J Neurophysiol 123: 682–694, 2020.
First published December 18, 2019; doi:10.1152/jn.00117.2019.

RESEARCH ARTICLE | *Sensory Processing*

A simple linear readout of MT supports motion direction-discrimination performance

Jacob L. Yates,[1,2,3] Leor N. Katz,[2,3,4] Aaron J. Levi,[2,3,7] Jonathan W. Pillow,[5,6] and Alexander C. Huk[2,3,7]

[1]*Brain and Cognitive Science, University of Rochester, Rochester, New York;* [2]*Center for Perceptual Systems, University of Texas at Austin, Austin, Texas;* [3]*Department of Neuroscience, University of Texas at Austin, Austin, Texas;* [4]*Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland;* [5]*Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey;* [6]*Department of Psychology, Princeton University, Princeton, New Jersey; and* [7]*Department of Psychology, University of Texas at Austin, Austin, Texas*

Submitted 19 February 2019; accepted in final form 4 December 2019

Yates JL, Katz LN, Levi AJ, Pillow JW, Huk AC. A simple linear readout of MT supports motion direction-discrimination performance. *J Neurophysiol* 123: 682–694, 2020. First published December 18, 2019; doi:10.1152/jn.00117.2019.—Motion discrimination is a well-established model system for investigating how sensory signals are used to form perceptual decisions. Classic studies relating single-neuron activity in the middle temporal area (MT) to perceptual decisions have suggested that a simple linear readout could underlie motion discrimination behavior. A theoretically optimal readout, in contrast, would take into account the correlations between neurons and the sensitivity of individual neurons at each time point. However, it remains unknown how sophisticated the readout needs to be to support actual motion-discrimination behavior or to approach optimal performance. In this study, we evaluated the performance of various neurally plausible decoders, trained to discriminate motion direction from small ensembles of simultaneously recorded MT neurons. We found that decoding the stimulus without knowledge of the interneuronal correlations was sufficient to match an optimal (correlation aware) decoder. Additionally, a decoder could match the psychophysical performance of the animals with flat integration of up to half the stimulus and inherited temporal dynamics from the time-varying MT responses. These results demonstrate that simple, linear decoders operating on small ensembles of neurons can match both psychophysical performance and optimal sensitivity without taking correlations into account and that such simple read-out mechanisms can exhibit complex temporal properties inherited from the sensory dynamics themselves.

NEW & NOTEWORTHY Motion perception depends on the ability to decode the activity of neurons in the middle temporal area. Theoretically optimal decoding requires knowledge of the sensitivity of neurons and interneuronal correlations. We report that a simple correlation-blind decoder performs as well as the optimal decoder for coarse motion discrimination. Additionally, the decoder could match the psychophysical performance with moderate temporal integration and dynamics inherited from sensory responses.

MT; perceptual decision making; population coding; visual motion

INTRODUCTION

To make a decision about the direction of visual motion, a primate observer likely integrates over large populations of weakly correlated middle temporal area (MT) neurons (Cohen and Newsome 2009; Shadlen et al. 1996). Such pooling over neurons can be described in current parlance in terms of a linear decoder (Haefner et al. 2013; Moreno-Bote et al. 2014). In this simple read-out mechanism, the sign of the weights of the decoder indicates a neuron's membership to a pool and the magnitude of the weight indicates each neuron's contribution to its respective pool and, ultimately, the decision. This weighted sum of neural activity produces a univariate "decision variable," which can then be compared with a threshold to generate an estimate of the stimulus class (in this case, one of two diametrically opposite directions of motion).

In theory, the amount an individual neuron should contribute to a perceptual decision depends on how much information it carries (i.e., its sensitivity) and its correlation with other neurons (Moreno-Bote et al. 2014). Additionally, the correlations between neurons may themselves contain information about the stimulus that is not present in the rates alone (Latham and Nirenberg 2005). Here, we define 'the optimal linear decoder' to be a logistic regression decoding model fit via maximum likelihood. This model pools information in a manner that is sensitive to the correlations in multineuron responses. In contrast, suboptimal decoders that ignore the correlation structure may perform as well as the optimal decoder while being potentially simpler to implement and likely easier to learn (Berens et al. 2012). Additional aspects of readout, such as temporal integration, may also limit (or not limit) performance on a specific psychophysical task depending on the temporal structure of the neural responses (Levi et al. 2018; Osborne et al. 2004). Which aspects of population responses carry information about the stimulus, and whether that information is used by downstream areas, is an unresolved question in systems neuroscience.

In this study, we assess how simple the readout can be to match the performance of whole observers and optimal stimulus decoding, in the context of a coarse motion-discrimination task, similar to that used in the development of classic "read-

Address for reprint requests and other correspondence: J. L. Yates, Brain and Cognitive Science, Univ. of Rochester, Rochester, NY 14627 (e-mail: jyates7@ur.rochester.edu).

out" schemes (a.k.a., decoders). Prior work evaluating the responses of pairs of neurons in the middle temporal area (MT) found that a simple pooling model could account for the psychophysical sensitivity (Shadlen et al. 1996). Follow-up work also considered how limited temporal integration might also be a part of the read-out scheme (Cohen and Newsome 2009). In the present study, we take advantage of newer methods to record from ensembles of neurons and consider how activity of up to 21 MT neurons recorded simultaneously might be read out, either with respect to optimal decoding or to quantitatively account for the animal's performance in the task (Katz et al. 2016; Yates et al. 2017).

Our decoding analyses, applied to simultaneously recorded MT neurons, revealed that a read-out mechanism did not need to know the full joint statistics of the population (i.e., interneuronal correlations) to perform as well as an optimal decoder (i.e., one that was correlation aware). We also found that a temporally flat decoder could approximate optimal performance, even though the MT responses that it integrated evenly were themselves time varying. Such an architecture could match the psychophysical performance of the animals using roughly half of the stimulus duration.

Finally, using a computational model that considers the effects of shared variability on stimulus encoding (Ecker et al. 2014), we provide a simple scheme that reconciles conflicting experimental results about the role of neural correlations in optimal decoding. These analyses highlight the potential role of anesthesia-induced shared variability in complicating decoding. Taken together, these results show that a relatively simple linear decoder operating on small populations of MT neurons exhibits performance that is quasi-optimal, and which is comparable to the psychophysical performance of the monkeys, in terms of both the overall accuracy and the time-varying sensitivity to different stimulus epochs. Thus the simple read-out schemes posited in earlier work hold up in light of richer multineuron measurements that are now attainable, despite observations in such data of large interneuronal correlations and time-varying sensitivity.

METHODS

Electrophysiological recordings. All surgical and electrophysiological methods have been described previously (Katz et al. 2016; Yates et al. 2017) and were performed in accordance with US National Institutes of Health guidelines and the Institutional Animal Care and Use Committee at The University of Texas at Austin. Recordings were performed in two rhesus monkeys (*Macaca mulatta*; one male, *monkey P*; one female, *monkey N*) aged 14 and 10 yr, weighing 10 and 7.7 kg, respectively. Electrophysiological recordings were performed using linear electrode arrays (Plexon Uprobe/Vprobe; spacing 50–150 μm). For data from *subject P*, offline spike sorting was performed by hand refinement of a standard clustering algorithm (Plexon Offline Sorter v3). For data from *subject N*, spike sorting was performed offline using custom MATLAB code by fitting a mixture of Gaussian models to clipped waveforms in a reduced dimensional space (Ecker et al. 2014). For data from both monkeys, sorting was refined by maximum a posteriori estimation of a model where the multielectrode voltage was the linear superposition of Gaussian white noise and the spike waveforms (Pillow et al. 2008, 2013). Unit isolation quality was established using waveform signal-to-noise ratio (SNR) (Kelly et al. 2007) and interspike interval distributions (Hill et al. 2011). We included both single units and well-defined multiunit clusters in our analyses (130 units total).

MT was identified physiologically by depth and sulcal anatomy (using gray/white boundaries) and functionally by receptive field size and a preponderance of directionally selective neurons. *Subject N* had a custom titanium chamber positioned over the right hemisphere (L9, P2). *Subject P* had had a Cilux chamber (Crist Instruments) over the left V1 for a posterior approach to MT (L17, P17). Both approaches yielded small populations of MT neurons with overlapping spatial receptive fields. The dorsal approach (*subject N*) produced more heterogeneity in direction tuning than the posterior approach (*subject P*), presumably because the dorsal penetrations traversed multiple direction columns, whereas the posterior penetrations crossed cortical lamina.

During each session, retinotopy and direction selectivity were first mapped by hand using drifting dot stimuli. Selectivity was then refined and quantified using a pair of protocols to confirm that all simultaneously recorded units overlapped with the discrimination stimulus. For 44 units, we used a dynamic flow field to measure the direction preference and spatial receptive field (Mineault et al. 2012). Spatial velocity fields were estimated using the spike-triggered average velocity at all spatial locations. For 96 MT units, we measured the tuning function by presenting drifting 100% coherence dots in 12 evenly spaced directions and counting spikes. Tuning was estimated by least-squares fitting of a von Mises function to the spike rate. We mapped 23 MT units using both methods, which established a tight correspondence between the measured direction tuning from both approaches. The majority of these data have been described in previous reports (Katz et al. 2016; Yates et al. 2017), with one additional session that was subsequently collected in *monkey N*, and the analyses described in the present report are novel with regard to other publications.

Stimulus, apparatus, and task. Stimulus, apparatus, and task have been described in full detail previously (Katz et al. 2016; Yates et al. 2017). Briefly, stimuli were presented using Psychophysics Toolbox (Brainard 1997) and synchronized with electrophysiology using PLDAPS (Eastman and Huk 2012). The liquid crystal display had a resolution of 1,920 × 1,080 pixels and a refresh rate of 60 Hz and was corrected to have a linear gamma. Monkeys viewed the stimulus from a distance of 118 cm such that the 55-in. screen subtended 100° of visual angle.

Monkeys were trained to discriminate the net direction of motion in a field of flickering and drifting Gabor patches. A trial began with the appearance of a fixation point. Once the monkey acquired fixation and held it for 400–1,200 ms (uniform distribution), two targets appeared and remained visible until the end of the trial. At 200–1,000 ms after target onset, the motion stimulus appeared, centered at an eccentricity of 5–7°. The motion stimulus consisted of seven consecutive motion pulses, each lasting 9 or 10 video frames (150 or 166 ms; pulse duration did not vary within a session). The strength and direction of each pulse was varied by changing the proportion of Gabors that were drifting in a given direction. Monkeys were rewarded for choosing the target consistent with the sign of the sum of the individual pulses. Each experimental session presented motion in directions that were matched with the preferred null axis of the neurons we were recording from; however, because the Gabors can only drift orthogonally to their carrier orientation, we refer to that axis as "left" and "right" for convenience.

