in sensory representations of motion. Sensory noise is distributed widely to downstream neurons. This leads to neuron−neuron correlations and temporal correlations in the spiking patterns of a population of neurons, which in turn render a single spike informative. Thus, neurons can transmit information based on the timing of single spikes, and single spikes could be a meaningful part of the neural code.

Materials and Methods

Experimental data

We have reanalyzed data reported in published studies (Medina and Lisberger, 2007, 2008, 2009; Joshua and Lisberger, 2014). The database comprised simultaneous recordings of eye movements and single-unit activity from 51 Purkinje cells in the floccular complex of the cerebellum, 31 FTNs in the brainstem, 88 non-FTN neurons in the region of the vestibular nucleus, and 54 neurons in the abducens nucleus. Methods for recording eye movements and neural responses, and for identifying the neurons, appear in the papers cited above.

We analyzed data obtained during pursuit tracking of step-ramp target motions (Rashbass, 1961) presented in discrete trials. Each trial began with the appearance of a stationary target that the monkey fixated for a random

interval of 500 − 1000 ms. The target then underwent a step in one direction and began ramp motion in the opposite direction. At the end of each trial, the target stopped and the monkey was required to fixate it for an additional 500 − 800 ms. Each recording session contained many repetitions (∼50 − 200) of one or a few target motions. We analyzed the trials that presented target motion at 10, 20, or 30 deg/s in the preferred direction of the neuron under study. The database included different blends of target speeds in different monkeys and structures, but we found no evidence that our results depended on target speed. In instances where we were given data for a single neuron for multiple speeds of target motion, there was a strong correlation in the results across target speeds.

Reducing the dimensionality of behavioral variation
We analyzed the trajectory of eye velocity from 50 to 200 ms after the onset of target motion for each pursuit eye movement. This comprises the interval before visual feedback can affect eye motion, and therefore our analysis considers only the open-loop interval that is driven by sensory estimates of target speed and velocity (Lisberger and Westbrook, 1985). To focus on the relationship between variation in neural spiking and eye speed, we analyzed eye velocity projected onto the direction of target motion. As shown previously (Osborne et al., 2005), the eye velocity during the open-loop interval varies from trial to trial. Our goal was to analyze whether variation in eye velocity is predicted by variation in the spike timing and patterns of cerebellar neurons, brainstem premotor neurons, and motoneurons.

To make the problem tractable, we reduced the 150-dimensional eye speed vector to a small number of principal components and described each eye movement response by the amplitudes of the first two principal components. We denote the eye velocity vector in each trial j by $E_j(t)$ and the mean eye velocity vector over repeated presentations of the same target motion by $\bar{E}(t)$. We then computed the behavioral variation vectors $e_j(t) \equiv E_j(t) - \bar{E}(t)$ and used them to calculate the 150×150 covariance matrix of behavioral variation C_e. Finally, we extracted the eigenvectors $P_k(t)$ of the covariance matrix and used them to represent the trial-by-trial variation in pursuit eye speed:

$$e_j(t) = \sum_{k=1}^{150} A_{jk} P_k(t) \qquad (1)$$

Here, the coefficient A_{jk} is the amplitude of the kth eigenvector for the eye speed in the jth trial. For further analysis, we normalized the coefficients A_{jk} to have unit variance.

For pursuit initiation, the first principal component described >85% of the total behavioral variation in the data (Fig. 2D), depending on the monkey and the data set. The low dimensional nature of pursuit is in good agreement with the >90% value reported by Osborne et al. (2005) for the first three principal components. Note a subtle difference between our analysis and that of Osborne et al.

(2005): we are considering only target motions along a single axis. We calculated the first two principal components of eye speed along the axis of target motion, whereas they performed principal component analysis on the two-dimensional trajectories of eye velocity. By reducing the dimensionality of the eye velocity and characterizing each movement according to the magnitude of the first two principle components, we characterized the overall variation in pursuit strength, but not the millisecond-by-millisecond fluctuations.

Linear estimators of behavioral variation
For each behavioral trial j, we represented the spikes as binary vectors in bins, $\vec{x_j}$. To establish the relationship between behavioral and neural variation for a single neuron, we subtracted the mean spike count across trials in each bin to obtain $\Delta \vec{x_j}$. We then asked how well $\Delta \vec{x_j}$ can predict the amplitudes of the top two principal components of behavioral variation in the same trial. We used linear estimators with weights $\vec{W}$ to estimate A_j as a function of the variation of each spike train, $\Delta \vec{x_j}$:

$$\hat{A_j} = \vec{W} \cdot \Delta \vec{x_j} \qquad (2)$$

We performed the analysis separately for the first and second principal components. We optimized $\vec{W}$ by minimizing the sum over trials j of the squared difference between the actual and predicted values of A_j: $\sum_j (\hat{A_j} - A_j)^2$. We performed the optimization a number of times, determined by the number of trials in each dataset. For each run of the optimization, we left one trial out and assessed the quality of the prediction by applying the best linear estimator to the single trial that was not used for the optimization. Performance was estimated by computing R^2 between the actual A_j and the predicted $\hat{A_j}$ on the omitted trial across all runs of the optimization. This linear reconstruction strategy parallels that originally used by Bialek et al. (1991) for time-dependent sensory signals; here we are trying to estimate just one component of the time-dependent motor output, so the analysis is simpler.

Creation of simulated spike trains with authentic statistics
We created a model of the responses of Purkinje cells and brainstem neurons by calculating simulated spike trains with essentially the same statistics as the real spike trains recorded in the monkeys. The model was a useful way to explore the limitations of our data analysis approach and to test hypotheses about the nature of the neural code for movement.

For each neuron, we fitted the mean firing rate during pursuit of step-ramp target motion with a linear model:

$$F(t) = C_0 + C_1 E(t - \Delta t) + C_2 \dot{E}(t - \Delta t) + C_3 \ddot{E}(t - \Delta t) \qquad (3)$$

where F, t, Δt, E, $\dot{E}$, and $\ddot{E}$ are the firing rate of a single neuron; time; a fixed latency between firing rate and eye

movement; and eye position, velocity, and acceleration along the on-direction of the neuron, respectively. This procedure resulted in a set of coefficients for each neuron we studied: C_0 is baseline firing rate with the eyes stationary at straight-ahead gaze; the values of C_{1-3} denote regression coefficients. As reported in the papers that included the original data, Equation 3 describes ∼90% of the variance in the mean responses of floccular Purkinje cells (Shidara et al., 1993; Medina and Lisberger, 2009) and brainstem neurons (Joshua and Lisberger, 2014) in terms of the parameters of the mean eye movements they drive. We next used the regression equation to predict the probability of spiking at different times during each trial of pursuit eye movements. We took advantage of the fact that the same regression equation used to fit the mean responses can account for a substantial traction of the trial-by-trial variation in neural responses (Medina and Lisberger, 2007; Joshua and Lisberger, 2014). For each behavioral trial, we computed the time-varying probability of spiking by applying the parameters of the fit between the average firing rate and eye kinematics (Eq. 5) to the actual eye kinematics in that trial. This is equivalent to the hypothesis that variability in motor output $E(t)$ is dominated by sources of noise that are "upstream" of the neurons we are studying, and hence these neurons have access to a signal that is almost perfectly correlated with the eye movements that the neurons ultimately drive. We generated spikes for each model behavioral trial using a procedure that mimicked the (sub-Poisson) coefficient of variation (CV) for each model neuron. We set the times of simulated spikes for each behavioral trial in a model neuron for each actual neuron. We started by placing a spike at a random time shortly after the end of the trial. Then we placed spikes iteratively one at a time at the next earlier time corresponding to the inverse of firing rate. We chose to start at the end of the trial so that spikes would tend to be placed before, rather than after, the behavior they drive, and to maximally desynchronize the spike trains by the time we had reached the interval used for most of our analyses, namely the initiation of pursuit. After placing each simulated spike, we randomly jittered its time by multiplying the interspike interval by a random sample from a gamma distribution with a standard deviation equal to the CV of the neuron being simulated. We created a refractory period by requiring each interspike interval to endure at least 2.5 ms. We verified that we obtained the same results with the variation in interspike intervals simulated by either a normal distribution or a nonhomogeneous gamma process. Finally, we subjected the simulated spike trains to the same analyses outlined above for the real spike trains. The use of the regression model and actual CV for each real neuron guaranteed that each model neuron would mimic the trial-by-trial neuron−behavior correlations (Medina and Lisberger, 2007; Joshua and Lisberger, 2014) in the real neuron whose parameters were used to create the simulated spike trains.

Results

The meaning of a single extra spike in cerebellar Purkinje cells

We analyzed recordings made from floccular Purkinje cells (and brainstem neurons) while monkeys tracked many repetitions of the same "step-ramp" target motion (Rashbass, 1961). The combination of a step and ramp of target position (Fig. 1A, downward arrow) allows the monkey to initiate smooth pursuit eye movement without an early saccadic eye movement and creates data where the responses to pursuit can be studied in isolation. Eye velocity records (Fig. 1D) show that the monkey initiated a brisk increase in smooth eye velocity ∼100 ms after the onset of target motion, and tracked accurately for ∼500 ms. Towards the end of the records, smooth eye velocity showed a gradual reduction as a prescient prediction that the target would stop. At the end of the target motion, the monkey fixated the stationary target accurately so that he could receive a reward for successful completion of the trial. We analyzed the eye movements and spikes in the interval from 50 to 200 ms after the onset of target motion.

Both the eye velocity of pursuit (Fig. 1B) and the spike trains of neurons in the cerebellum (Fig. 1C) and brainstem varied considerably from trial-to-trial, even though the target motion always was the same (Osborne et al., 2005; Medina and Lisberger, 2007; Joshua and Lisberger, 2014). Our premise is that understanding the neural basis for the variation might be broadly informative about the relationship between spike trains and behavior, and might reveal the nature of the neural code in these premotor neurons.

We started by using principal component analysis to reduce the dimension of the eye speed variation in the initiation of smooth pursuit (Fig. 1B). The first and second principal components of eye speed (Fig. 1G) accounted for more than 85% and 10% of the variation in the speed of pursuit initiation, respectively (Fig. 1H). Thus, the coefficients of the first and second principal components provided a nearly full account of the variation in the behavioral responses, in agreement with Osborne et al. (2005). We quantified the spike train $\vec{x_j}$ with much greater temporal resolution. We represented $\vec{x_j}$ according to the presence or absence of spikes in bins 5 ms in width. Thus, we started by evaluating the relationship between a high-dimensional representation of the neural responses and a one- or two-dimensional representation of the behavior.

We related the variation in each neuron's spiking activity (Fig. 1C) to the variation in the animal's behavior (Fig. 1B) in two steps. First, we computed the best linear estimator (Eq. 2) of the magnitude of the first or second principal component. Then, we correlated the amplitudes of the principal components predicted by the best linear estimator with the amplitudes of the principal components measured from the data. For floccular Purkinje cells studied during pursuit of target motion at 10, 20, or 30 deg/s, the distributions of correlations were strongly positive for both the first and second principal components (Fig. 1E,F). The results were similar for all target speeds. Further, if a Purkinje cell's spike train predicted behavioral
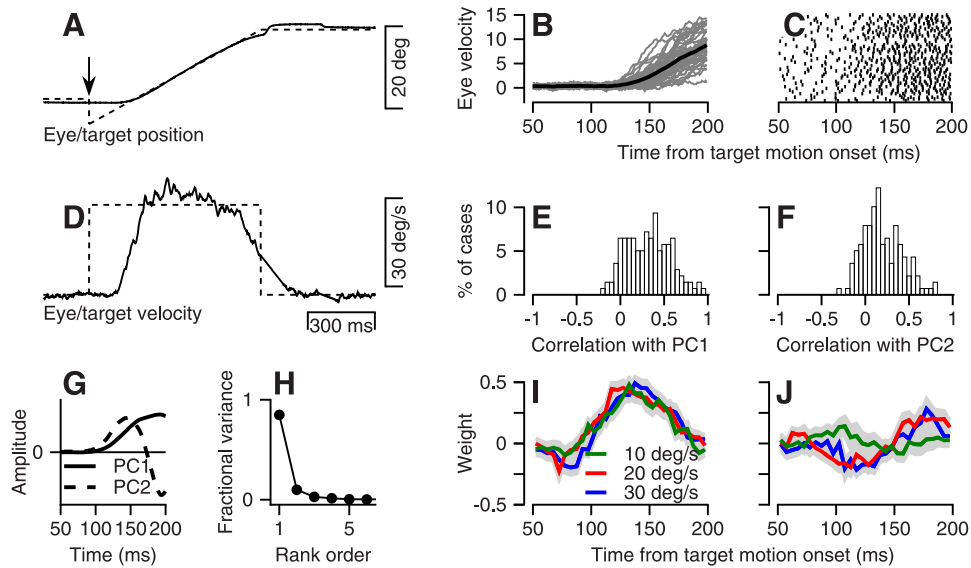
Figure 1 Example of an experimental trial and of neural and behavioral variation. ***A***, Superimposed eye and target position for a step-ramp target motion. Dashed and continuous traces show target and eye motion. ***B***, Eye velocity responses during the initiation of pursuit for recordings from a Purkinje cell. Gray and black traces show individual trial responses and the mean response. ***C***, Raster from a typical Purkinje cell during pursuit initiation. Each line shows the response in one behavioral trial, and each tick shows the time of one simple-spike. ***D***, Superimposed eye and target velocity for a step-ramp target motion. Dashed and continuous traces show eye and target motion. ***E***, ***F***, Distributions across the full sample of Purkinje cells of the correlation between actual and predicted magnitudes of the first (***E***) or second (***F***) principal component (PC1 and PC2, respectively) of the behavior. ***G***, First two principal components resulting from dimension reduction of the variation in eye velocity at the initiation of pursuit. ***H***, The amount of total behavioral variance explained by leading principal components for the variation in eye speed at the initiation of pursuit. ***I***, ***J***, Each trace shows the weights for an optimized linear estimator ($\vec{W}(t)$ in Eq. 2) relating the presence of an extra spike at time t on the horizontal axis to the amplitude of the first (***I***) or second (***J***) principal component of eye movement variation. Different colors show results for a single Purkinje cell during pursuit at three different target speeds. All times are relative to the onset of target motion.

variation well in movements at one target speed, then its spike trains also predicted behavioral variation in movements at other target speeds (data not shown).

The weights that worked best for the linear estimator ($\vec{W}$ in Eq. 2) comprise a time-varying vector that indicates
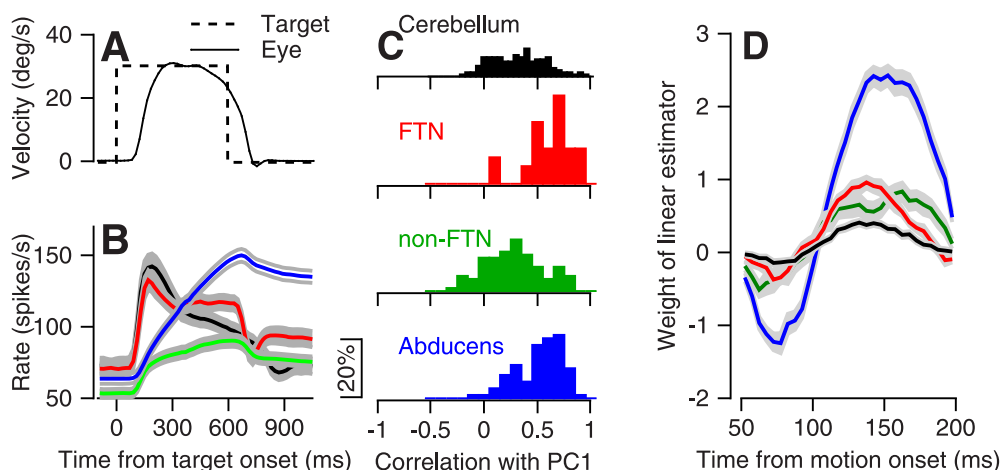
Figure 2 Prediction of eye speed by a single extra spike in four groups of cerebellar and brainstem neurons. ***A***, ***B***, Superimposed means of eye and target velocity for a typical recording session (***A***) and average firing rates (***B***) for the different brainstem neurons. Different colors show the mean firing rate across the different populations, and the gray ribbons show the SEM. ***C***, Distributions across the full samples of different neuron classes of the correlation between actual and predicted magnitudes of the first principal component (PC1) of the pursuit behavior. ***D***, Weight of spikes as a function of time for the best linear estimator. In ***B***, ***C***, and ***D***, black, red, green, and blue show data for floccular Purkinje cells, FTNs, non-FTN vestibular neurons, and abducens neurons. The traces in B were taken from a figure in an earlier paper (Joshua et al., 2013).

the relative contribution of spikes at different times to our best estimate of the value of one of the principal components of the behavior. The shape of the weight functions was similar across target speeds for Purkinje cells (different colored traces in Fig. 1I,J), but different between the two top principal components (Fig. 1, I vs J). The weight functions did not follow the shapes of the principal components (or derivatives thereof). In general, spikes that occurred $120 - 160$ ms after the onset of target motion were weighted most strongly by the linear estimator.

The weights in the optimal linear estimator have a scale, as in Figure 1I, peaking at $\sim$0.5. Referring to Equation 2, and keeping in mind that the amplitudes of principal components have been normalized, this scale means that one extra spike predicts a fluctuation in the component of the behavior up to one half of the standard deviation in the behavior. We found a few examples of floccular Purkinje cells where a single extra spike predicted a full standard deviation of the behavioral variation. Thus, the presence or absence of a few spikes in a 100 ms interval in floccular Purkinje cells is sufficient to predict nearly the full range of the behavioral variation.

The meaning of a single extra spike in brainstem premotor neurons and motoneurons

Some neurons in the brainstem showed quite large correlations between the actual magnitude of the first principal component of the behavior and the magnitude predicted by the linear estimator (Eq. 2). Among brainstem neurons, floccular target neurons receive monosynaptic inhibition from the floccular complex of the cerebellum and were identified by single-shock stimulation in the floccular complex (Lisberger et al., 1994). The sample of abducens neurons we analyzed probably included mainly motoneurons that innervate the lateral rectus muscle. The distributions of correlations (Fig. 2C) were shifted toward higher values in FTNs and abducens neurons compared to Purkinje cells or non-FTN vestibular neurons. Average correlations for Purkinje cells, FTNs, non-FTN vestibular neurons, and abducens neurons were 0.32, 0.61, 0.27, and 0.51, respectively.