Neurophysiological analyses. Population decoding required a sufficient number of simultaneously recorded MT units. We included sessions for analysis as long as there were five or more MT units with sensitivity for direction (d-prime) > 0.2. This yielded 12 of 24 possible sessions. Spike times were binned at a 10-ms resolution. For visualization purposes, the peristimulus time histograms (PSTHs) in Fig. 2 were smoothed with a 50-ms boxcar filter. The tuning curves in Fig. 2 were fit with a von Mises function of the form

$$r(\theta) = a + b\exp\{-\kappa[\cos(\theta - \theta_{\text{pref}}) - 1]\}. \quad (1)$$

Spike count correlations (R_{sc}) were computed using the Pearson correlation coefficient of the stimulus-conditioned spike counts. We conditioned on the stimulus in two ways and report both. First, we computed R_{sc} using trials that were generated with exactly the same stimulus seed (i.e., frozen noise). Additionally, we computed R_{sc} by z-scoring spike counts for each net motion strength and then pooled all trials before computing the correlation coefficient. This second method allowed us to directly compare the real data correlations with shuffled data, both of which include a range of stimulus values.

Population decoder. We evaluated the population-level representation of motion direction using regularized logistic regression (Bishop 2006). This decoder is optimal in a maximum-likelihood sense for any case where the log-likelihood ratio is a linear function of the neural responses. This includes Poisson noise and Gaussian noise with fixed covariance as special cases. The probability that the direction, y, on any trial, t, was rightward given the decoding weights, $\mathbf{w}$, and neural response, $\mathbf{r}_t$, is

$$p(y_t = \text{right} \mid \mathbf{w}, \mathbf{r}_t) = \frac{1}{1 + \exp(-\mathbf{w}^{\mathrm{T}}\mathbf{r}_t)}, \quad (2)$$

where $\mathbf{w}^{\mathrm{T}}\mathbf{r}_t$ is the dot product between a single trial vector of spike counts (augmented by 1 to capture bias) and the weights. We estimated the weights using L2 regularization via the glmnet toolbox (Friedman et al. 2010):

$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} \sum_t \{y_t \mathbf{w}^{\mathrm{T}}\mathbf{r}_t - \ln[1 + \exp(\mathbf{w}^{\mathrm{T}}\mathbf{r}_t)]\} + \lambda \| \mathbf{w} \|_2^2, \quad (3)$$

where $\mathbf{r}$ is a matrix of spike counts on each trial augmented by a column of 1's to capture a bias term, $\mathbf{w}$ is a vector of weights (one for each neuron and one for the bias), y is a vector of the direction on each trial (1 for right, 0 for left). The hyperparameter, λ, dictates how much regularization is applied to the weights and was chosen using cross-validation within the training set (Friedman et al. 2010).

The optimal decoder operated on the total spike count of each simultaneously recorded neuron on each trial in a window from motion onset to 100 ms after motion offset. We compared the performance of this decoder with that of a correlation-blind (CB) decoder that did not have knowledge of the joint spiking statistics of the population. The CB decoder was estimated by shuffling the trial identity for each neuron within each condition for the training set. The CB decoder was then evaluated on data that were not shuffled (Latham and Nirenberg 2005).

We also tested whether the inclusion of quadratic terms [e.g., $\sqrt{\boldsymbol{r}(i)\boldsymbol{r}(j)}$] improved decoding performance (Fig. 3*D*). Because the instantaneous motion energy of our stimulus is known (Yates et al. 2017), we shuffled trials conditioned on the empirical signal strength of each trial (as opposed to the expected). Our shuffling procedure is as follows: First, we subdivided trials into conditions based on the net signal strength (signed number of Gabors) into m conditions, where m is the number of trials in the data set divided by the number of neurons recorded. The resulting number of conditions ranged from 16 to 35, with a median of 21.5 discretized signal strengths. Second, we shuffled trials within each of these conditions, enforcing that no pairs had spikes that came from the same trial. To ensure that we preserved the sensitivity (mean response and variance) of each neuron while only breaking the correlation structure, we reshuffled trials until there was no significant change in d-prime for any neuron.

To evaluate the effect of temporal integration on decoding performance, we trained the decoder with increasing window size in 10-ms steps (Fig. 5). To evaluate the instantaneous sensitivity of the population over time, we retrained the weights in 10-ms steps, counting spikes over a sliding 100-ms window (Fig. 6). In each case, we trained the decoders using the net direction of motion on each trial. To evaluate the effect of correlations on the instantaneous decoder, we shuffled spike counts (within the sliding 100-ms window) by conditioning on the exact value of each pulse separately. This procedure had the effect of shuffling correlations while matching the mean and variance for a particular pulse. Of course, each individual pulse only affects spike counts for a small time window (Yates et al. 2017), so we repeated this shuffling for each pulse separately and then performed our sliding window decoding analysis. These two steps produced seven pulse-dependent shuffled-accuracy traces. We stitched the seven traces together to form a single accuracy trace by extracting the values from 50 to 200 ms after the onset of each pulse (this window was selected based on the time-lagged effect of each pulse, estimated using pulse-triggered reverse correlation; Yates et al. 2017). This stitched accuracy trace was then directly compared with the instantaneous decoder trained on real (unshuffled) data.

To measure the optimal temporal weights (Fig. 6), we trained a spatiotemporal decoder that had temporal weights for each neuron. We used 50-ms time bins, which resulted in 25 temporal parameters for each neuron. We regularized the estimates of the weights by smoothing in time using a Gaussian prior over w, $\mathbf{w} \sim N(0,C)$, where $C^{-1} = \mathbf{w}^{\mathrm{T}}\mathrm{D}\mathbf{w}$ and D is a block-diagonal derivative matrix such that D penalizes the squared difference between parameters over time, but not across neurons. We estimated λ using cross-validation within the training set. We resampled training data sets with replacement 1,000 times and evaluated decoding accuracy on the withheld data for each sample. We then corrected for any potential overfitting using the ".632+ bootstrap" (Efron and Tibshirani 1997). This method provides a low variance estimate of the true decoding performance (Yousef et al. 2004). We generated 95% confidence intervals both for the weights and accuracy from the bootstrap distribution.

We evaluated the performance of single neurons at different stimulus strengths by evaluating the performance of a linear classifier trained to discriminate the net direction of the stimulus using only the spikes from that neuron in the same way that we measured the population accuracy. This approach diverges from classic estimation protocols for neurometric functions that typically use the area under the receiver operating characteristic (ROC) curve to quantify percent correct (Britten et al. 1992; Uka and DeAngelis 2004). In this task, which only has one stimulus presentation per trial, the area under the ROC curve would require the assumption of a hypothetical "anti-neuron" to be interpreted as the percent correct. This assumption is problematic for our population recordings, which often include neurons of opposite preferred directions; thus it was unnecessary to explicitly build in an "anti-population."

Psychometric/neurometric threshold. To quantify psychophysical and neuronal thresholds, we fit a cumulative Weibull function to the accuracy of the monkeys and decoders:

$$p(\text{correct}) = 1 - 0.5\exp\left[-\left(\frac{s}{\alpha}\right)^{\beta}\right], \quad (4)$$

where s denotes the net motion strength, α is the threshold on motion strength (at the 82% level), and β is the slope of the function. We used the fitted values of α for comparisons of neuronal and psychophysical threshold. The cumulative Weibull was fit by maximizing the likelihood of individual trials assuming Bernoulli distributed responses. For the neural decoders, trial responses were stitched together using the average response for each trial on the bootstrapped test sets.

Psychophysical/decoder kernel. To measure the contribution of each pulse to the monkey's choice (or the decoder output) on each trial, we again used logistic regression. In this case, logistic regression measures the psychophysical weights of the monkey for each pulse (Katz et al. 2016; Yates et al. 2017). The probability of the monkey's choice on a single trial is given by

$$p(Y \mid X, \beta) = \frac{\exp(Yb)}{1 + \exp(b)}, \quad (5)$$

where $b = \beta_0 + \sum_{i=1}^{7} \beta_i X_i$ and the choice $Y \in \{0,1\}$. X is a vector of the seven pulses. This model was fit using glmfit in MATLAB. Error

bars were derived from the matrix of partial second derivatives (i.e., used the square root of the diagonal of the inverse Hessian). This model is equivalent to the decoding model, except the weights are over the motion pulses and the decoded value is the monkey's choice on each trial. We applied this analysis to the choice output from the decoder to measure the contribution of each pulse to the decoder output. This is analogous to the slope of a neurometric function except expanded over time.