In the rest of our paper, we describe spikes as having predictive value when the activity of a neuron allows us, as observers, to predict a quantifiable fraction of motor variation. We regard the degree of prediction as a continuum, and in no sense are we implying that any single neuron or spike causes behavior by itself, in isolation from other, correlated neurons and spikes. The difference in predictive value of a single spike across neuron types was not tightly linked to the time-varying profile of firing rates (Fig. 2B). Purkinje cells and FTNs both had early transients and steady sustained firing rate, but differed in the predictive values of their spikes. Non-FTN vestibular neurons and abducens neurons both showed ramp increases in firing rate throughout the duration of pursuit, but differed in the predictive values of their spikes. Single spikes had the highest predictive values in FTNs and abducens neurons, but these neurons have quite different time-varying profiles of firing rate. Finally, the shapes of the average linear weighting functions were similar across the neuron

populations (Fig. 2D), even though the functions differed in amplitude. The time of the peak weight shifted slightly later as signals moved downstream from floccular Purkinje cells to FTNs to abducens neurons.

Similar predictive value of single spikes in real and simulated spike trains

Why does a single extra spike convey information about the impending movement? Is it a natural consequence of the sensory origin of motor noise and the nature of the code for movement in premotor pathways and motoneurons? What properties of neuronal spiking or responses create a situation where a single extra spike is informative?

The current understanding of the pursuit system is that vision estimates the relevant parameters of target motion (e.g., speed, direction, and time of motion onset) and passes those estimates to the neurons we are studying here (Osborne et al., 2005). Estimation errors are passed along as well, and will generate trial-by-trial variations in neural firing (and behavior), even if the neurons are completely deterministic and follow their sensory inputs perfectly. The presence of trial-by-trial correlations between neural responses and eye-movement behavior in the pursuit system (Medina and Lisberger, 2007; Schoppik et al., 2008; Hohl et al., 2013; Joshua and Lisberger, 2014) implies that variation in the sensory estimates is distributed fairly uniformly to all downstream neurons, so that the responses of neurons within a particular population are correlated with each other (Medina and Lisberger, 2007; Schoppik et al., 2008; Huang and Lisberger, 2009).

The trial-by-trial correlations with motor behavior occur because the neurons in question are driving smooth eye movement, rather than vice versa. But the quantitative relationship to the eye movements provides a convenient tool for understanding the firing rate of the neurons. In addition, each neuron's spike trains fluctuate due to local noise that is independent in each neuron. The local, independent noise reduces the size of neuron−behavior correlations, but can be eliminated by averaging across a population and therefore does not contribute materially to variation in pursuit. Thus, the situation for premotor neurons in the oculomotor brainstem and cerebellum is homologous to that for MT neurons: their spike trains reflect a combination of systematic trial-by-trial top-down fluctuations that may be applied to all neurons and local noise (Goris et al., 2014).

We next use simulated spike trains to show that the sensory-induced variation in neural responses supports the informative nature of single spikes. Local noise in the timing of individual spikes, on the other hand, tends to defeat the informative nature of single spikes. We find that simulated spike trains (see **Materials and Methods, above**) have as much predictive value for the behavioral variation as did the real spike trains. For example, the example Purkinje cell and FTN in Figure 3, A and D, showed correlations of 0.74 and 0.88 between the actual magnitude of the first principal component of eye speed and the prediction of the optimal linear estimator (i.e., Eq. 2). Identical analysis of the simulated spike trains for the
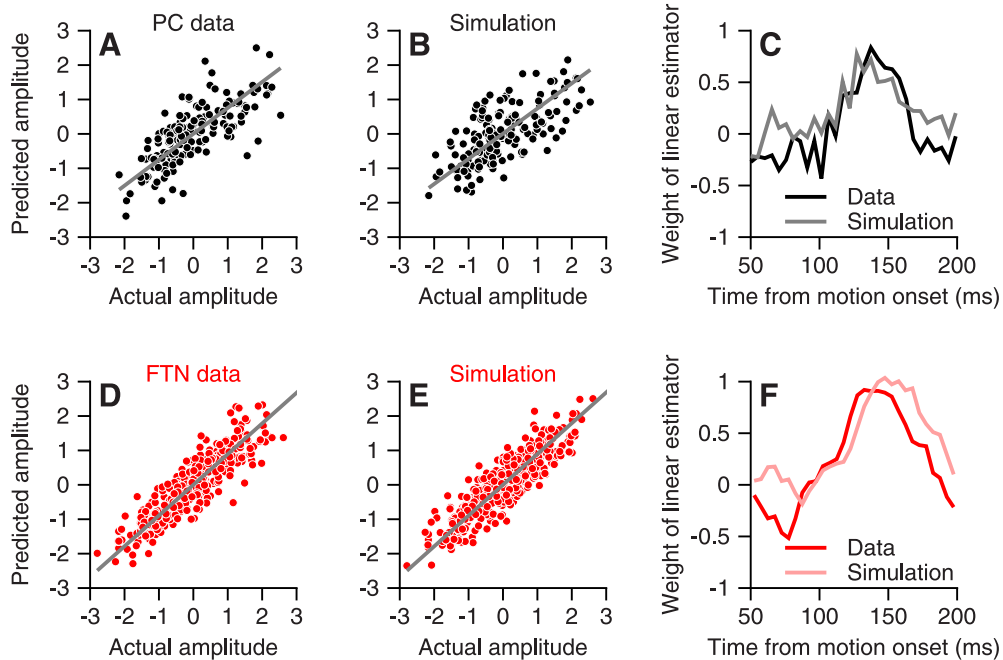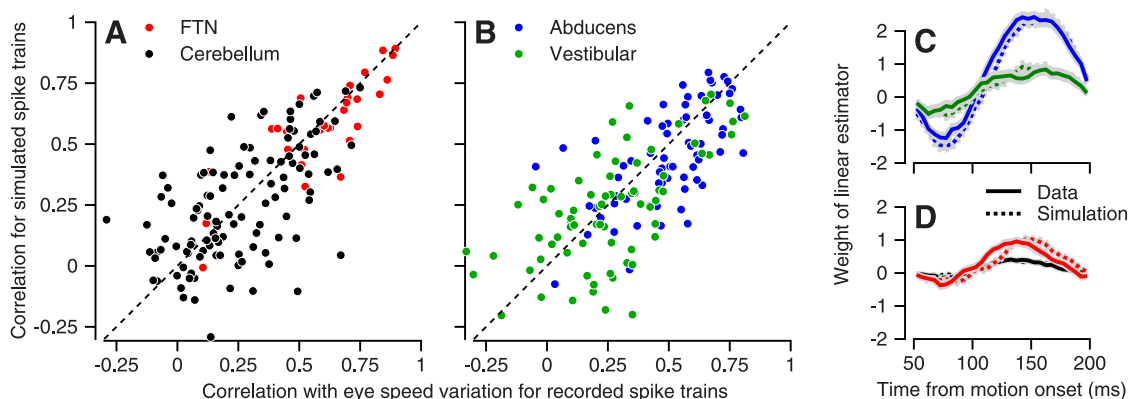
Figure 3 Comparison of the predictive value of a single extra spike in real versus simulated spike trains for a floccular Purkinje cell and an FTN. Results for Purkinje cells are in black and FTNs are in red. **A, D**, Predicted versus actual amplitude of the first principal component (PC) of pursuit behavior based on the optimal linear estimator from the real spike trains. **B, E**, Predicted versus actual amplitude of the first principal component of pursuit behavior based on the optimal linear estimator from the simulated spike trains. In all four graphs, each symbol shows the results from a single behavioral trial. **C, F**, Each trace shows the weights for an optimized linear decoder ($\vec{W}(t)$ in Eq. 2) relating the presence of an extra cell spike at time *t* on the horizontal axis to the amplitude of the first principal component of eye movement variation. Dark and light curves show results for real data versus simulated spike trains.

same two neurons revealed correlations of 0.67 and 0.88 (Fig. 3B,E). Further, the weights in the linear estimators were very similar for the real and simulated spike trains of both neurons (Fig. 3C,F).

Across the full samples of floccular Purkinje cells, FTNs, non-FTN vestibular neurons, and abducens neurons, there was a close relationship between the correlations

for real versus simulated single spikes. The scatter plots in Figure 4, A and B, show that neurons with strongly versus weakly predictive single spikes in the real spike trains had similarly predictive single spikes in the simulated spike trains. In addition, for all four groups of neurons, the average linear estimators from the simulated data followed the same time course as the average linear estima-

Figure 4 Similarity of the predictive value of a single extra spike for real and simulated spike trains across all neurons in our sample. **A, B**, Each symbol summarizes the results for a single neuron and plots the correlation between the optimal predictor based on simulated spike trains against the correlation based on real spike trains. **C, D**, Average weights in the optimal linear estimators for the four groups of neurons, plotted as a function of the time of a spike. Continuous and dashed curves show weights based on the real and simulated spike trains, respectively. Throughout the figure, black, red, green, and blue indicate analyses for floccular Purkinje cells, FTNs, non-FTN vestibular neurons, and abducens neurons, respectively.

tors from the real data (Fig. 4C,D). There is one subtle difference between Figures 3 and 4 in the linear estimators for the simulated data. In Figure 3, C and F, the linear estimators for the simulated data do not have the early negative component found in the estimators for the real data, in the interval from 50 to 100 ms after target motion onset. In Figure 4, C and D, we have modified the procedure for creating the simulated spike trains slightly so that the linear estimators now include the early negative lobe. We did so by measuring the magnitude of slow fluctuations in spike count across trials during fixation in the real neurons, and including those fluctuations in the calculations for the simulated spike trains.

The analysis of simulated spike trains emphasizes that the single spikes are informative about the impending pursuit eye speed because of the organization of the sensory-motor system. Available data imply that sensory estimation errors are distributed fairly uniformly to all floccular Purkinje cells. The resulting shared trial-by-trial variation among Purkinje cells leads to sensory-driven variation in pursuit initiation, and a strong trial-by-trial correlation between Purkinje cell firing rate and eye movement (Medina and Lisberger, 2007). We take advantage of that trial-by-trial correlation to simulate the spike trains of Purkinje cells on the basis of the eye movement in each behavioral trial. Eye movement does not cause Purkinje cell firing in real life, but this procedure yields spike trains that are equivalent to those produced by the shared sensory drive. The match between the statistics of the real and simulated spike trains in relation to motor variation implies that the smallness of local noise and the wide distribution of trial-by-trial sensory estimation errors to the Purkinje cell population are responsible for the predictive power of a single spike. The same logic explains the agreement of the analyses of real and simulated spike trains for neurons that lie between the cerebellum and the eye muscles.

Conditions that allow single spikes to be informative
The observation that we can emulate the predictive value of a single spike with simulated spike trains opens up an approach to understand the properties of a neuron's spiking that render a single extra spike informative. For a representative FTN and a representative abducens neuron, we used simulated spike trains based on different parameters to assess how the predictive value of single spikes varied as a function of signal and noise, defined as the magnitude of the neural response and the value of the CV of the simulated spike trains. In the data, the values of CV during the period of fixation before target motion varied systematically across neuron populations. It averaged 0.35, 0.20, 0.14, and 0.05 for floccular Purkinje cells, FTNs, non-FTN vestibular neurons, and abducens neurons, respectively.

For each model neuron, we simulated the time-varying probability of spikes using the regression coefficients from the best-fitting kinematic model (Eq. 3), defining that as a gain of one. We then systematically varied both the CV and a gain factor that multiplied the probability of spiking for a given trial. We than generated 10 sets of

simulated spikes for each gain and CV, where a set of simulated spikes included all the trials in the dataset. For each set of simulated spikes, we computed the best linear estimator (Eq. 2) of the first principal component of the eye velocity behavior. We averaged across the 10 sets to estimate the mean and standard deviation of the correlation between the actual and predicted magnitude of the first principle component. Finally, we plotted the results as a function of gain and CV (Fig. 5A,B).

The analysis of both the simulations and the data from neurons implies that a neuron's signal-to-noise ratio is the major determinant of how informative its single spikes are about trial-by-trial variations in behavior. For models based on both representative neurons in Figure 5, increasing the value of CV caused large decreases in the correlation between actual and predicted magnitude of the first principal component (Fig. 5A). Conversely, decreasing CV to zero caused the correlation to increase, reaching a value close to one for the FTN and of 0.75 for the abducens neuron. For reference, the actual CV of these two neurons was 0.158 for the FTN and 0.058 for the abducens neuron and the values of the correlation for the data were 0.88 and 0.65. Scaling the gain of the spiking probability for every trial at the CV from the data had the opposite effect. Higher gains led to higher values of correlation coefficient (Fig. 5B), and both neurons achieved correlations of 1.0 when the gain was high enough. We conclude that single spikes are likely to be most informative in neurons that respond strongly to a stimulus or movement, or that have very regular firing.

It is noteworthy that the abducens neuron achieved a correlation coefficient of one for high gain, but not for zero CV. The reason is that some noise is added through the aliasing caused by binning the spikes, and the binning noise is a much larger fraction of the total variance of spiking for the abducens neuron versus the FTN. We were able to demonstrate the larger noise due to binning in model spike trains with CV of zero and no trial-by-trial variation in behavioral output (results not shown).

Analysis of the real data from cerebellar and brainstem neurons confirms and extends the predictions based on simulated spike trains with different gains and coefficients of variation. The CV of each individual neuron during fixation did not predict the correlation between the real and predicted magnitudes of the first principal component of the behavioral variation, either within or between populations of neurons (Fig. 5D). The amplitude of the neural signals, defined as the peak of the average firing rate during the initiation of pursuit minus the baseline firing rate during fixation, predicted the correlation somewhat within some of the neural populations, but did not predict the differences between populations (Fig. 5E). In contrast, the signal-to-noise ratio, defined as the amplitude of the neural signal divided by the CV, was quite predictive. Data from all four neuron populations followed approximately a single relationship, where higher signal-to-noise ratios were accompanied by a larger correlations between the real and predicted magnitudes of the first principal component of the behavioral variation (Fig. 5C,F). Thus,
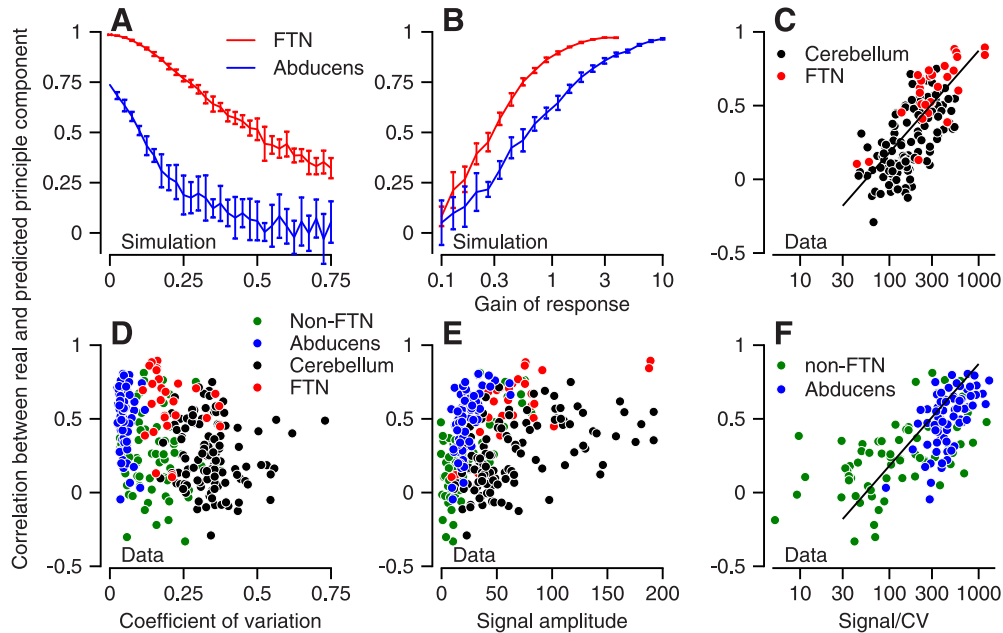
Figure 5 Effect of discharge variability and overall signal magnitude on the predictive value of a spike train. *A, B*, Each graph plots the mean and standard deviation of the correlation between the real and predicted magnitude of the first principal component of eye speed based on simulated spike trains. Data are plotted as a function of the coefficient of variation (*A*) or the gain (*B*) of the artificial spike train. Red and blue symbols indicate analysis of simulations based on a representative FTN or abducens neuron. *C–F*, Scatter plots, where each symbol shows the results of analyzing the data of a single neuron, and the different plots show the correlation between the real and predicted magnitude of the first principal component as a function of the coefficient of variation of the spike trains (*D*), the amplitude of the neural signal (*E*), and the ratio of the signal amplitude divided by the coefficient of variation (*C, F*). Different colors show results for different classes of neurons.

the signal-to-noise ratio of a neuron's spiking during the initiation of pursuit defines the degree to which a single spike can be informative about variations in the impending behavior (see also Schoppik et al., 2008).

Finite datasets underestimate the predictive value of a spike train

The average weight functions across neurons within a population are quite smooth (Fig. 2D), but the functions for individual neurons were quite noisy. The noisiness gave us some concern about whether our fitting procedure was limited by the small size of the datasets. For some neurons, the datasets provided only ∼50 trials of data and the weighting functions had 30 parameters. Our concern was that our "leave one out" strategy might be leading to underestimates of how well a linear estimator could predict the variation in pursuit eye speed. Our logic was that the optimization procedure might be degrading the weighting functions by trying to fit noise in the timing of spikes. The resulting weighting functions would be suboptimal for the test trial that was left out of the optimization.

We tested the effects of finite datasets by simulating the spike trains (see **Materials and Methods, above**) for larger numbers of trials. We reused the same eye velocity traces multiple times but added noise to calculate the spike trains independently each time. For all groups of neurons, both the mean weighting functions and the individual traces were much less noisy when we used more trials (compare second and third column of Fig. 6). The

correlations between real and predicted amplitudes of the first principle component were higher for almost all neurons when the analysis used 1000 trials versus for the real data (first column of Fig. 6), and were still higher when the analysis used 10,000 trials (not shown). We showed already (Fig. 4A,B) that we obtained the same correlation from simulated trials as we did with the real trials, as long as trial numbers were equated. Thus, the correlations between the real and predicted amplitude of the first principle component in the analysis of the actual data provide lower bounds on the actual correlations that would have been obtained from the linear weighting function if we had acquired much larger datasets.