Shared variability simulation. To understand the effects of shared variability on optimal linear decoding, we simulated neuronal populations under different amounts of stimulus-driven (i.e., information limiting; Moreno-Bote et al. 2014) and stimulus-independent shared noise. Our population model is inspired by the Gaussian-process factor analysis model used to describe the effects of anesthesia in Ecker et al. (2014). The response of n neurons on each trial, t, is an $[n \times 1]$ vector of spike counts, $\mathbf{y}(t)$, resulting from a Poisson draw from a vector of spike rates for that trial, $\mathbf{r}(t)$:

$$\mathbf{y}(t) \sim \text{Poisson}[\mathbf{r}(t)]. \quad (6)$$

Regardless of the number of neurons, the population spike rates are driven by a mixture of only two scalar variables: the stimulus value, $s(t)$, and the stimulus-independent latent (e.g., level of anesthesia), $a(t)$, on each trial. Each neuron's spike rate is a weighted sum of these two inputs, which is then passed through a static nonlinearity:

$$\mathbf{r}(t) = f\{\mathbf{A}^{\mathrm{T}}[s(t) + \xi, \lambda a(t)] + \mathbf{b}\}, \quad (7)$$

where ξ is a Gaussian random variable that adds shared noise to the stimulus representation, $\xi \sim N(0,\sigma)$, $a(t)$ is the value of a shared latent on each trial, and $\mathbf{A}$ is a $[2 \times n]$ matrix that describes how each neuron weighs the two latents (a and s). The total magnitude of anesthesia is multiplied by a scalar, λ. $\mathbf{b}$ is an additive offset for each neuron that scales the baseline firing rate. f is a quadratic nonlinearity, which was chosen to match the square-root transform applied to the data in Ecker et al. (2014). Expanding the quadratic reveals an interaction term between the anesthesia and stimulus latents, indicating that they interact both additively and multiplicatively to affect firing rates. For the simulations in Fig. 7, $s(t)$ was drawn randomly from two values (left or right) with equal probability. $a(t)$ was a Gaussian random variable with zero mean. a was smoothed with a 100-trial Hanning window to create slow fluctuations. $\mathbf{A}$ was generated from a random uniform distribution size 2 by N neurons ($N = 20$ in our simulations to match population sizes from recent empirical data (Chaplin et al. 2018; Yates et al. 2017)). $\mathbf{b}$ was fixed to 1 for all neurons, because our results are independent of this term. We simulated 10,000 trials of weak left or right motion and generated spike counts for populations with different levels of shared stimulus noise (ξ) and stimulus-independent noise (λ). We then performed our primary decoding analysis (train on real vs. train on shuffled, test on real) on these simulated populations. We chose this model to simulate the effects of anesthesia because it has been successful at describing the effects of anesthesia in primary visual cortex with a single latent (Ecker et al. 2014) and because it allows us to directly parameterize information-limiting noise along a single stimulus latent.

RESULTS

We measured the activity of up to 21 simultaneously-recorded MT neurons (median = 10) while monkeys performed a motion-discrimination task (Fig. 1*A*). On each trial, the monkey indicated his or her choice about the net direction of motion with an eye movement to one of two targets (Katz et al. 2016; Yates et al. 2017). The units composing our population exhibited responses during motion that are strongly dependent on the net motion strength (Fig. 1*B*) and which closely resemble the responses to low-coherence moving dots reported previously (Britten et al. 1993). The sensitivity of neurons with responses like these have been studied extensively in single-unit or cell pair recordings during a similar motion-discrimination paradigm (Bair et al. 2001; Britten et al. 1992; Cohen and Newsome 2009; Zohary et al. 1994).

A large proportion of MT neurons are directionally selective when recorded with single electrodes (Albright 1984), and we found a similar preponderance of directionally selective cells in our array recordings. Figure 2*A* shows the responses of 16 simultaneously recorded neurons to different directions of motion. These cells were also responsive in a direction-selective manner during the motion-discrimination task (Fig. 2*B*). We used a generic population decoding approach (Berens et al. 2012; Georgopoulos et al. 1986) to study the representation of motion in the joint activity patterns of simultaneously recorded units. The logistic regression decoder we employed takes a weighted sum of spike counts and passes that through a sigmoid function to get the probability of a rightward direction for each trial (Fig. 1*C*). The monkey's task can be thought of in the same classification framework (Bishop 2006): given the noisy response of MT neurons, which of the two possible task classes was most likely (e.g., left or right)? Figure 1*D* depicts the conceptual framework.

A simple, correlation-blind, linear decoder performs as well as the optimal decoder. Before evaluating whether a neural readout could gain information from the correlation structure, we first established that interneuronal correlations were present in our data set. We analyzed 579 cell pairs from the 12 sessions and found that spike count correlations calculated on frozen noise trials were mostly positive (Fig. 3*A*), with the mean correlation significantly larger than zero ($r = 0.098$; $P < 0.001$, t test). In theory, optimal linear decoding requires knowledge of the covariance of neurons; however, it is possible that the accuracy of a decoder would be insensitive to suboptimal weighting (Berens et al. 2012).

To investigate whether knowledge of the full joint statistics of the population is required to maximize accuracy, we compared a decoder trained on the real data with one that was trained on shuffled data. Shuffling the trials of each neuron within each stimulus condition has the effect of breaking the covariance structure of the simultaneously recorded neural ensemble (Graf et al. 2011). We shuffled trial identities for each neuron within each motion strength. We confirmed that mean spike count correlations for shuffled data were not statistically different from zero [mean shuffled $R_{sc} = -0.001 \pm 0.0017$, $P = 0.65414$, signed rank (579) = 130,854], whereas within-condition correlations on real data were indistinguishable from those of frozen trials [mean $R_{sc} = 0.094 \pm 0.0053$, $P = 0.1713$, paired signed rank (579) = 89,464]. Figure 3*B* shows the relationship between tuning preference similarity and spike count correlations for our data and the same analysis after shuffling.

Figure 3*C* shows the decoder weights trained on shuffled data compared with the weights trained on real data. The difference in learned weight patterns for the full decoder and the correlation-blind (CB) decoder were subtle, and the decoders agreed on 95.5 ± 0.81% of trials. These subtle differences were not sufficient to produce a meaningful difference in direction discrimination (decoding) accuracy. Figure 3*D* shows the performance of the shuffled decoder compared with the optimal decoder. The accuracy of the shuffled decoder was not significantly different from that of the optimal decoder [geo-

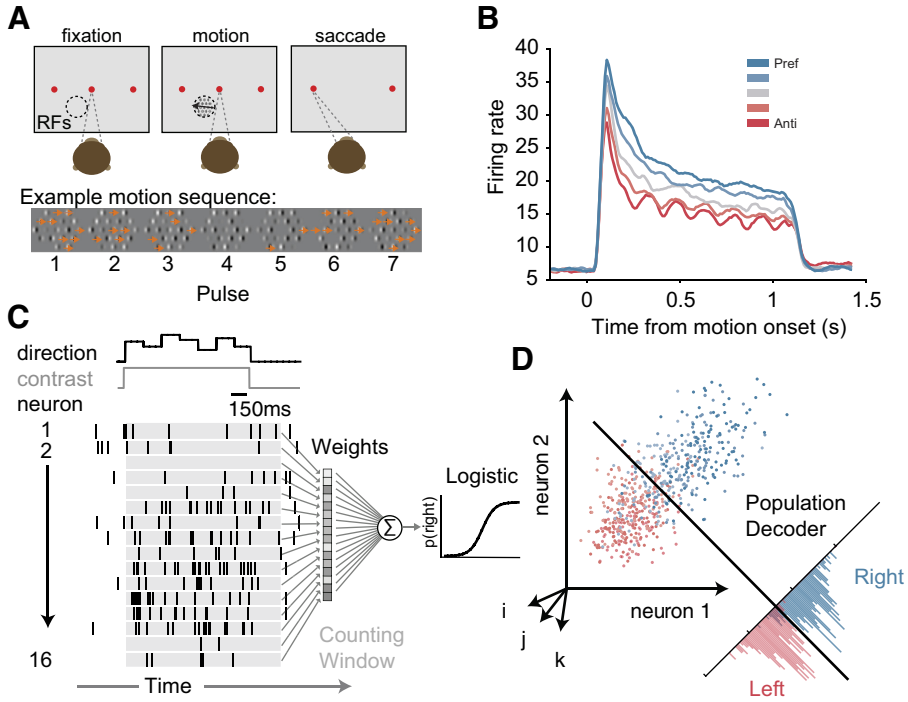

Fig. 1. Motion-discrimination task, middle temporal area (MT) responses, and decoding approach. *A*: the monkey holds fixation while viewing a fixed-duration motion-pulse stimulus and then waits for a go signal before indicating his or her choice about the net direction of motion with an eye movement to 1 of 2 saccade targets. The motion stimulus on each trial consists of seven 150-ms-long motion pulses, where the strength of each pulse is given by the proportion of drifting Gabor elements. An example motion pulse sequence is shown at *bottom*. RF, receptive field. *B*: the average responses of all MT neurons contributing to the population analyses ($n = 130$) were modulated by the net motion strength. Average firing rates are sorted into quantiles of the net motion strength (indicated by color and shade). Pref and Anti, preferred and anti-preferred direction of motion. *C*: example single-trial neural responses and decoding model. The optimal linear decoder integrates spikes over time (gray-shaded window) and then takes a weighted sum of those spike counts. The resulting scalar is passed through a sigmoid function to give the probability of rightward motion [p(right)] on each trial. The spikes on this trial come from the 16 simultaneously recorded MT neurons shown in Fig. 2. *D*: the spike count of this ensemble of 16 neurons can be represented as a point in a 16-dimensional space, where each axis is the spike rate of one of the neurons (*neurons 1, 2, i, j*, and *k*). The weighted sum of the ensemble activity (i.e., the linear decoder) projects the high-dimensional population activity to a one-dimensional decision axis.

metric mean ratio: 1.0041 ± 0.0041, signed rank(11) = 18, $P = 0.109$] and did not differ on any individual session. We also compared the optimal linear decoder to one with quadratic terms and found that accounting for pairwise interactions between neurons did not improve performance (Fig. 3*E*). We confirmed this similarity in performance between the optimal linear decoder and the CB decoder was not a product of regularization by investigating the hyperparameters from our fits. Cross-validated λ values for the two models were small (0.0868 [0.026,1.479] and 0.157 [0.029,2.292] for optimal and CB, respectively; brackets indicate 95% confidence intervals) and not significantly different from each other ($P = 0.344$, bootstrapped t test).