**Linear weighting functions versus spike counts**

The shapes of the linear estimators for the first principal component of eye speed (Figs. 1I, 2D, 3C,F, 4C,D) predict that spikes are more informative about the first principal component of eye movement variations if they occur near the center of the analysis window versus at the start or the end. Thus, the analysis with the linear estimator predicts that the time of a neuron's spikes can be important; simply counting spikes might be less informative than a linear weighting. However, the weighting function has a smooth envelope, so that small variations in spike timing may not have much effect. To resolve this, we next used a rectangular weighting function to test directly, by counting spikes in different time windows, how much the location or duration of the analysis window matters.
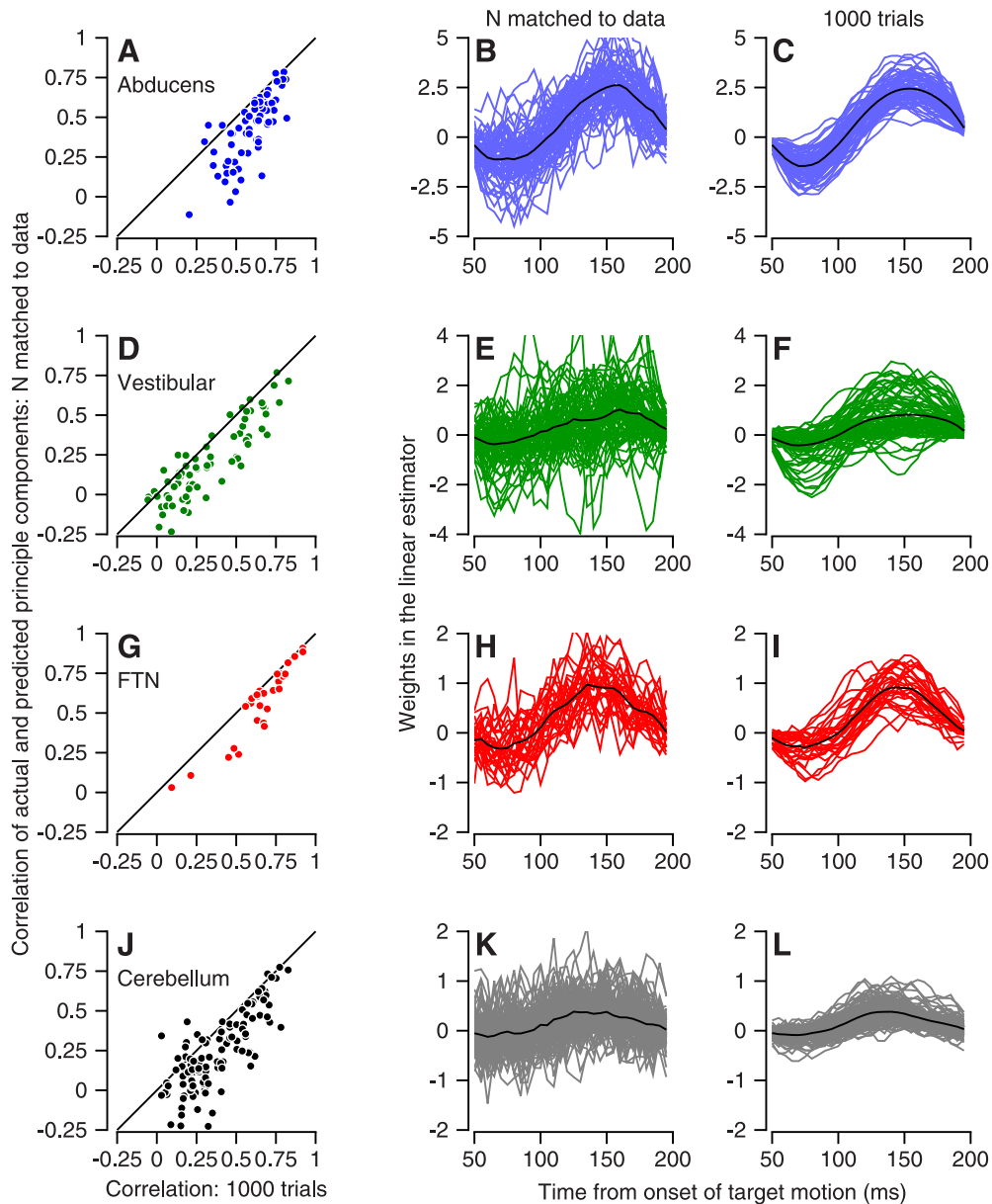
Figure 6 Demonstration that computing the best linear estimator with finite datasets causes over-fitting and underestimates the predictive value of a spike train. Each row shows data for a different set of neurons. Left, Each graph contains one symbol for each neuron and shows scatter plots of the correlation of actual and predicted magnitude of the first principal component of the data. Each point plots the correlation obtained using the same number of trials available in the real data on the *y*-axis and the correlation obtained using 1000 trials of simulated data on the *x*-axis. Middle and right, Each graph plots the time course of the optimal linear estimator when we used the actual data (middle) or 1000 simulated trials (right). The black curve in each graph shows the average across all neurons and the gray and colored traces show the individual neurons.

Spike counts in brief windows centered 125 ms after the onset of target motion predicted the magnitude of the first principal component almost as well as did the linear weighting function. For each neuron type, the correlation between the predicted and actual magnitude of the first principal component increased as a function of the duration of the counting interval, and reached an asymptote for intervals of duration 40 to 100 ms (Fig. 7). For Purkinje cells, windows as short as 40 ms provided nearly asymptotic correlations, mean-

ing that a good account of the variation in motor behavior can emerge from just a few spikes (0.04 s × 100 spikes/s = 4 spikes). The asymptotic correlation in Purkinje cells matched the correlation provided by the linear estimator for the actual data, while for the other neurons the asymptotic correlation was somewhat smaller than that provided by the linear estimator. Interestingly, counting spikes is the least informative for abducens neurons, probably because of edge effects created by their very regular spiking pattern.
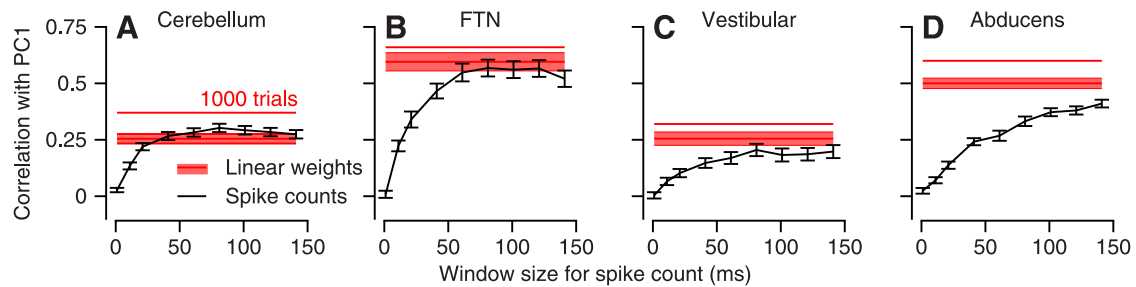
Figure 7 Prediction of eye speed by spike counts in windows of different duration. Each graph plots the correlation between the output of different estimators and the amplitude of the first principal component of eye speed. Pink swatches are the lower bounds for the best linear estimator (Eq. 2). Black lines are from the spike count in windows of different duration centered 125 ms after the onset of target motion. The red line labeled "1000 trials" is the estimate of the upper bound of the results from the best linear estimator. Different graphs show data for different neuron types.

Recall that the performance of the optimal linear estimator might be worse than that of the spike counting model simply because these models differ in their complexity. Taking account of the problems of over-fitting (Fig. 6) improves performance relative to what we can infer from the limited number of trials in our experiments. With a dataset of 1000 trials, the linear estimator predicts performance that is slightly better than the asymptotic performance of spike counting in windows up to 100 ms in duration (horizontal red lines in Fig. 7). Thus, counting spikes in an appropriate small time window predicts the first component of behavioral variation only slightly less well than does an optimal linear estimator.

Estimation based on Purkinje cells' spike counts in sliding windows of different durations confirmed that counts were most informative if they occurred in the middle of the initiation of pursuit, ~140 ms after the onset of target motion (Fig. 8A). The peak correlation between the actual and predicted magnitude of the first principal component of eye speed increased as a function of the duration of the analysis interval. For an analysis interval of 41 ms, the correlation from counting spikes matched that from the linear estimator, but remained less than we expected from the linear estimator with larger dataset (red line labeled "1000 trials").

The number of spikes in a 41 ms analysis window has a clear relationship to the eye velocity at the initiation of pursuit. In Figure 8D, the average trajectory of eye velocity as a function of time shows a more rapid rise as the number of spikes in the analysis window increased from three to six. The correlation between spike count and the magnitude of the first principle component of the variation in eye speed is clearly visible in a scatter plot of all trials from a single Purkinje cell (Fig. 8E). By normalizing the y-axis of graphs like Figure 8E for the standard deviation of the magnitude of the first principle component, we can use the slope of the regression line to estimate the impact of a single spike on the variation in eye speed, in units of standard deviations per spike. The distributions of sensitivity for the populations deviated significantly from zero for Purkinje cells, FTNs, and abducens neurons ($p < 10^{-5}$ for all populations, signed rank test), with average values of 0.31, 0.70, and 0.66, respectively (Fig. 8F).

Finally, the data from the 5 ms analysis window are particularly intriguing because the firing rate of 125 spikes/s at the peak of the correlation (gray shading in Fig. 8B) corresponds to either one or zero spikes in the analysis window on most trials (Fig. 8C). Two spikes occurred very rarely and three spikes almost never occurred. Thus, the peak correlation of almost 0.1 shows that one spike in one Purkinje cell at the right time is informative about 1% of the variance in the eye speed. The predictive power of spikes seems to accumulate across time, however. The presence or absence of a single spike is substantially less informative than is the number of spikes in a larger window, or the spike train convolved with an optimal linear estimator.

Counting spikes across time versus counting spikes across neurons

The similarity of the statistics of the real and simulated spike trains opens up the possibility of using the simulated spike trains to ask questions about the neural code that are not accessible given the current size of datasets. Because the responses of different neurons are linked through their trial-by-trial correlation with behavior, we can generate simulated population responses for many variants of the behavior with confidence that the simulated populations will accurately represent the statistics of real populations, if we could record them.

We compared the predictions from counting spikes across a population of neurons versus counting spikes across time in a single neuron. In each case, we asked how well the spike count predicted the magnitude of the first principal component of eye speed. As before, we created a population of *N* neurons with *M* trials of target motion. For the *M* trials, we used the different trajectories of the initiation of pursuit taken from the data and the approach outlined in **Materials and Methods**, above, to generate $N \times M$ spike trains. An example of the results appears in Figure 9, which plots the magnitude of the correlation as a function of the time of the center of the counting interval. We obtained better correlations between the first principal component of eye speed and spikes counts in a 40 ms window in individual neurons (Fig. 9, blue traces), versus counts from 40 different neurons in a 1 ms window (Fig. 9, green traces). The difference was present in all four groups of neurons we studied.
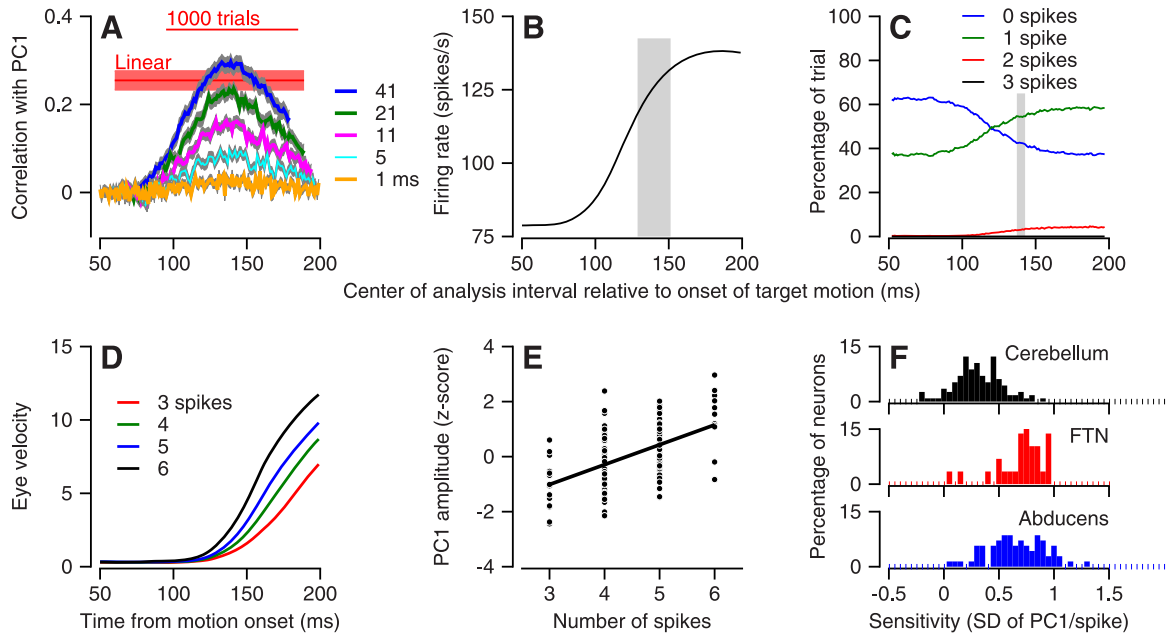
Figure 8 Prediction of eye speed by spike counts in floccular Purkinje cells. **A**, Correlation between actual and predicted amplitude of the first principal component of eye speed as a function of the duration and time of the analysis interval. Different colored traces show results for analysis intervals of different durations. Pink band labeled "Linear" shows the correlation obtained with the best linear estimator. Red line labeled "1000 trials" shows the expected correlation for the best linear estimator if we had acquired 1000 repetitions of the target motion. **B**, Average firing rate of the population of Purkinje cells as a function of time. Gray bar shows a 20 ms analysis window centered at 140 ms. The gray bar indicates the time of the peak correlation with the first principal component (PC1) in **A**. **C**, Number of spike counts in a 5 ms analysis interval as a function of the center of the interval. Different color traces show numbers for different spike counts. **D**, Averages of eye velocity as a function of the time from onset of target motion for trials in one example Purkinje cell. Different colors show eye velocities for trials with three, four, five, or six spikes in the 41 ms interval centered 140 ms after the onset of target motion. **E**, Scatter plot showing the z-scored amplitude of the first principle component of eye speed as a function of the number of spikes in the same interval as **D**. Each symbol shows data from one trial. The regression line defines the sensitivity in units of standard deviations per spike. $R = 0.53$ for the symbols and the regression equation was $y = 0.72x - 3.18$. **F**, Distributions of sensitivity for populations of floccular Purkinje cells, FTN, and abducens neurons.

To be more systematic, we next evaluated the correlations between the actual and predicted magnitudes of the first component of eye speed for spike counts in bins of width ranging from 1 to 96 ms and for populations of 96 to 1 model neurons. The colored images in Figure 10 show the correlations between actual and predicted magnitude of the first principal component of eye speed for each combination. The graphs include superimposed iso-correlation contours to show the tradeoff between population size and duration of the analysis window. The contours have slopes that range from −2 to −4 in their steep negative regions, indicating that decreases in the width of the analysis window had to be compensated by at least a doubling of the population size to maintain the predictive value of spike counts.

The results in Figures 9 and 10 depend on the degree of variation in the interspike intervals. For Poisson spike trains, it did not matter whether we counted spikes in a N millisecond window for one neuron versus a 1 ms window for N neurons (data not shown). However, for the sub-Poisson spike trains of neurons in the oculomotor cerebellum and brainstem, it was better to count across a larger window in a smaller number of neurons. Thus, under the circumstances that exist in the oculomotor cerebellum and brainstem, the spike train in a brief window for a single neuron predicts behavior better than does a population average with high-temporal resolution.

Discussion

Our goal was to use detailed analyses of spike trains on a fine temporal scale during individual movements to understand the neural code for motor control. We studied neurons whose firing rate, defined as the inverse of the interspike interval, encodes trial-by-trial variation in the impending motor behavior (Medina and Lisberger, 2007; Joshua and Lisberger, 2014). Our analysis reveals that the presence and timing of a single extra spike in a single behavioral trial is informative about the impending eye movement in many classes of neurons in the oculomotor system. They include Purkinje cells in the cerebellar floccular complex, FTNs in the brainstem that are last-order interneurons, other premotor neurons in the region of the vestibular nucleus, and neurons in the abducens nucleus.

Possible neural codes in the motor system

Our analysis reveals the features of spike trains that could constitute the neural code in the final motor pathways for pursuit eye movements. Variation in a spike train of the order of a single extra spike is informative about the variation in behavior. The time of the extra spike matters. Spikes are
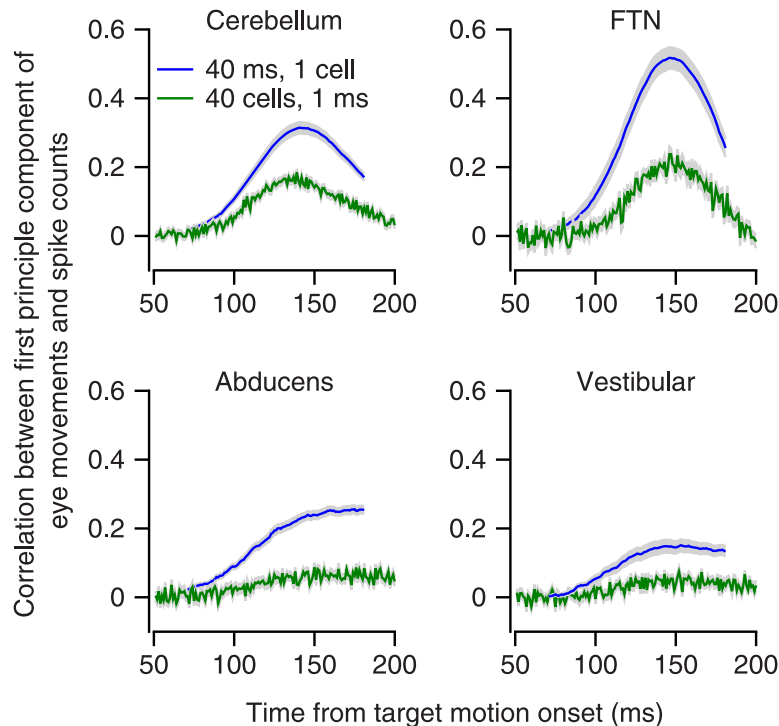
Figure 9 Comparison of predictions made by spike counts over 40 ms in time versus space. Each panel shows data for a different group of neurons. Within each panel, the traces show the correlation between the actual and predicted amplitude of the first principal component of eye speed as a function of the time at the center of the analysis window. Blue traces show the results for counting spikes in a 40 ms window; the traces end at 180 ms because of the width of the analysis window. Green traces show the result for counting spikes in a 1 ms window across the simulated spike trains of 40 different neurons.

more informative if they occur around the center of the 100 ms time interval during which pursuit initiation occurs, versus at the start or the end of the interval. Simply counting spikes is almost as informative as the optimal linear filter, but only if the analysis window is centered near the middle of the initiation of pursuit. Thus, simple counting of spikes and the weighted spike counting provided by the linear estimator agree about the amount of information in spike counts and the importance of the time of the spikes. We suspect that the importance of the time of spikes does not indicate a feature of the decoder used to convert neural spikes into eye movements. Instead, we think that the failure of time−translation invariance is a by-product of our analysis of behavioral variation. If we are trying to map spikes into the amplitude of the first mode of behavioral variation, the mapping may not be invariant to time translations.