With the assumption that the monkey uses one set of pooling weights for all signal strengths, the above analyses indicate that correlation-blind decoding is indistinguishable from optimal decoding in this task. However, it is possible that using different decoders at each signal strength could produce better performance than a single global decoder. To test this, we trained and tested separate decoders at five quantiles of the net motion strength. We then evaluated that a correlation-blind decoder performed the same as the optimal decoder at all stimulus strengths. The geometric mean ratio of the optimal decoder to the shuffled decoder was 1.008 ± 0.2, 1.005 ± 0.15, 1.01 ± 0.017, 1.003 ± 0.006, and 1.008 ± 0.0049 for each of the five motion strengths tested, from weak to strong, respectively. Across sessions, the optimal decoder performed significantly better at the strongest signal value [$P = 0.0005$, signed rank(11) = 78]; however, the other signal strengths showed no significant differences [$P = 0.677$, signed rank(11) = 45; $P = 0.791$, signed rank(11) = 43; $P = 0.233$, signed rank(11) = 55; $P = 0.129$, signed rank(11) = 59]. Additionally, for individual sessions, the optimal decoder was never outside the 95% confidence intervals of the CB decoder for any stimulus value. There was a weak but significant dependence of the optimal-to-blind accuracy ratio on the number of neurons [linear model $\beta_1 = 0.0022 \pm 0.0007$, $t(58) = 3$, $P = 0.00389$], suggesting that if we were to record from more neurons, it is possible that we would see some dependence on correlations for these local (stimulus dependent) decoders. We then compared the performance of multiple stimulus-dependent decoders with that of a single, global decoder. To summarize performance across signal strengths, we fit a cumulative Weibull function to the trial-by-trial performance and estimated the threshold. We compared the threshold stimulus for the global decoder with the threshold for the local decoders. The thresholds were not significantly different [geometric mean ratio: 1.025 ± 0.08, $t(11) = 0.865$, $P = 0.405$, 2-sided t test]. This was true for a correlation-blind global decoder, as well [geometric mean ratio: 0.97 ± 0.1, $t(11) = 0.4173$, $P = 0.685$]. This indicates that a single set of pooling weights performs indistinguishably from multiple

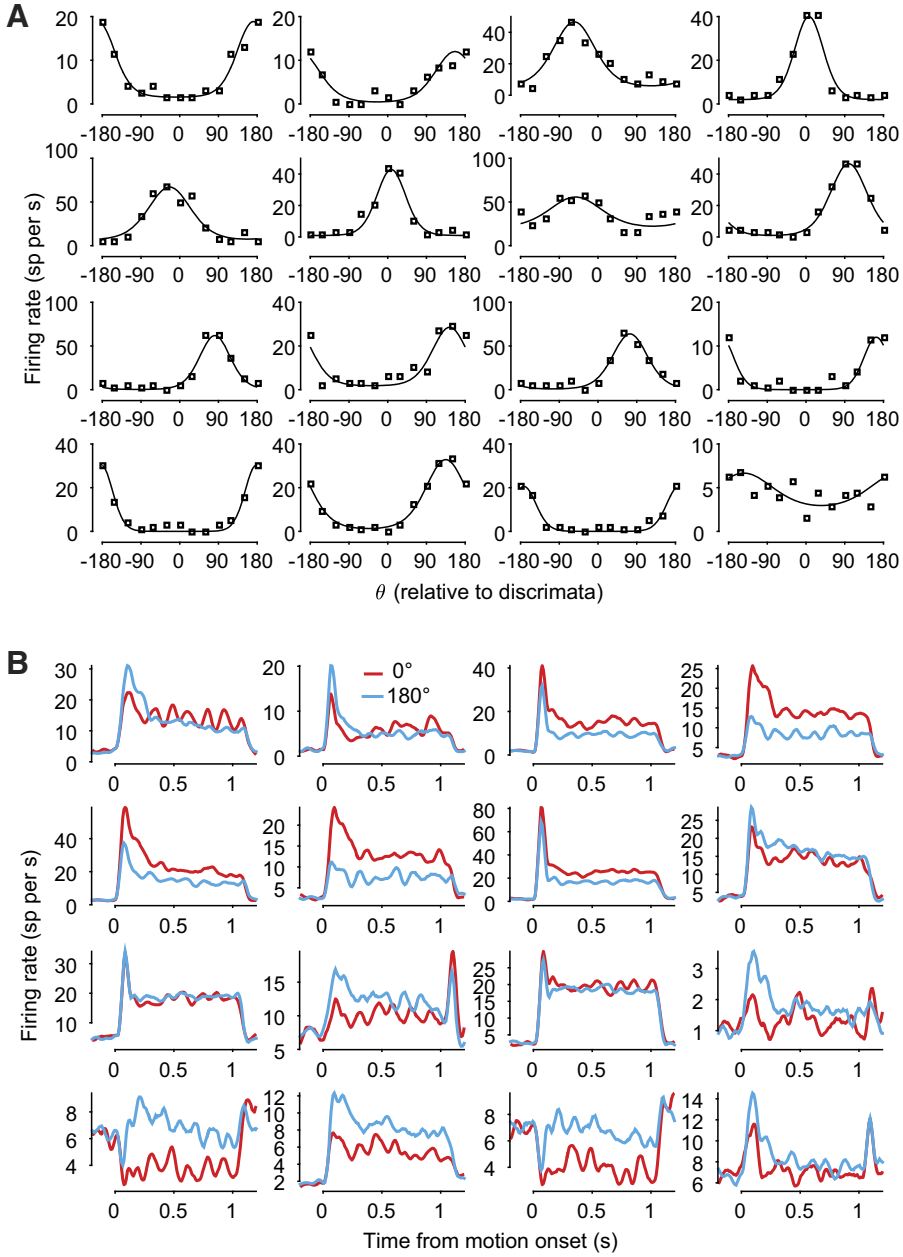

Fig. 2. Example tuning functions and responses of 16 simultaneously recorded middle temporal area (MT) neurons. *A*: firing rate (sp per s, spikes/s) as a function of direction for each of the 16 MT neurons as measured during a separate mapping protocol. Points indicate the measured firing rate, and lines are fits from a von Mises function. *B*: direction-sorted peristimulus time histograms (averaged across signal strengths) for the same neurons as in *A*, measured during the main motion discrimination task.

stimulus-dependent decoders and that the global decoder can ignore neural correlations. Having established that the decoders we evaluated could effectively be insensitive to the joint firing patterns of MT neurons despite conventional levels of interneuronal correlation, we compared the performance of the optimal linear decoder with the sensitivity of single neurons and the psychophysical sensitivity of the monkeys.

Comparison of psychophysical and neurometric performance. We compared the performance of the optimal linear decoder to the monkey's accuracy on each session. We did so in two ways: by comparing total accuracy and by comparing thresholds inferred by fitting the decoder output (on test data) with psychometric functions. The first is a measure of total performance, and the second reveals sensitivity to the stimulus. Figure 4*A* shows the approach for an example session. Across all sessions, the population decoder performance unsurprisingly depended on the performance of the best single neuron (linear model: $\beta_1 = 1.08 \pm 0.18$, $t = 5.9024$, $P = 0.0002$), but the population consistently outperformed the best single neuron (Fig. 4*B*). The best single neuron provides a lower bound on the performance of a neural population decoder, and our results show that adding other neurons provides additional information.

To establish a baseline for the range of sensitivities supported by the units in our data set, we compared the total accuracy of each unit to the monkey's behavior (Fig. 4*C*). We also compared the ratio of the neurometric and psychometric thresholds (Fig. 4*D*). Forty-five of 130 neurons had neurometric-to-psychometric ratios > 10 and were excluded. The geometric mean of the neurometric-to-psychometric ratio for the remaining neurons was significantly greater than 1 [2.6608 ± 0.36, $t(83) = 10.2636$, $P = 1.971 \times 10^{-16}$], indicating that the monkey performed better than the individual units, on average (even after exclusion of the worst-performing

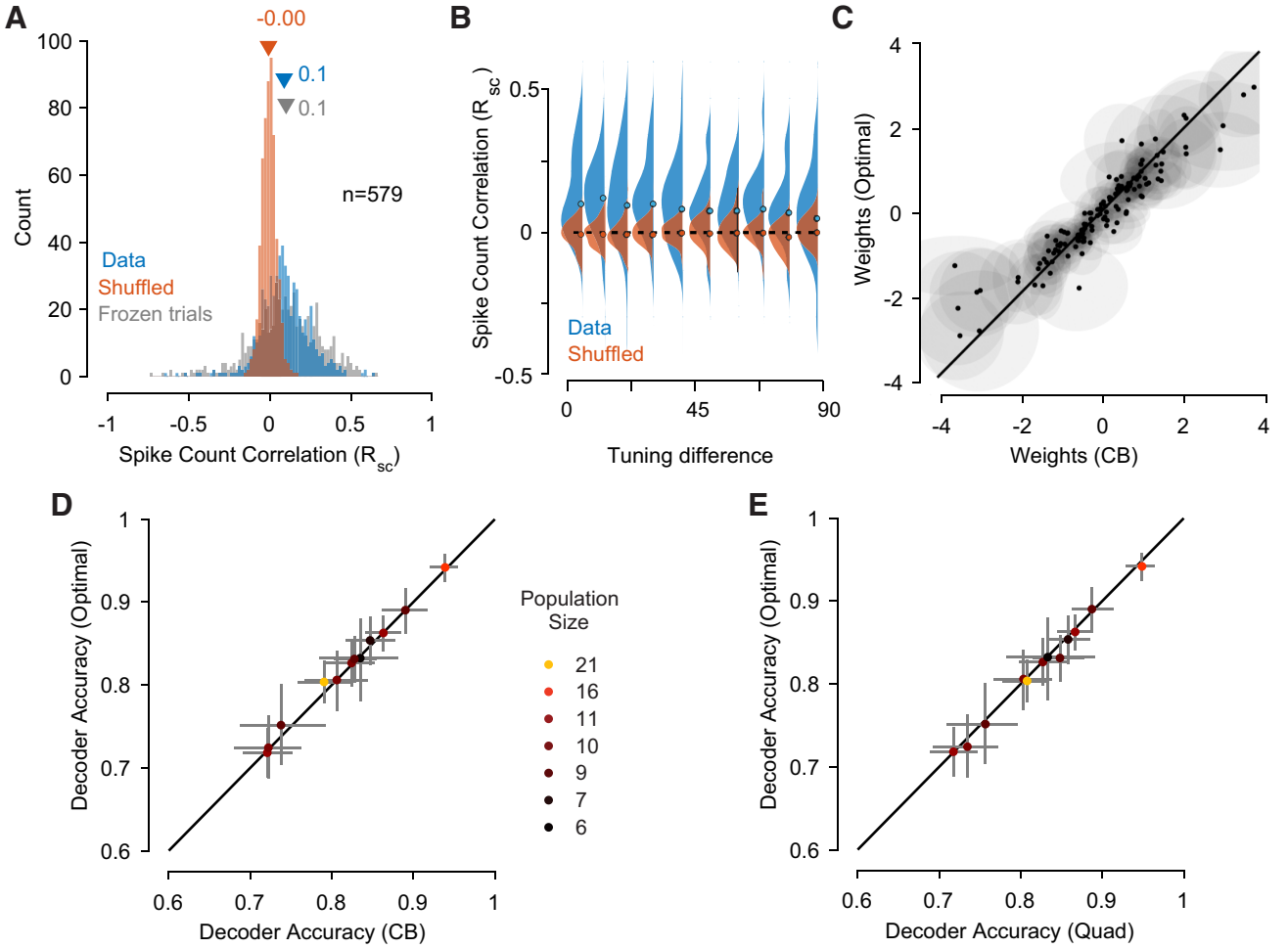

Fig. 3. Decoder performance. *A*: pairwise noise correlation measured on repeat trials (gray) and when conditioned on signal strength (blue) for 579 middle temporal area (MT) cell pairs were largely positive. Inverted triangles indicate the average noise correlation (0.1). Shuffling trial identities removed the correlation structure (orange). *B*: count correlations depend on the difference in preferred tuning for pairs (blue). Shuffling trials breaks this relationship (orange). *C*: individual neuron weights for the correlation-blind (CB) decoder and the optimal decoder were highly correlated. Error clouds are 1 SD of a Gaussian fit to the individual bootstraps. Oriented structure in the error clouds means that the exact value of the weights depended on the particular random training data more than the shuffling. *D*: comparison of accuracy of a CB decoder with accuracy of a decoder trained on the full response distribution. Error bars are bootstrapped 95% confidence intervals. Color of each point indicates the population size for that session. *E*: comparison of accuracy of a decoder having quadratic interaction terms (Quad) with accuracy of the full linear decoder. Error bars are bootstrapped 95% confidence intervals. Colors indicate the population size for each session. R_{sc}, spike count correlation.