The ideal duration of the window for counting spikes in our data was related to the discharge regularity of the particular group of neurons. For Purkinje cells, which are approximately half-Poisson, counting spikes beyond a 40 ms window increases the predictive power for variation in the behavior only slightly. Even counts of zero or one spike in a 5 ms window were (slightly) informative about the variation in eye speed. Spike counts in vestibular and abducens neurons predicted variation in the first principle component of eye speed in a way that improved more gradually as window size increased up to 60 or 150 ms. The need for longer windows in the latter neurons may be due to aliasing noise created by their more regular dis-

charge: FTNs are approximately one-quarter-Poisson while abducens neurons are ~1/20th-Poisson. Purkinje cells and FTNs are second-to-last-order and last-order interneurons. Thus, even for neurons that are close to the end of a sensory-motor pathway, spike counts in very brief windows provide asymptotic information about variation in behavior, and the time of the center of the counting window matters. For such short windows, the difference between spike counting and temporal codes starts to blur.

In the brainstem and cerebellum, it has been typical to characterize the neural code by calculating firing rate. In practice, firing rate is difficult to assess from a single instance of a spike train from a single neuron. Therefore, the preferred approach has been to average across repeated movements, something the nervous system cannot do. The field takes refuge in the notion that the nervous system can reconstruct firing rate in real time by counting spikes in small bins across single spike trains from many neurons in a neural population. Our analysis of simulated spike trains suggests that refuge may be insubstantial. Counting spikes of an individual Purkinje cell across a properly positioned 40 ms window is considerably more informative than counting spikes of 40 Purkinje cells in a 1 ms window. This difference results from the sub-Poisson nature of the firing, but may have general implications because few neurons in the brain are Poisson on all time scales. We note that synaptic potentials with a long time constant would take advantage of the
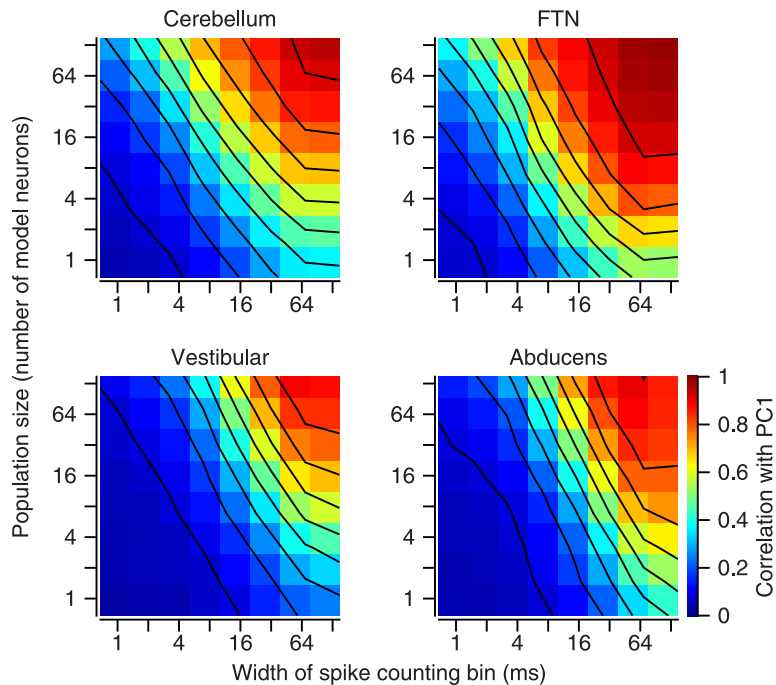
Figure 10 Trade-offs between the duration of the spike counting bin and the size of the population of model neurons. Different color plots show results based on different populations of neurons. Each pixel uses color to indicate the correlation between actual and predicted magnitude of the first principle component of eye speed, for the population size and bin width on the *y*- and *x*-axes. The black lines show iso-probability contours.

informative nature of accumulating spikes across time versus across neurons.

The filter that creates firing rate from a single spike train is nonlinear, messy, and difficult for a nervous system to implement. If the firing rate cannot be reconstructed optimally by counting spikes across neurons, then perhaps we should consider the neural code in terms of the occurrence of small numbers of spikes in restricted time windows. This view is supported by the analysis of real spike trains in our paper. It also agrees with the finding that information about target direction asymptotes in short time windows after MT neurons fire just a few spikes (Osborne et al., 2004), and with an analysis showing that the pattern of spikes and silences across small populations of MT neurons is more informative than simple spike counts (Osborne et al., 2008). It is worth noting that the original example of linear decoding of spike trains (Bialek et al., 1991) also was in a situation where the encoding of signals is nonlinear: decoding can be simpler than encoding.

What makes a single spike informative?
Our analysis of simulated spike trains implies that the strong neuron−behavior correlations of brainstem and cerebellar neurons (Medina and Lisberger, 2007; Joshua and Lisberger, 2014) are linked tightly to the predictive value of a single extra spike. We simulated the spike trains by assuming that the variation in both the spike trains of neurons and the eye movement behavior on a given trial result from a single source, namely variation in sensory estimates of target speed. Because the neural and motor

responses share a noise source, they are correlated. The correlation allows us to use the eye movement to create realistic model spike trains on each simulated trial, even though the real cause-and-effect relationship is in the opposite direction: trial-by-trial variations in the neural spike trains cause the trial-by-trial variations in eye movement.

The proposed correlations across the Purkinje cell population could, in principle, arise from local circuit properties in the cerebellar cortex. However, three facts argue for a sensory origin in terms of estimates of target velocity that are shared across the full population of Purkinje cells. First, more than 90% of the trial-by-trial variation in the initiation of pursuit projects onto the parameters of target motion: speed, direction, and time of onset (Osborne et al., 2005). Second, many individual neurons in area MT and in the smooth eye movement region of the frontal eye fields show trial-by-trial correlations between firing rate and the eye velocity at the initiation of pursuit (Hohl et al., 2013; Schoppik et al., 2008). If trial-by-trial variation in pursuit behavior arose downstream from MT, then trial-by-trial correlations between cortical neuron activity and pursuit velocity would not be expected (Schoppik et al., 2008). Third, we do not see evidence in the brainstem of separate, parallel noise sources from the two sides of the cerebellum (Joshua and Lisberger, 2014), as we would expect to if motor variation arose locally in the cerebellum.

The presence of a shared noise source to all neurons in a population, and its propagation all the way to the motor output, means that the probability of spiking as a function of time should follow highly correlated trajectories in all the model neurons in a given population. The spike trains of the

entire population should covary. The predictive value of a small number of spikes in a specific time window results from the dominance of the upstream sensory noise source over more local or private sources of noise in the individual neurons; a necessary corollary of this is that variations in the different neurons must be correlated. Thus, the organization of the sensory-motor system for pursuit allows small numbers of spikes to be a critical symbol in the neural code for motor control, even without any explicit mechanism for controlling spike timing or synchronizing spiking across neurons in a population. The predictive value of a single extra spike for the motor output really is a report of small errors in sensory estimation.

Our results align with the earlier analysis of the motion-sensitive neurons of the fly visual system during the complex, dynamic stimuli found in natural settings. In this system, de Ruyter van Steveninck and Bialek (1988) showed that single spikes and short spike sequences point to distinguishable brief motion trajectories. Bialek and Zee (1990) showed that the same finding emerges even for model neurons with Poisson spiking modulated by time-dependent signals. Bialek et al. (1991) showed that the information carried by each spike could be added up to give a running estimate of a velocity waveform with a precision close to the limits set by noise in the sensory periphery. More abstractly, Strong et al. (1998) measured the information carried by the spike trains of fly motion neurons, and Brenner et al. (2000) decomposed the spike trains into contributions from single spikes and intervals. We do not think it is fair to dismiss the results from the fly motion system because flies use a relatively small number of neurons to represent visual motion. The wide-field motion sensors in large blowflies take inputs from thousands of other neurons, as do neurons in the mammalian cortex. Importantly, Strong et al (1998) did the same analyses for data on the response of primate MT neurons to repeated presentations of a dynamic random dot stimulus (Britten et al., 1993). Under these dynamic conditions, MT neurons carry essentially the same number of bits per spike as do the fly neurons.

The dynamic nature of the stimulus/response situation is a critical piece of the informativeness of single spikes. Indeed, with quasi-static stimuli or movements, it is hard to imagine single spikes being very important. The work in flies pushes us away from thinking about static inputs and focuses attention on dynamic stimuli and short time scales. The initiation of smooth pursuit in monkeys is again a dynamic behavior that occurs over a brief time interval. In both cases, the relevant neurons generate only handful of spikes over the relevant time interval, so it seems hard to escape the conclusion that single spikes could be predictive.