neurons, as described above). However, across sessions, the threshold of the population decoder was not significantly different from the monkey's performance [geometric mean ratio of accuracy = 1.043 ± 0.30, P = 0.677, signed rank(11) = 45.00, 2-sided Wilcoxon signed-rank test]. This was true of the total accuracy, as well. The population accuracy was 0.99 ± 0.04% of the monkey's performance, on average. On 4/12 sessions, the population significantly outperformed the monkey. On 3/12 sessions, the monkey significantly outperformed the neural population. These results support the insight that although single units in our data set were not sufficiently sensitive to support the psychophysical accuracy, small ensembles of neurons could surpass the accuracy level of the monkey.

One possible explanation for the neural decoders outperforming psychophysical accuracy is that the monkeys did not use all of the stimulus to inform their choices (Cohen and Newsome 2009; Kiani et al. 2008; Mazurek et al. 2003). Given that possibility, we next evaluated the temporal integration required to support the psychophysical behavior of the monkey.

Consequences of temporal dynamics in MT for decoding performance. Previous studies have demonstrated that primate observers likely do not use the full time the stimulus is available to make up their mind (Cohen and Newsome 2009; Kiani et al. 2008; Mazurek et al. 2003). Because the decoders we evaluated thus far have all used the entire motion epoch, it is possible that the ability of a neural decoder to outperform the monkey is due to suboptimal temporal integration by the monkey (Cohen and Newsome 2009). However, in a parallel to interneuronal correlations, the presence of temporal autocorrelation can limit benefits of integration, meaning that performance saturates with time (Goris et al. 2018; Osborne et al. 2004). As such, we evaluated the amount of integration time required to maximize performance and compared that to the monkey's performance.

We evaluated the performance of an optimal decoder on a cumulative window from motion onset until 100 ms after

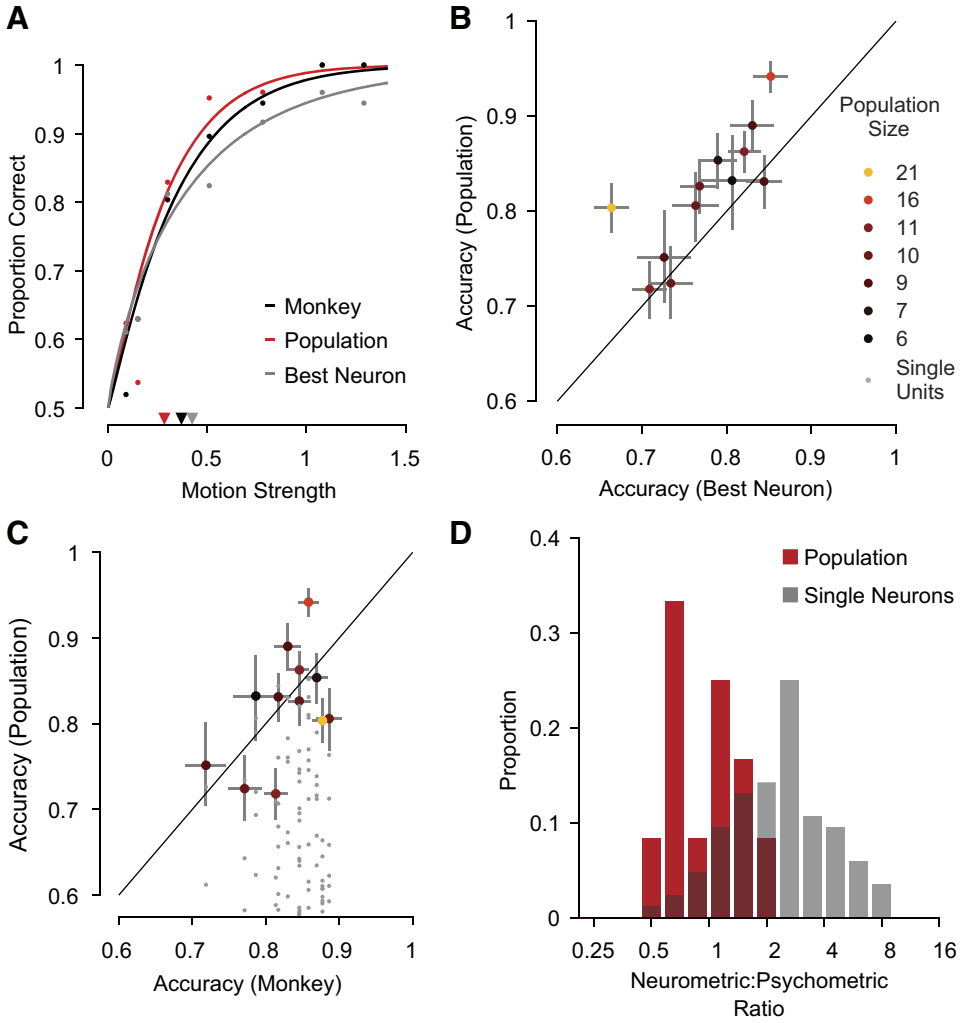

Fig. 4. Comparison of neural decoder and psychophysical performance. *A*: example-session psychometric (black) and neurometric functions (gray = best single neuron, red = population decoder). Inverted triangles indicate threshold performance (82% correct). *B*: comparison of accuracy of the population decoder with accuracy of the best single unit from the corresponding session. Color indicates population size, as in Fig. 3. *C*: proportion correct for the monkey plotted against the cross-validated accuracy for the population decoder (colored points) and single neurons (gray points). Color indicates population size. *D*: ratio of the neurometric-to-psychometric threshold at the 82% level (as described in *A*). Values < 1 indicate the neurons are outperforming the monkey.

motion offset (Fig. 5*A*). The decoder showed diminishing returns for integration, saturating as early as 225 ms into motion with a mean saturation time of 610 ± 62 ms (mean $\pm$ SE). We then compared the cumulative decoder to the psychophysical performance and found, on average, 548 ± 94 ms were required to match the monkey's performance, similar to the typical average reaction time of monkeys performing a similar motion-discrimination task (Cohen and Newsome 2009; Huk and Shadlen 2005; Kiani et al. 2008). Because our stimulus has independent motion pulses, we were able to characterize the pulses that contribute most to behavior, which has been used in the past to constrain the temporal integration of the animal during fixed duration tasks such as ours (Kiani et al. 2008). Use of logistic regression to measure the psychophysical kernel revealed the monkeys used motion early in the trial most to guide their choice (Fig. 5*B*), consistent with many prior measurements of monkey behavior (Kiani et al. 2008; Nienborg and Cumming 2009; Yates et al. 2017). We then ran the same analysis on choices generated by the optimal neural decoder that perfectly integrated MT (Fig. 5*B*). Although perfect integration of MT might be expected to produce flat psychophysical weighting, this decoder's kernel exhibited substantial early weighting. This happens because MT activity has effectively reweighted the stimulus, producing at least some of

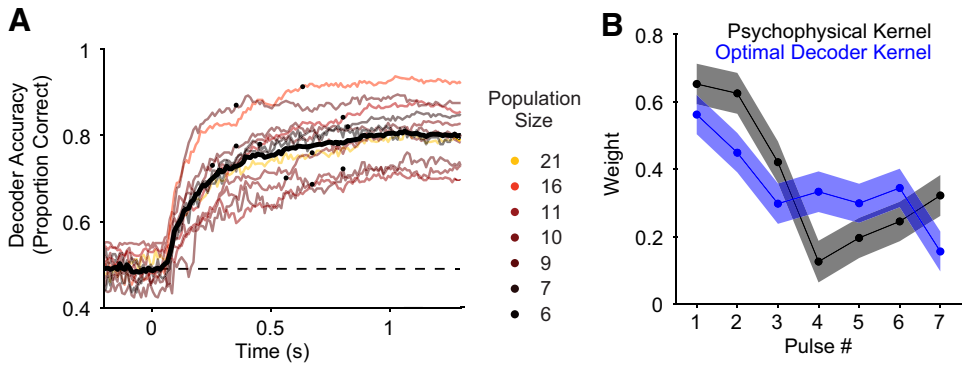

Fig. 5. Temporal integration and stimulus weighting. *A*: accuracy of direction decoding over different-sized counting windows for each session (hot color map) and session average (black). Saturation point is indicated by black point on each trace. *B*: psychophysical kernel (black) compared with the same analysis applied to the output of the optimal decoder that perfectly integrates middle temporal area throughout the entire stimulus (blue).

the time-varying weight pattern of the psychophysical kernel (Yates et al. 2017).

Temporal sensitivity to motion. We reported previously that direction selectivity in MT decays throughout motion viewing (Yates et al. 2017), and in the present study we showed that perfectly integrating real MT neurons produces early weighting. To further explore early weighting in MT, we evaluated the temporal information available in the population by decoding from the instantaneous spike rate in a 100-ms sliding window. We performed this analysis using a fixed set of weights trained on the trial spike count and using weights that were retrained in each time bin. We additionally confirmed the results of our shuffling analysis at this finer temporal resolution using a decoder that was trained on spikes that were shuffled by the value of each pulse independently. Shuffling each pulse produced seven accuracy traces, which we stitched together after accounting for the time lags affected by each pulse (see METHODS). Figure 6*A* shows the stitched shuffled accuracy compared with the instantaneous decoder as well as the fixed decoder. Even at a fine temporal resolution, the correlation-blind decoder performed indistinguishably from the optimal decoder. The geometric mean ratio of accuracy was 0.993 ± 0.002, and across sessions there was no significant difference [$P = 0.85775$, $t(1,715) = -0.17927$, t test, tested on all time points from all sessions]. The fixed decoder showed similar temporal fluctuations to the instantaneous decoder (Fig. 6*A*), and the fixed weights performed almost as well as the instantaneous weights ($97 \pm 0.2\%$ of fixed). This similarity was due to the stability of the representation of direction in MT across time. Regardless of when we trained a decoder (as long as it was during motion), the performance across time was roughly the same (Fig. 6*B*).

The instantaneous decoding analysis revealed that the population appeared most sensitive immediately following motion onset (Fig. 6, *A* and *B*). This is not due to the stimulus itself, which was equally informative about the net direction over time, but rather reflects the dynamics of the representation in MT. Across our sessions, the peak performance was reached, on average, 285 ms after motion onset, and all sessions reached peak performance within the first 500 ms (Fig. 6*C*). To assess the relative sensitivity of motion early in the trial compared with late in the trial, we evaluated the cross-validated accuracy of a decoder trained on the first half of motion and compared it with one trained on the second half. Figure 6*D* shows that in almost all sessions (10/12), responses from the first half of motion viewing were significantly more sensitive to the net direction than responses from the second half.

We wondered whether this change in sensitivity would make early-stopping or overweighting spikes early in the stimulus epoch optimal based solely on the sensitivity of the population. To test this, we trained a decoder with temporal weights for each neuron. If the pattern of weights decreased over time, we would conclude that the optimal weighting pattern to read out the population was early weighting. In contrast to this, the temporal weights were largely flat, suggesting the optimal weight pattern for decoding direction is to integrate all time with roughly equal weight (Fig. 6*E*).

Shared variability explains when correlation-blind decoding is optimal. To date, at least three other empirical studies have asked whether correlation-blind decoding is optimal for cortical neural populations (Berens et al. 2012; Chaplin et al. 2018; Graf et al. 2011). To our knowledge, both studies performed with anesthetized animals found that correlation-blind decoding is not optimal (Chaplin et al. 2018; Graf et al. 2011), whereas both studies of awake animals (including this study) found that correlation blind decoding is effectively optimal.

To understand the potential effects of anesthesia on the optimal decoding weights, we extended a model from Ecker et al. (2014) to simulate the effects of anesthesia on shared variability. We then ran our primary decoding analyses on the

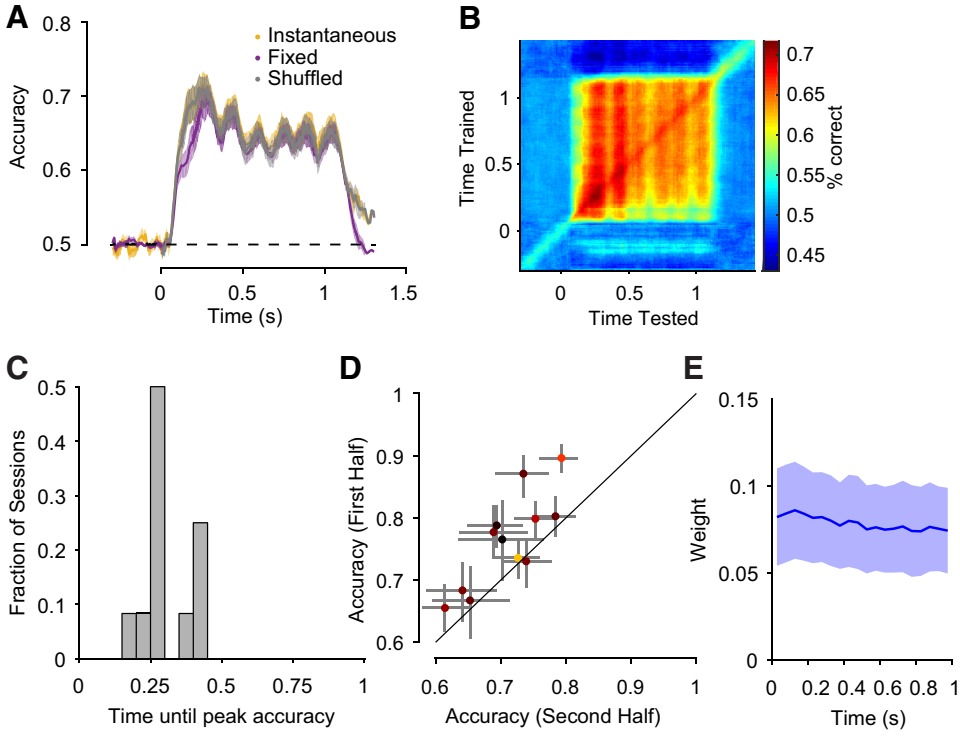

Fig. 6. Temporal sensitivity and stability of the optimal decoder. *A*: accuracy of direction decoding of the instantaneous decoder plotted over time (yellow) and comparison with accuracy of a fixed decoder (purple) and a correlation-blind instantaneous decoder (gray) averaged across all sessions. *B*: stability of the encoding of direction was determined by training the classifier at one point in time (*y*-axis) and testing at a second point in time (*x*-axis). Decoding accuracy is indicated by the color at each *x*,*y* coordinate. Vertical bands during motion correspond to the motion pulses and confirm that the weights were stable for different training epochs. *C*: fraction of sessions that reach their peak decoding performance at the time bin on the *x*-axis. Mean peak performance is at 285 ms. *D*: comparison of decoding accuracy of spikes from the first half of motion viewing with accuracy of the second half. Color of each point indicates accuracy as shown by color map in *B*. *E*: average temporal weights from the full temporal decoder (see METHODS).

simulated data. Figure 7*A* shows the shared variability model for our simulations. The population response on any trial is an *N*-dimensional vector, where *N* is the number of neurons. The firing rate of each neuron is a weighted combination of two latent variables: stimulus and anesthesia. Regardless of the number of neurons in the population, only mixtures of these two latent variables (and independent Poisson noise) contribute to their responses. Thus different values of the stimulus will drive the entire population along a one-dimensional vector embedded in the high-dimensional population space. We refer to this vector as the "stimulus axis." Similarly, the level of anesthesia drives the population response along a one-dimensional "anesthesia axis." As a result of this formulation, we can add noise (variability) to the latent value of the stimulus or anesthesia independently. However, because of the quadratic nonlinearity, $f(x)$, anesthesia can both add and act as a gain on the stimulus-driven response of a neuron. Importantly, only noise along the stimulus axis limits performance of the decoder. Noise added to the stimulus latent (i.e., adding variability along the stimulus axis) will be indistinguishable from the stimulus itself. Therefore, additional variability along the stimulus axis cannot be accounted for by the optimal linear decoder and will impair its performance (Kanitscheider et al. 2015; Moreno-Bote et al. 2014). This has the additional effect that the decoder can ignore this correlation structure because it cannot be separated from signal (Pitkow et al. 2015).

To illustrate this point, we visualized how variability along the stimulus and anesthesia axes affects the optimal decoding weights for a population of only two neurons. Figure 7*B* shows an example case where variability induced by anesthesia changes the optimal decoder. The joint responses for two neurons are shown for two stimuli and two levels of anesthesia. In this example, *neuron 2* is selective for the stimulus, whereas *neuron 1* is not (the stimulus axis is parallel to the firing rate axis for *neuron 2*). Without anesthesia, the optimal decoder would weigh *neuron 2* and ignore *neuron 1*. The addition of anesthesia [$f(a)$] induces variability that drives both *neuron 1* and *neuron 2* to fire together, which increases the total variability and adds positive noise correlation. Because this variability is independent of the stimulus axis and drives both neurons, the optimal decoding weights can account for its influence on *neuron 2* by looking at *neuron 1*. Thus the optimal decoder now includes a weight for *neuron 1*, even though alone that neuron carries no information about the stimulus. In this way anesthesia-induced covariance can shift the optimal decoder. In contrast, if anesthesia (or any other source of variability) is orthogonal to the stimulus axis (Fig. 7*C*), the optimal decoder does not shift because there is no way to discriminate anesthesia effects from the stimulus.

We explored the effects of these two sources of variability for a population of 20 simulated neurons (population size picked to match published data sets). The purpose of this