Generalization to other neural systems?

To what extent can we expect that a single extra spike will be meaningful in neurons in other brain areas and sensory-motor systems? Churchland et al. (2006) showed neuron−behavior correlations in the premotor cortex, raising the possibility that a single spike might be informative in the most responsive of premotor cortex neurons. Manette and Maier (2004) performed an analysis related to ours on motor cortex neurons and came to similar, but slightly less specific, conclusions about a code based on the timing of small numbers of spikes. A clear example of the importance of single spikes is provided by the temporally sparse responses related to specific syllables in one group of neurons in an area of the songbird telencephalon called HVC (Hahnloser et al., 2002). The climbing fiber input to the cerebellum is widely viewed as an event detector, where a single spike is informative, even with recent evidence that the climbing fiber input can provide a graded signal (Najafi and Medina, 2013; Yang and Lisberger, 2014). There is evidence for a temporal code based on synchrony in the whisker sensation system (Jadhav et al., 2009; Bruno, 2011). Finally, there are multiple examples where the timing of single spikes is informative about the properties of a visual stimulus (Victor and Purpura, 1996; de Ruyter van Steveninck and Bialek, 1988). Still, there also are some counterexamples (Miura et al., 2012).

Our simulations suggest that single spikes will predict the impending behavior best in neurons with either large response magnitudes or low interspike interval variability. For a single spike train to account for 9% of the variance in motor activity ($r = 0.3$), the ratio of response magnitude to the coefficient of variation will have to be ∼150 (Fig. 7C,F). A cortical neuron with a CV of 1.0 would need to have a response magnitude of 150 spikes/s. On this basis alone, we would expect a single spike train to be informative about the impending smooth eye movement for the most responsive of MT neurons, if the neuron also has strong neuron−neuron correlations with its neighbors, which many MT neurons do (Huang and Lisberger, 2009; Hohl et al., 2013; Lee and Lisberger, 2013).

We have shown that a few spikes in small temporal windows could provide a meaningful code for motor behavior in neurons that are close to the final motor output. Still, the neural code will be identified by knowing how the response of a population of neurons actually is decoded, not just how it could be decoded. The answer will be in the timing of the barrage of spikes in a neural population, processed through potentially nonlinear mechanisms that determine the magnitude, time course, and spatiotemporal integration of postsynaptic potentials. For example, Chaisanguanthum and Lisberger (2011) used a model that was based on the timing of spike arrivals at a neural decoder to solve a problem that had been formulated previously in terms of the firing rates of neurons. Their model performed the same decoding of target speed as any number of traditional decoders that are based on firing rate, including maximum likelihood decoders (Deneve et al., 1999; Jazayeri and Movshon, 2006) and simple vector averagers (Salinas and Abbott, 1997; Groh, 2001; Churchland and Lisberger, 2001; Hohl et al., 2013). We imagine that similar solutions based on the timing of small numbers of spikes are likely to be possible for many coding and decoding problems in the brain.

References

Barlow HB (1956) Retinal noise and absolute threshold. J Opt Sci Am 46:634−639. Medline

Barlow HB, Levick WR, Yoon M (1971) Responses to single quanta of light in retinal ganglion cells of the cat. Vision Res Suppl 3:87–101. Medline

Bialek W, Zee A (1990) Coding and computation with neural spike trains. J Stat Phys 59:103–115. CrossRef

Bialek W, Rieke F, de Ruyter van Steveninck RR Warland D (1991) Reading a neural code. Science 252:1854–1857. Medline

Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck RR (2000) Synergy in a neural code. Neural Comp 12:1531–1552. Medline

Britten KD, Shadlen MN, Newsome WT, Movshon JA (1993) Responses of neurons in macaque MT to stochastic motion signals. Vis Neurosci 10:1157–1169. Medline

Britten KD, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100. Medline

Bruno R (2011) Synchrony in sensation. Current Opinion Neurobiology 21:701–708. CrossRef Medline

Chaisanguanthum KS, Lisberger KS (2011) A neurally efficient implementation of sensory population decoding. J Neuroscience 31:4868–77. CrossRef Medline

Churchland MM, Lisberger SG (2001) Shifts in the population response in the middle temporal visual area parallel perceptual and motor illusions produced by apparent motion. J Neurosci 21:9387–402. Medline

Churchland MM, Afshar A, Shenoy KV (2006) A central source of movement variability. Neuron 52:1085–1096. CrossRef Medline

Deneve S, Latham PE, Pouget A (1999) Reading population codes: a neural implementation of ideal observers. Nat Neurosci 2:740–745. CrossRef Medline

de Ruyter van Steveninck R, Bialek W (1988) Real–time performance of a movement sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. Proc R Soc London Ser B 234:379–414. CrossRef

Goris RL, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. Nat Neurosci 17:858–865. CrossRef Medline

Groh JM (2001) Converting neural signals from place codes to rate codes. Biol Cybern 85:159–165. Medline

Hahnloser RH, Kozhevnikov Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419:65–70. CrossRef

Hohl SS, Chaisanguanthum KS, Lisberger SG (2013) Sensory population decoding for visually guided movements. Neuron 79:167–179. CrossRef Medline

Huang X, Lisberger SG (2009) Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. J Neurophysiol 101:3012–3030. CrossRef Medline

Jadhav SP, Wolfe J, Feldman JE (2009) Sparse temporal coding of elementary tactile features during active whisker sensation. Nat Neurosci 12:792–800. CrossRef Medline

Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. Nat Neurosci 9:690–696. CrossRef Medline

Joshua M, Lisberger SG (2014) A framework for using signal, noise, and variation to determine whether the brain controls movement synergies or single muscles. J Neurophysiol 111:733–734. CrossRef Medline

Joshua M, Medina JF, Lisberger SG (2013) Diversity of neural responses in the brainstem during smooth pursuit eye movements constrains the circuit mechanisms of neural integration. J Neurosci 33:6633–6647. CrossRef Medline

Lee J, Lisberger SG (2013) Gamma synchrony predicts neuron-neuron correlations and correlations with motor behavior in extrastriate visual area MT. J Neurosci 33:19677–19694. CrossRef Medline

Lisberger SG (2010) Smooth pursuit eye movements: sensation, action, and what happens in between. Neuron 66:477–491. CrossRef Medline

Lisberger SG, Westbrook LE (1985) Properties of visual inputs that initiate horizontal smooth pursuit eye movements in monkeys. J Neurosci 5:1662–1673. Medline

Lisberger SG, Pavelko TA, Broussard DM (1994) Responses during eye movements of brainstem neurons that receive monosynaptic inhibition from the flocculus and ventral paraflocculus in monkeys. J Neurophysiol 72:909–927 [PMC][.CrossRef].

Manette OF, Maier MA (2004) Temporal processing in primate motor control: relation between cortical and EMG activity. IEEE Trans Neural Netw 15:1260–1267. CrossRef Medline

Medina JF, Lisberger SG (2007) Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. J Neurosci 27:6832–6842. CrossRef Medline

Medina JF, Lisberger SG (2008) Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. Nat Neurosci 11:1185–1192. CrossRef Medline

Medina JF, Lisberger SG (2009) Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum. J Neurophysiol 102:2039–54. CrossRef Medline

Miura K, Mainen ZF, Uchida M (2012) Odor representations in olfactory cortex: distributed rate coding and decorrelated population activity. Neuron 74:1087–1098. CrossRef Medline

Newsome WT, Wurtz RH, Dursteler MR, Wurtz RH (1985) Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. J Neurosci 5:825–840. Medline

Najafi N, Medina JF (2013) Beyond "all-or-nothing" climbing fibers: graded representation of teaching signals in purkinje cells. Front Neural Circuits 7:115. 10.3389/fncir.2013.00115] Medline

Osborne LC, Bialek WB, Lisberger SG (2004) Time course of information about motion direction in visual area MT of macaque monkeys. J Neurosci 24:3210–3222. CrossRef Medline

Osborne LC, Lisberger SG, Bialek W (2005) A sensory source for motor variation. Nature 437:412–416. CrossRef Medline

Osborne LC, Palmer SE, Lisberger SG, Bialek W (2008) The neural basis for combinatorial coding in a cortical population response. J Neurosci 28:13522–13531. CrossRef Medline

Rashbass C (1961) The relationship between saccadic and smooth tracking eye movements. J Physiol 159:326–338. Medline

Salinas E, Abbott LF (1997) Vector reconstruction for firing rates. J Comput Neurosci 1:89–107. Medline

Schoppik D, Nagel KI, Lisberger SG (2008) Cortical mechanisms of smooth eye movements revealed by dynamic covariations of neural and behavioral responses. Neuron, 58:248–260. CrossRef Medline

Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci 16:1486–1510. Medline

Shidara M, Kawano K, Gomi H, Kawato M (1993) Inverse-dynamics model eye movement control by purkinje cells in the cerebellum. Nature 365:50–52. CrossRef Medline

Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W (1998) Entropy and information in neural spike trains. Phys Rev Lett 80:197–200. CrossRef

Valbo AB (1995) Single afferent neurons and somatic sensation in humans. In The cognitive neurosciences (Gazzaniga M, ed.), pp 237-252.Cambridge MA: MIT Press.

Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-state analysis. J Neurophysiol 76:1310–1326. Medline

Yang Y, Lisberger SG (2014) Plasticity and cerebellar motor learning are graded by the duration of complex-spikes. Nature 510:529–532. CrossRef Medline