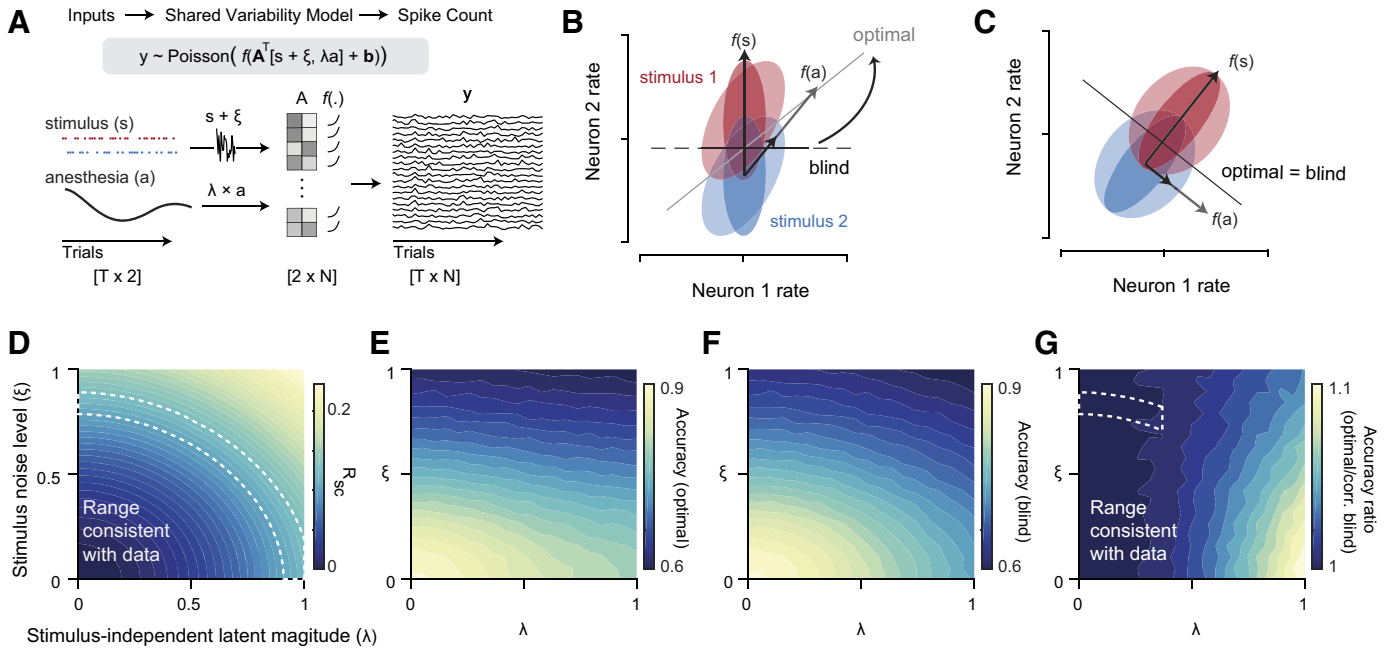

Fig. 7. Shared variability explains when correlation-blind decoding is optimal. *A*: shared variability model (after Ecker et al. 2014). Neural responses on each trial are driven by 2 latents (stimulus and anesthesia) and independent Poisson noise. Population firing rates are a weighted sum of these 2 latents passed through a static nonlinearity. Adding noise, ξ, to the stimulus latent, s, creates variability along the stimulus manifold. Scaling the anesthesia latent, a, with λ scales the magnitude of the stimulus-independent latent. *B*: intuition for the effect of anesthesia depicted for a 2-neuron case. The joint responses of 2 neurons to 2 stimuli are depicted as colored ellipses (red and blue for the 2 stimulus conditions). Only *neuron 2* is selective for the stimulus. Anesthesia drives both neurons together, increasing total variability and inducing positive correlations (transparent ellipses). The optimal decoder takes advantage of this information. *C*: same as in *B*, but both neurons are driven more for *stimulus 1* (red) and have noise correlations along the stimulus manifold, $f(s)$. Anesthesia affects both neurons the same, but its effects are orthogonal to the stimulus axis and do not impact the weights of the optimal decoder. Anesthesia has no effect on decoding performance and does not limit information. *D–G*: results of simulations from model shown in *A* for $N = 20$ neurons. *D*: average spike count correlations (R_{sc}) as a function of noise along the stimulus axis, ξ, and the magnitude of anesthesia, λ. Values of ξ and λ are normalized to range from 0 to 1. Either latent produces the same R_{sc} independently, and both latents together produce a higher R_{sc}. Our data are consistent with the contours within the white-dashed boundary. *E*: proportion correct of the optimal decoder in the same space. Accuracy depends heavily on the amount of noise along the stimulus axis and less so on anesthesia levels. *F*: proportion correct for the correlation-blind decoder. Accuracy depends similarly on the amount of stimulus noise, but depends more on the magnitude of anesthesia. *G*: ratio of decoding performance (optimal to blind) shows that when the noise along the stimulus axis dominates, the 2 decoders perform similarly.

exercise was to explore whether the simple intuition presented in Fig. 7, *B* and *C*, can explain the discrepancy in results between anesthetized and awake monkeys for realistic empirical populations. We built in an assumption that anesthesia affects all neurons to different degrees, but with the same sign, as was true of most neurons in Ecker et al. (2014). This means their spike rates move up and down together (i.e., anesthesia induces positive noise correlations). Figure 7*D* demonstrates how spike count correlations (R_{sc}) depend on additional variability along the stimulus and/or anesthesia axis. More of either source of variability increases the average R_{sc} in the population. The dashed white lines show the range of noise correlations that are consistent with our data set. Of course, because our monkeys were awake and behaving, the anesthesia axis is really any stimulus-independent shared variability. The range indicates that the observed noise correlations in our data are consistent with a range of stimulus-axis noise and anesthesia-axis noise. We then ran our decoding analysis on the simulated population responses for each of these variability levels. We trained both optimal and correlation-blind decoders on the stimulated responses. Both decoders were tested on the same trials and only differed by whether the training set was shuffled or not. Figure 7*E* shows the performance of the optimal decoder depends heavily on the amount of stimulus noise and a small amount on the level of anesthesia. In contrast, the performance of the correlation-blind decoder depends much more on the amount of anesthesia (Fig. 7*F*) because it does not account for the covariance induced by anesthesia. Taking the ratio of these two decoders shows that they only perform differently when the variability from anesthesia is greater than the stimulus-dependent variability (Fig. 7*G*). In other words, if information-limiting correlations are present at a greater level than stimulus-independent shared variability, correlation-blind decoding is effectively optimal. A similar result has been described analytically (Pitkow et al. 2015), but our simulations exhibit the exact pattern of results observed in existing data sets, reconciling the different conclusions.

DISCUSSION

We analyzed the performance of a simple read-out scheme (i.e., a population decoder) that relied on the activity of an ensemble of MT neurons to perform a direction discrimination task. We found that reading out direction from the multineuron MT response was quantitatively unaffected by whether or not the decoder was trained with the use of data that preserved or removed the correlations between neurons. This suggests that the brain mechanisms for performing direction discrimination can be relatively simple, i.e., without access to estimates of the interneuronal correlation structure.

Several previous studies have focused on whether decoding performance is limited by correlations (MorenoBote et al. 2014; Pitkow et al. 2015; Zohary et al. 1994) or improved (Franke et al. 2016; Zavitz et al. 2019; Zylberberg 2017; Zylberberg et al. 2016). Our findings do not conflict with that foundational point. The decoder is not improved by being "correlation aware," but the performance of increasingly large populations would be bounded if the population covariance contained information-limiting structure (Moreno-Bote et al. 2014). In contrast, the question we addressed in the present study is whether there was extra information to be gleaned in the joint responses of neurons. Although we did not find evidence for this in our analyses, it is entirely possible that sensitivity of the correlation structure could be useful in other tasks or decoding exercises.

In addition, we found that a population decoder with simple temporal properties could also approximate both optimal decoding and the performance of the observers. Moderate temporal integration was sufficient to account for the perceptual sensitivity of the monkeys, even though the activity of single units was not. This core finding thus generalizes earlier work that had related motion discrimination performance to the activity of single neurons (Britten et al. 1992; Parker and Newsome 1998). That prior work generally found that the activity of individual neurons was close to perceptual sensitivity, although it considered only well-targeted neurons for whom the stimulus was optimized (Cohen and Newsome 2009). In contrast, we considered larger samples of neurons, all of which were driven selectively, but often suboptimally, by the stimulus (because, of course, a single stimulus could not be perfectly suited to the selectivities of all individual neurons).

We have previously reported a decrease in direction selectivity throughout motion viewing (Yates et al. 2017), and our present decoding analyses confirm a corresponding population-level increase in sensitivity early in motion viewing, which may be boosted by a variance quenching immediately following the stimulus onset (Churchland et al. 2010). Thus even a simple (i.e., temporally stationary) decoder exhibited better performance early in trials compared with late, but this was inherited simply from the time-varying sensitivity of the neurons. Although our observation of increased sensitivity early in motion viewing may provide additional motivation for decision processes that preferentially weigh early parts of the stimulus (e.g., Drugowitsch et al. 2012; Mazurek et al. 2003), we found that the accuracy of optimal temporal integration was only modestly different from that of a simpler, temporally stationary decoder. Thus it seems reasonable to conclude that the brain can implement temporally flat decoding, because it is not far from optimal in this task. This temporally flat scheme still exhibits the time-varying signature of the MT responses evident in its output, at the level of psychophysical behavior.

Prior work on MT during direction discrimination has famously focused on the relation between MT responses and choices (Britten et al. 1996; Cohen and Newsome 2009; Haefner et al. 2013; Shadlen et al. 1996). Although an analogous set of decoding exercises similar to what we have performed for stimulus direction could be performed to "decode" choice, there are a number of reasons why we have eschewed that exercise in the present study. First, our primary goal was to assess how simple decoders might approximate optimal performance, a concept that is not applicable to the prediction of choice from neural responses. Second, since the publication of the original reports of choice-correlated fluctuations of MT response, a number of alternative interpretations of this relation have been put forth (Bondy et al. 2018; Lange and Haefner 2017; Nienborg and Cumming 2009). Likewise, the mathematical frameworks for relating the activity of neural ensembles to choices have become increasingly sophisticated (Panzeri et al. 2017). Both of these issues point to the use of a variety of distinct analyses that are not as simple as linear readout of the responses of MT, and which even question the assumption that the choice-correlated components of MT responses are reflec-

tions of a feedforward read-out scheme. For these reasons, the value of multineuron recordings for understanding the relation between MT and choices will need to be addressed in a separate report.

In contrast to our findings, prior work in anesthetized monkeys has found that knowledge of the correlation structure was necessary to optimally decode orientation from V1 (Graf et al. 2011) and direction from MT (Chaplin et al. 2018). In the present work, we put forth one explanation for this difference in results, demonstrating that shared covariance can render correlation-blind decoding effectively optimal. Specifically, we show that correlation-blind decoding matches the performance of an optimal decoder when information-limiting correlations are substantial. If, on the other hand, anesthesia induces covariance that is not information limiting, an optimal (correlation aware) decoder can outperform a simpler decoder by discounting these effects.

There are of course limitations to our approach. Previous studies have modeled the effects of anesthesia explicitly with shared gain (Ecker et al. 2014; Goris et al. 2014), we instead used a latent variable model. This offered the advantage of explicitly modeling the dimensionality of the stimulus encoding, but additive and multiplicative effects are entangled in this formulation. Although the intuitions from our model are likely to be general, more detailed characterizations of the effects of anesthesia will surely alter matters quantitatively. Future studies using recent methods to recover such latent variables from data will undoubtedly be useful for addressing questions about how neural variability contributes to perceptual decisions (Zhao and Park 2017).

Taken together, these results apply richer multineuron data to bolster the theoretical framework for reading out direction from MT that was developed in the context of single-neuron recordings. The seminal preceding studies were critical for developing a quantitative framework and considering the impacts of statistical structure and read-out schemes. Our population decoding exercises reinforce this framework by considering simultaneous measurements of multiple neurons and confirm the basic idea that simple readout of MT for estimating coarse direction is quantitatively near optimal and accounts for various aspects of psychophysical sensitivity. It remains to be seen if simple schemes for readout generalize to other stimulus features and tasks, even within MT (Purushothaman and Bradley 2005).

GRANTS

This research was supported by a Howard Hughes Medical Institute International Student Research Fellowship (to L. N. Katz), a McKnight Foundation grant (to J. W. Pillow), National Institutes of Health (NIH) Grant R01EY017366 (to J. W. Pillow and A. C. Huk), and NIH Training Grants T32DA018926 and T32EY021462. J. L. Yates is an Open Philanthropy fellow of the Life Sciences Research Foundation.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

J.L.Y., L.N.K., J.W.P., and A.C.H. conceived and designed research; J.L.Y., L.N.K., and A.J.L. performed experiments; J.L.Y. analyzed data; J.L.Y., L.N.K., J.W.P., and A.C.H. interpreted results of experiments; J.L.Y. prepared figures; J.L.Y. drafted manuscript; J.L.Y., L.N.K., A.J.L., J.W.P., and A.C.H. edited and revised manuscript; J.L.Y., L.N.K., A.J.L., J.W.P., and A.C.H. approved final version of manuscript.

REFERENCES

Albright TD. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol* 52: 1106–1130, 1984. doi:10.1152/jn.1984.52.6.1106.

Bair W, Zohary E, Newsome WT. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci* 21: 1676–1697, 2001. doi:10.1523/JNEUROSCI.21-05-01676.2001.

Berens P, Ecker AS, Cotton RJ, Ma WJ, Bethge M, Tolias AS. A fast and simple population code for orientation in primate V1. *J Neurosci* 32: 10618–10626, 2012. doi:10.1523/JNEUROSCI.1335-12.2012.

Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

Bondy AG, Haefner RM, Cumming BG. Feedback determines the structure of correlated variability in primary visual cortex. *Nat Neurosci* 21: 598–606, 2018. doi:10.1038/s41593-018-0089-1.

Brainard DH. The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997. doi:10.1163/156856897X00357.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13: 87–100, 1996. doi:10.1017/S095252380000715X.

Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992. doi:10.1523/JNEUROSCI.12-12-04745.1992.

Britten KH, Shadlen MN, Newsome WT, Movshon JA. Responses of neurons in macaque MT to stochastic motion signals. *Vis Neurosci* 10: 1157–1169, 1993. doi:10.1017/S0952523800010269.

Chaplin TA, Hagan MA, Allitt BJ, Lui LL. Neuronal correlations in MT and MST impair population decoding of opposite directions of random dot motion. *eNeuro* 5: ENEURO.0336-18.2018, 2018. doi:10.1523/ENEURO.0336-18.2018.

Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13: 369–378, 2010. doi:10.1038/nn.2501.

Cohen MR, Newsome WT. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29: 6635–6648, 2009. doi:10.1523/JNEUROSCI.5179-08.2009.

Drugowitsch J, Moreno-Bote R, Churchland AK, Shadlen MN, Pouget A. The cost of accumulating evidence in perceptual decision making. *J Neurosci* 32: 3612–3628, 2012. doi:10.1523/JNEUROSCI.4010-11.2012.

Eastman KM, Huk AC. PLDAPS: a hardware architecture and software toolbox for neurophysiology requiring complex visual stimuli and online behavioral control. *Front Neuroinform* 6: 1, 2012. doi:10.3389/fninf.2012.00001.

Ecker AS, Berens P, Cotton RJ, Subramaniyan M, Denfield GH, Cadwell CR, Smirnakis SM, Bethge M, Tolias AS. State dependence of noise correlations in macaque primary visual cortex. *Neuron* 82: 235–248, 2014. doi:10.1016/j.neuron.2014.02.006.

Efron B, Tibshirani R. Improvements on cross-validation: the .632+ bootstrap method. *J Am Stat Assoc* 92: 548–560, 1997. doi:10.2307/2965703.

Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, da Silveira RA. Structures of neural correlation and how they favor coding. *Neuron* 89: 409–422, 2016. doi:10.1016/j.neuron.2015.12.037.

Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 33: 1–22, 2010. doi:10.18637/jss.v033.i01.

Georgopoulos AP, Schwartz AB, Kettner RE. Neuronal population coding of movement direction. *Science* 233: 1416–1419, 1986. doi:10.1126/science.3749885.

Goris RL, Movshon JA, Simoncelli EP. Partitioning neuronal variability. *Nat Neurosci* 17: 858–865, 2014. doi:10.1038/nn.3711.

Goris RL, Ziemba CM, Movshon JA, Simoncelli EP. Slow gain fluctuations limit benefits of temporal integration in visual cortex. *J Vis* 18: 8, 2018. doi:10.1167/18.8.8.

Graf AB, Kohn A, Jazayeri M, Movshon JA. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14: 239–245, 2011. doi:10.1038/nn.2733.

Haefner RM, Gerwinn S, Macke JH, Bethge M. Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci* 16: 235–242, 2013. doi:10.1038/nn.3309.

Hill DN, Mehta SB, Kleinfeld D. Quality metrics to accompany spike sorting of extracellular signals. *J Neurosci* 31: 8699–8705, 2011. doi:10.1523/JNEUROSCI.0971-11.2011.

Huk AC, Shadlen MN. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci* 25: 10420–10436, 2005. doi:10.1523/JNEUROSCI.4684-04.2005.

Kanitscheider I, Coen-Cagli R, Kohn A, Pouget A. Measuring Fisher information accurately in correlated neural populations *PLoS Comput Biol* 11: e1004218, 2015. [Erratum in *PLoS Comput Biol* 12: e1004945, 2016.] doi:10.1371/journal.pcbi.1004218.

Katz LN, Yates JL, Pillow JW, Huk AC. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature* 535: 285–288, 2016. doi:10.1038/nature18617.

Kelly RC, Smith MA, Samonds JM, Kohn A, Bonds AB, Movshon JA, Lee TS. Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J Neurosci* 27: 261–264, 2007. doi:10.1523/JNEUROSCI.4906-06.2007.

Kiani R, Hanks TD, Shadlen MN. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci* 28: 3017–3029, 2008. doi:10.1523/JNEUROSCI.4761-07.2008.

Lange RD, Haefner RM. Characterizing and interpreting the influence of internal variables on sensory activity. *Curr Opin Neurobiol* 46: 84–89, 2017. doi:10.1016/j.conb.2017.07.006.

Latham PE, Nirenberg S. Synergy, redundancy, and independence in population codes, revisited. *J Neurosci* 25: 5195–5206, 2005. doi:10.1523/JNEUROSCI.5319-04.2005.

Levi AJ, Yates JL, Huk AC, Katz LN. Strategic and dynamic temporal weighting for perceptual decisions in humans and macaques. *eNeuro* 5: ENEURO.0169-18.2018, 2018. doi:10.1523/ENEURO.0169-18.2018.

Mazurek ME, Roitman JD, Ditterich J, Shadlen MN. A role for neural integrators in perceptual decision making. *Cereb Cortex* 13: 1257–1269, 2003. doi:10.1093/cercor/bhg097.

Mineault PJ, Khawaja FA, Butts DA, Pack CC. Hierarchical processing of complex motion along the primate dorsal visual pathway. *Proc Natl Acad Sci USA* 109: E972–E980, 2012. doi:10.1073/pnas.1115685109.

Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. *Nat Neurosci* 17: 1410–1417, 2014. doi:10.1038/nn.3807.

Nienborg H, Cumming BG. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459: 89–92, 2009. doi:10.1038/nature07821.

Osborne LC, Bialek W, Lisberger SG. Time course of information about motion direction in visual area MT of macaque monkeys. *J Neurosci* 24: 3210–3222, 2004. doi:10.1523/JNEUROSCI.5305-03.2004.

Panzeri S, Harvey CD, Piasini E, Latham PE, Fellin T. Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. *Neuron* 93: 491–507, 2017. doi:10.1016/j.neuron.2016.12.036.

Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21: 227–277, 1998. doi:10.1146/annurev.neuro.21.1.227.

Pillow JW, Shlens J, Chichilnisky EJ, Simoncelli EP. A model-based spike sorting algorithm for removing correlation artifacts in multi-neuron recordings. *PLoS One* 8: e62123, 2013. doi:10.1371/journal.pone.0062123.

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454: 995–999, 2008. doi:10.1038/nature07140.

Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A. How can single sensory neurons predict behavior? *Neuron* 87: 411–423, 2015. doi:10.1016/j.neuron.2015.06.033.

Purushothaman G, Bradley DC. Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8: 99–106, 2005. doi:10.1038/nn1373.

Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16: 1486–1510, 1996. doi:10.1523/JNEUROSCI.16-04-01486.1996.

Uka T, DeAngelis GC. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42: 297–310, 2004. doi:10.1016/S0896-6273(04)00186-2.

Yates JL, Park IM, Katz LN, Pillow JW, Huk AC. Functional dissection of signal and noise in MT and LIP during decision-making. *Nat Neurosci* 20: 1285–1292, 2017. doi:10.1038/nn.4611.

Yousef WA, Wagner RF, Loew MH. Comparison of non-parametric methods for assessing classifier performance in terms of ROC parameters. *Applied Imagery Pattern Recognition Workshop (AIPR'04)*. 33: 190–195, 2004. doi:10.1109/AIPR.2004.18.

Zavitz E, Yu HH, Rosa MG, Price NS. Correlated variability in the neurons with the strongest tuning improves direction coding. *Cereb Cortex* 29: 615–626, 2019. doi:10.1093/cercor/bhx344.

Zhao Y, Park IM. Variational latent gaussian process for recovering single-trial dynamics from population spike trains. *Neural Comput* 29: 1293–1316, 2017. doi:10.1162/NECO_a_00953.

Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370: 140–143, 1994. doi:10.1038/370140a0.

Zylberberg J. Untuned but not irrelevant: the role of untuned neurons in sensory information coding (Preprint). *bioRxiv* 134379, 2017.

Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. Direction-selective circuits shape noise to ensure a precise population code. *Neuron* 89: 369–383, 2016. doi:10.1016/j.neuron.2015.11.019.